Society for the Study
of Economic Inequality

Working Paper Series

# Nonparametric analysis of heterogeneous multidimensional fairness

**Domenico Moramarco**

**Bram De Rock**

ECINEQ 2022 621

ECINEQ 2022 621
2022 September
www.ecineq.org

# Nonparametric analysis of heterogeneous multidimensional fairness

**Domenico Moramarco**
*ECARES, Université Libre de Bruxelles*
**Bram De Rock**
*ECARES, Université Libre de Bruxelles*

**Abstract**

*The paper proposes a framework to assess fairness in multidimensional distributions while respecting individual preferences. We characterize a simple measure - Equivalent Advantage - that captures the distance from the current outcome to the potentially individual specific norm outcome. We introduce a nonparametric approach to partially identify our measure via set identification of individual indifference curves. Our methodology is illustrated by analyzing multidimensional fairness in Belgium using the MEqIn database. Despite the set identification, we show that our analysis of the Equivalent Advantage distribution allows for interesting insights on multidimensional inequality, poverty and opportunity distribution.*

Keyword: fairness, multidimensional inequality, poverty, equality of opportunity, Equivalent Advantage, individual preferences, nonparametric analysis

JEL Cassification: D30, D63, I14, I32

# Nonparametric analysis of heterogeneous multidimensional fairness*

Bram De Rock[†] and Domenico Moramarco[‡]

September 2022

**Abstract**

The paper proposes a framework to assess fairness in multidimensional distributions while respecting individual preferences. We characterize a simple measure - Equivalent Advantage - that captures the distance from the current outcome to the potentially individual specific norm outcome. We introduce a nonparametric approach to partially identify our measure via set identification of individual indifference curves. Our methodology is illustrated by analyzing multidimensional fairness in Belgium using the MEqIn database. Despite the set identification, we show that our analysis of the Equivalent Advantage distribution allows for interesting insights on multidimensional inequality, poverty and opportunity distribution.

**Keywords**: fairness, multidimensional inequality, poverty, equality of opportunity, Equivalent Advantage, individual preferences, nonparametric analysis

*This paper makes use of the MEqIn dataset, collected by a team of researchers from Université catholique de Louvain, KU Leuven, Université libre de Bruxelles, and University of Antwerp. The collection of the MEqIn data was enabled by the financial support of the Belgian Science Policy Office (BELSPO) through grant BR/121/A5/MEQIN (BRAIN MEqIn).

[†]ECARES, Université Libre de Bruxelles and Department of Economics, University of Leuven (KU Leuven). Av. F. Roosevelt 50, B-1050 Bruxelles, Belgium. E-mail: bram.de.rock@ulb.be. Bram De Rock gratefully acknowledges FWO and FNRS for their financial support.

[‡]ECARES, Université Libre de Bruxelles. Av. F. Roosevelt 50, B-1050 Bruxelles, Belgium. E-mail: domenico.moramarco@ulb.be. Domenico Moramarco gratefully acknowledges FNRS for their financial support.

**JEL codes**: D30, D63, I14, I32.

# 1 Introduction

Inequality and fairness in opportunities have always been of prominent interest in and outside academia. The majority of the recent research on distributive analysis continues to focus on income and wealth inequality (see Piketty & Saez, 2006; Piketty *et al.*, 2006; Alvaredo *et al.*, 2017; Piketty *et al.*, 2019, for some examples). At the same time, several prominent economists have been interested in the distribution of other relevant outcomes like health (Banerjee *et al.*, 2004; Deaton, 2008) or income opportunities (Chetty *et al.*, 2014).

Although it is intuitive that one should ideally acknowledge the several dimensions simultaneously, this poses an empirical challenge on how to construct a measure that is easy to compute and readily to interpret. We tackle this issue in our paper.

**Multidimensional inequality.** The multidimensionality of individual well-being is of course not new to the economic literature, which pleads for a definition of social welfare indicators that go beyond GDP (see Stiglitz *et al.*, 2009; Fleurbaey & Blanchet, 2013, for a rich argumentation). This multidimensionality nature has also been affecting the analysis of social justice, and the literature has been, for a long time, investigating the possibility of measuring distributive phenomena like inequality and poverty from a multidimensional point of view (see Maasoumi, 1986; Atkinson & Bourguignon, 1987; Tsui, 1995, 1999, for early contributions).

After more than two decades of contributions on multidimensional inequality, several methodologies have been proposed (see Andreoli & Zoli, 2020, for a recent survey). Two are clearly the most prominent. The first approach measures inequality in each of the dimension to then define a weighted average of the various inequality indices. The second approach starts by a multidimensional measure of individual well-being and proceeds with measuring inequality in individual welfare. While the first approach is the one followed by, for instance, the World Bank in computing the Human Development Index (HDI) (UN, 2018), economists tend to prefer the second one (Bosmans *et al.*, 2009).

Independently of the approach, to measure multidimensional inequality one needs to choose how to weight each dimension. This decision is arbitrary in principle but crucial for the validity of the results and the consequent policy recommendations. In this sense, one of the main challenges in constructing a measure of multidimensional fairness is the definition of the 'right' system of weights (see Decancq & Lugo, 2012, 2013; Fleurbaey & Blanchet, 2013, for a discussion). An attractive alternative is to allow for heterogeneity in the weights by using individual preferences to determine them (see Piacquadio, 2017; Decancq *et al.*, 2019; Maniquet & Neumann, 2021; de Frahan & Maniquet, 2021; Fleurbaey & Van der Linden, 2021, for some recent applications).

**Theoretical contribution.** In this paper we measure social injustice by allowing for multiple dimensions of well-being and at the same time by respecting individual preferences. To account for the multidimensionality, we compare the empirical (or observed) distribution, with a normative (or fair) one. As underlined by Cowell *et al.* (2013), this is not an unusual approach, since most, if not all, measures of inequality or poverty can be interpreted as the distance between the current income distribution and a fair one.[1] Recently, Almås *et al.* (2011) extended this approach to more flexible definitions of the norm distribution. In this paper, we extend this methodology to a multidimensional setting, to analyze not only multidimensional equality but also poverty, equality of opportunity or more complex definitions of distributive justice. This also includes, among other things, that the normative distribution can be individual specific.

To account for the individual preferences, we assume to have data allowing to approximate the individual preferences related to the several dimensions of well-being. To obtain robust and reliable estimates of these preferences, we introduce an easy to implement nonparametric methodology to partially identify individual indifference curves. Our approach is in line with Manski (2003) and Blundell *et al.* (2008), and their applica-

[1] In the case of inequality, these are respectively the Lorenz curve and the first diagonal. For poverty, the current distribution can be approximated by the TIP curve (Jenkins & Lambert, 1997), where the fair one is a horizontal line at the poverty threshold.

tion in the revealed preference analysis (see Crawford & De Rock, 2014, for a survey), since we abstain from parametrizing preferences and therefore only obtain set identification. Differently form the standard revealed preference approach, that uses expenditure data, we make use of elicited willingness-to-pay information. This type of stated preferences are easy to collect and, in a multidimensional setting with outcomes like health, they are sometimes the only possible data source (see Decancq & Nys, 2021, for a related recent contribution).

We provide axiomatic justification for a simple and intuitive measure of individual (dis)advantage with respect to the socially acceptable outcome: the Equivalent Advantage. Individuals are advantaged whenever they prefer the current outcome to the fair one and we measure (dis)advantages as the distance between the fair outcome and the individual indifference curve at the current outcome. The Equivalent Advantage can be interpreted as the amount of income we should add/remove to the fair outcome to make it indifferent to the current one. Our set-identified individual indifference curves, allow us to estimate upper and lower bounds of Equivalent Advantage, which can be considered the most optimistic and pessimistic assessment of an individual situation.

**Empirical application.** To implement our new methodology we make use of the Belgian MEqIn data. For simplicity we restrict our attention to only consumption and health. To demonstrate the versatility of our approach, we consider three fair distributions. For the first distribution, which defines social justice as outcome equality, the fair outcome of each individual is the combination of average consumption and average health in the population. The second distribution defines fairness as absence of relative poverty so that the current outcome of each individual is fair if she prefers it to the one containing 60% of the median consumption and 60% of the median level of health. The third fair distribution implements the opportunity egalitarian paradigm (see Ramos & Van de Gaer, 2016, for a survey), imposing the fair outcome of an individual to be the average outcome among individuals that exert the same level of effort and make the same choices. This latter distribution also illustrates that the norm distribution should not be the same for all individuals.

Our analysis highlights that a quarter of the Belgian population is multidimensionally poor, and about half has an outcome that they consider worse than the the average one in the population, or in the group of individuals with similar characteristics. However, the multidimensional nature of our approach does not simply lead to the appearance of new disadvantaged individuals. As we will discuss it provides instead a different picture of distributive justice, which can be hardly replicated when disregarding individual preferences.

In this different picture, we still observe a clear gender gap, with women being more likely to suffer from welfare losses caused by the existing social injustices. In line with the life-cycle patterns, individuals in adult age (in particular between 40 and 49 years old) tend to be better off, and the risk and intensity of disadvantage is significantly higher among the old people. Not surprisingly, employment status is highly correlated with Equivalent Advantage, and unemployment puts people at high risks of welfare losses. Although these results are stable within each of the three main regions of Belgium (Brussels, Flanders and Wallonia), we also find that the economic development of Brussels and Flanders has induced less preferable distributions of Equivalent Advantages, compared to Wallonia.

**Outline.** The rest of the paper is organized as follows. Section 2 introduces a general theoretical framework and defines our Equivalent Advantage measure. Section 3 introduces the data and the three normative distributions we consider. Section 4 introduces our nonparametric methodology to set identify individual indifference curves. Section 5 discusses our obtained empirical results. Section 6 concludes the paper. Additional details on results and data are reported in the online Appendix.

# 2 Equivalent Advantage

We start by introducing some notation and discussing the basic axioms that a measure of individual (dis)advantage should satisfy. Subsequently we introduce our measure of Equivalent Advantage and show it is the only measure satisfying our axioms. We end by briefly discussing how to compare distributions of Equivalent Advantage.

## 2.1 Preliminaries

We observe the $K$-dimensional outcome $x^i \in \mathbb{R}_+^K$ of each individual $i$ in a population $\mathcal{N}$ of dimension $n$. Individuals are assumed to have rational, continuous, convex and monotone preferences $R^i$, which are strictly monotone in a numéraire outcome. For simplicity we call it income and we let it be the first of the $K$ dimensions.[2] Let $\mathfrak{R}$ be the preference domain; for any two outcomes $x, x' \in \mathbb{R}_+^K$ and $R^i \in \mathfrak{R}$, $x \; P^i \; x'$ denotes strict preference of $x$ over $x'$, $x \; R^i \; x'$ is its weak version and $x \; I^i \; x'$ denotes indifference. We do not assume to observe the whole map of indifference curves generated by individual preferences. Rather, for each individual $i$, we assume to observe only the set of outcomes she considers indifferent to $x^i$, denoted $I(x^i, R^i)$. Hence, our assumption on individual preferences must only hold locally, for the indifference curve at the current outcome.

We call $X = \left\{x^i \in \mathbb{R}_+^K : i \in \mathcal{N}\right\}$ the outcome distribution and $\mathcal{R} = \{R^i \subset \mathbb{R}_+^K \times \mathbb{R}_+^K : i \in \mathcal{N}\}$ the preference distribution. We assume that there exists a fair/norm distribution $Y = \left\{y^i \in \mathbb{R}_{++}^K : i \in \mathcal{N}\right\}$ where $y^i$ is the fair/norm outcome of individual $i \in \mathcal{N}$. At this stage, we take $Y$ as given; we only impose it to depend on the outcome distribution we would like to assess, and to be composed of strictly positive outcomes. A simple example of a fair distribution is the egalitarian one in which $y^i = \left(1/n \sum_{j \in \mathcal{N}} x_1^j, ..., 1/n \sum_{j \in \mathcal{N}} x_K^j\right)$ for all $i \in \mathcal{N}$. In this example, the fair outcome of each individual is the average one, where $x_k^j$ denotes the $k$th entry of the vector $x^j$. More in general, $Y$ can assign different outcomes to different individuals depending on the distributive principle(s); specific examples will be discussed in the empirical application.

The triplet $(X, \mathcal{R}, Y)$ describes the population of interest and any individual $i \in \mathcal{N}$ is represented by $(x^i, R^i, y^i) \in (X, \mathcal{R}, Y)$. It is straightforward to state that if the outcome $x^i$ coincides with its fair counterpart $y^i$, then the individual is in a fair (or socially acceptable) situation. In this paper, we strengthen this concept by claiming that individual evaluations should be respected. In particular, if one considers the current and the fair outcomes equally good - $x^i \; I^i \; y^i$ - then we should also conclude that her situation is fair. Our approach is in line with Dworkin (1981a,b) who proposes to

[2] For convenience, the graphs will represent income on the vertical axis.

hold individuals responsible for their preferences when analyzing fairness; Fleurbaey & Tadenuma (2014) call this the consumer sovereignty principle.

## 2.2 Axioms

We are interested in measuring how each individual situation deviates from the fair one. That is, we want to assess the individual's unfair advantage or disadvantage via a function $a : \mathbb{R}_{+}^{K} \times \mathfrak{R} \times \mathbb{R}_{++}^{K} \rightarrow \mathbb{R}$, such that $a\left(x^{i}, R^{i}, y^{i}\right)$ is the advantage/disadvantage of an individual with outcome $x^{i}$ and preferences $R^{i}$, with respect to a fair outcome $y^{i}$. A simple way of measuring unfair (dis)advantages consists of looking at the difference between the utilities at current and fair outcomes. However, this measure would be sensitive to the specific utility function implemented, so such comparisons are not robust (see Fleurbaey & Zuber, 2021, for a recent contribution on this issue). More important, we assumed limited knowledge of individual preferences - we observe only one indifference curve - which does not allow us to adequately define a utility function.

Our alternative measure for (dis)advantage should naturally satisfy some basic axioms. The first is a normalization one, which simply states that our measure is zero whenever the observed and fair outcome coincide.

**Axiom 1.** *Normalization (Norm) - For all* $(x^{i}, R^{i}, y^{i}) \in (X, \mathcal{R}, Y)$*, if* $x^{i} = y^{i}$*, then* $a\left(x^{i}, R^{i}, y^{i}\right) = 0$.

The following axiom characterizes our principle of respect for preferences. In words, we impose that two equivalent outcomes should induce the same (dis)advantage.

**Axiom 2.** *Indifference (Ind) - For all* $\left(x^{i}, R^{i}, y^{i}\right), \left(x^{j}, R^{j}, y^{j}\right) \in (X, \mathcal{R}, Y)$*, if* $y^{i} = y^{j}$ *and* $I(x^{i}, R^{i}) = I(x^{j}, R^{j})$*, then* $a\left(x^{i}, R^{i}, y^{i}\right) = a\left(x^{j}, R^{j}, y^{j}\right)$.

This axiom restricts (dis)advantages to depend on the fair outcome and the indifference contour set at the current outcome. In this sense, this introduces an anonymity principle by treating individuals $i$ and $j$ the same when they are identical with respect to these two objects. At the same time, by focusing on a single individual, we have the desired requirement that indifferent outcomes are associated with the same advantage.

The last axiom considers a situation in which individuals only differ in one dimension from the fair outcomes, which in turn leads to a natural comparison. Let us denote with $\mathbf{0}_{K-1}$ the $K-1$-dimensional vector of zeroes.

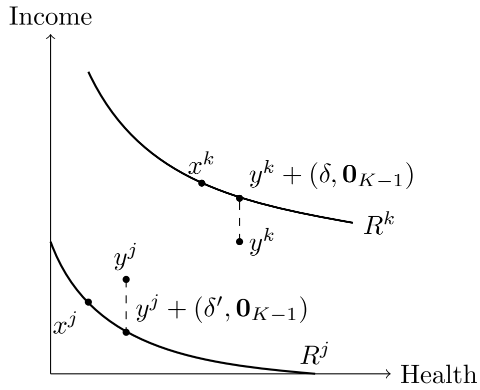
**Axiom 3.** *Fairness Comparison (FC) - Let* $(x^i, R^i, y^i), (x^j, R^j, y^j) \in (X, \mathcal{R}, Y)$*. Assume there exists* $\delta^i, \delta^j \in \mathbb{R}$ *such that* $x^i - y^i = (\delta^i, \mathbf{0}_{K-1})$ *and* $x^j - y^j = (\delta^j, \mathbf{0}_{K-1})$*. Then,* $\delta^i \geq \delta^j$ *implies* $a(x^i, R^i, y^i) \geq a(x^j, R^j, y^j)$*.*

Note that this axiom nests the unidimensional analysis of (income) inequality and poverty. More specifically, it can be framed as situations in which the social evaluator is concerned with the distribution of income but not of other outcomes, so that $x_j^i = y_j^i$ for all $j = 2, .., K$ and $i \in \mathcal{N}$. FC requires that in such situations, we can behave as if the other dimensions were irrelevant and implement the standard unidimensional criterion.

Two concluding remarks are in order. First, FC abstracts from weighting the different dimensions. That is, when comparing $x^i$ to $x^j$, the exact value of the other dimensions is irrelevant. So for instance, comparing income of healthy individuals happens in the same way as comparing income of unhealthy individuals. Whether a minimum level of health (or any other outcome) is sufficient for an individual to always be considered advantaged (or never disadvantaged) or if trade offs between dimensions should be considered, is a normative requirement that should enter the definition of $Y$, rather than the individual measure of advantage.

Second, FC is a natural requirement to impose when there is an obvious or desirable numéraire. As a consequence, the advantage measurement and the ranking between individuals, will be sensitive to the choice of the numéraire. At the same time, given the multidimensional nature of our measure, using income as a numéraire helps the interpretation and comparison of (dis)advantages. Related to this, an alternative version of FC can be formulated by considering a situation in which all fair outcomes are proportional to the current one (see Decancq *et al.*, 2019). For the sake of a clear exposition, we abstract from this proportional alternative or considering a different numéraire.

Figure 1: The Equivalent Advantage in the income-health space.




**Description** The EA is computed as the income distance between the fair outcome and the individual indifference curve. In this example, $EA_k = \delta > 0 > \delta' = EA_j$

## 2.3 Equivalent Advantage

In this paper we focus on a function that evaluates individual advantages in terms of the numéraire.

**Definition 1.** *The Equivalent Advantage (EA) is a function* $a : \mathbb{R}_+^K \times \mathfrak{R} \times \mathbb{R}_{++}^K \to \mathbb{R}$ *such that* $a(x, R, y) = \delta \iff z - y = (\delta, \mathbf{0}_{K-1})$ *and* $zIx$.

To simplify the notation, for all $(x^i, R^i, y^i) \in (X, \mathcal{R}, Y)$, let $EA_i \equiv a(x^i, R^i, y^i)$. This definition states that Equivalent Advantage assesses unfairness by using the income we should add/remove to the fair outcome to make it indifferent to the current one. To put it differently, $EA_i > 0$ means that, for agent $i$, the current outcome is equivalent to enjoying an income surplus with respect to the fair outcome $y^i$. Reversely, a negative EA means that the current situation is equivalent to renouncing to part of the fair income. Geometrically, the EA is measured along the segment, parallel to the numéraire dimension, that connects the individual fair outcome with the indifference contour set at the current outcome (see Figure 1).

Note that strict monotonicity of preferences in the numéraire outcome ensures $EA$ to be well defined. Given this, it is straightforward to verify that EA satisfies the three axioms we discussed above. The following theorem states that it is also the only measure (up to increasing transformations) that does so.

**Theorem 1.** *For all* $(X, \mathcal{R}, Y) \in \left(\mathbb{R}^K_+\right)^n \times (\mathfrak{R})^n \times \left(\mathbb{R}^K_{++}\right)^n$, *Equivalent Advantage is the only measure that satisfies Norm, Ind and FC.*

*Proof.* Let $a$ be a measure that satisfies the three axioms. Since preference are strictly increasing in the numéraire, for any $i \in \mathcal{N}$ there exists a unique $z^i \in I(x^i, R^i)$ such that $z^i - y^i = (\delta^i, \mathbf{0}_{K-1})$ for some $\delta^i \in \mathbb{R}$. Therefore, any $(X, \mathcal{R}, Y)$ is associated to a $(Z, \mathcal{R}, Y)$ in which the conditions of FC are binding.

Since $a$ satisfies FC, for all $i, j \in \mathcal{N}$, $\delta^i \geq \delta^j$ implies $a\left(z^i, R^i, y^i\right) \geq a\left(z^j, R^j, y^j\right)$. Therefore, in $(Z, \mathcal{R}, Y)$ we can define a linear order of individuals in terms of advantage. This order can be represented by a function $a : \mathbb{R}^K_+ \times \mathfrak{R} \times \mathbb{R}^K_{++} \to \mathbb{R}$ such that, for all $(z^i, R^i, y^i) \in (Z, \mathcal{R}, Y)$, $a\left(z^i, R^i, y^i\right) = f(\delta^i)$, where $f : \mathbb{R} \to \mathbb{R}$ is an increasing function.

By Ind, $a\left(z^i, R^i, y^i\right) = a\left(x^i, R^i, y^i\right)$, so that measuring advantages in $(X, \mathcal{R}, Y)$ or $(Z, \mathcal{R}, Y)$ is a matter of indifference.

Finally, Norm imposes $f(0) = 0$. Therefore, without loss of generality, we can set $f(\delta) = \delta$ for all $\delta \in \mathbb{R}$. The resulting measure coincides with EA from Definition 1.

□

We want to make two final remarks. First our EA measure is closely related to the equivalent income idea since we evaluate outcomes by looking at equivalent ones (see Samuelson, 1974; Samuelson & Swamy, 1974; Fleurbaey & Blanchet, 2013). Next, our measure is constructed to assess if and to what extent individuals are obtaining a fair outcome and does not aim at making inter-personal well-being comparisons. For example, an individual may result more advantaged than another individual with strictly higher current outcome. This is, only in part, due to the respect for preferences. More in general, Equivalent Advantage and individual well being are two different concepts and our approach focuses on the former to evaluate social justice (see Mahler & Ramos, 2019; Decancq *et al.*, 2017, for examples of the other approach).

## 2.4 Ranking distributions

Equivalent Advantages take positive values whenever the individual prefers his current outcome to the fair one and becomes negative in the opposite case. From a purely distributional perspective, both advantages and disadvantages are source of unfairness since they both represent deviations from the fair outcome. However, between a population in which we observe only advantages and one in which there are only disadvantages, the former is a more desirable outcome. It goes without saying that, since Equivalent Advantages depend on the fair outcome distribution, the latter should be adequately chosen for the previous comparison to be meaningful.

Interpreting Equivalent Advantages as welfare benefits, we can follow the standard approach in the distribution analysis, and use an increasing and concave aggregation of the advantages as social evaluation function. Formally, let us denote with $\mathcal{S}$ the set of all increasing and Schur-concave social evaluation functions $S : \mathbb{R}^n \to \mathbb{R}$. Call $A$ the Equivalent Advantage distribution induced by $(X, \mathcal{R}, Y)$ and $B$ the one induced by $(X', \mathcal{R}', Y')$. If $S(A) \geq S(B)$, $S \in \mathcal{S}$, then we say that $(X, \mathcal{R}, Y)$ displays more social justice.[3] The functions in $\mathcal{S}$ implement the idea that advantages are desirable and can (partially) compensate for disadvantages. At the same time, the concavity of these function expresses aversion to the degree of inequality in the distribution of Equivalent Advantages.

A given function $S \in \mathcal{S}$ provides us with a test to compare fairness in different populations and construct a complete ranking. However, the choice of the 'right' social evaluation function is arbitrary in principle and a different $S$ can invert the ranking between two Equivalent Advantage distributions. This is a well-known issue in the distributive analysis literature and can be solved with the use of the generalized Lorenz curves.

For all increasingly ordered real vectors $A = (a(x^1, R^1, y^1), ..., a(x^n, R^n, y^n))$, the generalized Lorenz curve is the graph of the function

$$gL(p, A) = \frac{1}{n} \sum_{i=1}^{k} a\left(x^i, R^i, y^i\right) \tag{1}$$

[3] The implicit assumption when comparing two populations $(X, \mathcal{R}, Y)$ and $(X', \mathcal{R}', Y')$ is that $Y$ and $Y'$ are inspired by the same fairness principle.

with respect to $p = k/n$, for all $k = 0, 1, \ldots, n$. This graphical instrument is extremely useful to compare Equivalent Advantage distributions. Indeed, as shown by Shorrocks (1983), $S(A) \geq S(B)$ for all $S \in \mathcal{S}$ if and only if $gL(p, A) \geq gL(p, B)$ for all $p \in [0, 1]$. Therefore, we have a test to compare populations, or subgroups of the same population, in a robust and unambiguous way. In words, if the generalized Lorenz curve of a distribution $A$ is never below the one of $B$ (equivalently, $A$ generalized Lorenz dominates $B$) then, independently from the way in which we allow advantages to compensate disadvantages (i.e. the degree of inequality aversion), $A$ is preferable in terms of social Equivalent Advantage. In particular, for any share of the population, the (cumulative) average Equivalent Advantage in $A$ is higher than in $B$, so that people are always less unfairly disadvantaged in the former distribution.

# 3 Data and empirical set-up

This section introduces the data and empirical set-up that we use to implement our framework. In particular, we compute EA in a consumption-health space and analyze its distribution across different groups of the population. We use the Measuring Equivalent Income (MEqIn) database based on a survey conducted in Belgium in 2016, collecting detailed information on individual consumption, time-use, health, work and other aspects of life for people above 18 years old (see Capéau *et al.*, 2020, for more details).

## 3.1 Data

We focus on a bi-dimensional outcome distribution of personal consumption and health score.[4] The former is a measure of the monthly individual's private expenditures (e.g. food, clothes, leisure activities etc.) and a share of the public household expenditures (e.g. utilities and insurance). The latter is an index from 0 to 100, where 100 represents perfect health, constructed in the survey to summarize individual answers to detailed questions about physical and psychological health.

[4]We focus on personal consumption because it's the outcome used in the MEqIn survey to elicit WTP for perfect health.

In order to avoid outliers to bias our results, we remove from the sample all observations with consumption, or health score, in the first or last percentile of the relative distributions. Moreover, we do not consider individuals that are unemployed due to still studying. This choice is aimed at focusing on people with more freedom in terms of consumption: unemployed students are more likely to depend on parents for their consumption. We further restrict the dataset to those observations with information about time-use, number of visits to doctors and willingness-to-pay (WTP) for perfect health. Our final sample is composed of 2,778 individual observations. When relevant, we use individual weights to maintain the sample representative for the Belgian population. Online Appendix A contains descriptive statistics and more details on our selected variables.

## 3.2 Fair distributions

The first step to implement the proposed framework is the characterization of a fair distribution. To illustrate the versatility and generality of our method, we consider three different fair distributions that realize, respectively, multidimensional equality, absence of poverty and (ex post) equality of opportunity. Below, the observed outcome of an individual $i$ is a pair $(c_i, h_i)$ of consumption and health score, while the pair $(c_i^\star, h_i^\star)$ denotes the norm outcome for that individual.

**Multidimensional inequality.** The first norm distribution, which we call equal outcome (EO) distribution, is obtained by setting $(c_i^\star, h_i^\star) = (\mu_c, \mu_h)$ as the fair outcome of each individual, where $\mu_c$ is the average personal consumption and $\mu_h$ is the average health score in the population. The EA with respect to this norm outcome will be a measure of the gain or loss that the individual experiences with respect to the respective averages of both dimensions in the population. The average EA can therefore be interpreted as a measure of the average welfare gain or loss caused by the existence of multidimensional inequality. An inequality index applied to the distribution of Equivalent Advantages computed under the EO norm is therefore a measure of multidimensional inequality.

**Multidimensional poverty.** We call no relative poverty (POV) the second norm distribution, defined in order to analyze multidimensional poverty. This distribution implements the idea an outcome is socially acceptable whenever it is above a poverty bundle. Hence, the fair outcome of an individual with less than the poverty bundle -$(c_i, h_i) \ll (z_c, z_h)$ - is the poverty bundle itself.[5] We set $z_c$ and $z_h$ equal to 60% of the median outcome. Note that this way of defining the POV norm implements the standard focus axiom in the construction of the fair distribution. Moreover, the (always negative) EA with respect to this norm can be interpreted as a multidimensional poverty gap.

**Equality of opportunity.** The last fair distribution implements the opportunity egalitarian approach in its ex post interpretation (see Fleurbaey, 2008, for a deep discussion); we call it opportunity egalitarian (EOp) distribution. In a nutshell, this social justice principle claims that individuals making the same choice should obtain the same outcome. To illustrate that our method also allows us to bring this principle to the data, we have to make some (strong) assumptions related to individual choice.

To define the fair consumption level, our EOp principle is the following: individuals with the same level of education, that spend the same time in productive activities and live in a household with the same size should have the same consumption. We approximate these choices using: *(i)* $Edu_i$ which represents the level of education of individual $i$ via a categorical variable with values from 1 (primary/no education) to 5 (university/PhD degree); *(ii)* $WorkTime_i$ which takes values from 1 to 5 depending on the individual's quintile position in the cumulative distribution of the amount of weekly time spent in paid work, housekeeping, child care and care for other household and non household members; *(iii)* $HhSize_i$ that is the number of persons forming the household, and takes values from 1 to 5 where the former means that the individual constitutes a one-person household while the latter value includes household composed of 5 or more persons.[6] This allows us to divide the population in 125 groups of individuals and for each

[5]We write $(c_i, h_i) \ll (z_c, z_h)$ if $c_i < z_c$ and $h_i < z_h$. Next, if, for example, $c_i > z_c$ while $h_i < z_h$ then the fair outcome is $(c_i, z_h)$.

[6]The use of a categorical version of this and other variables is aimed at increasing the number of individuals in each responsibility group (Brunori *et al.*, 2019).

individual the norm consumption is then equal to the average consumption in the group she belongs to. In other words, the EOp distribution removes the within group inequality in consumption due to circumstances out of individual control like individual innate talents, parent's education, luck and so on. At the same time, it preserves the consumption differences induced by different level of education, work time and household size.

To define the fair health score, our EOp principle is the following: individuals (without disability or chronic diseases) who go as often to visit doctors and spend the same time taking care of themselves are entitled to have the same health score. We approximate these two choices using: *(i)* $PersCareTime_i$ which takes values from 1 to 5 depending on the individual's quintile position in the cumulative distribution of the amount of weekly time spent in leisure and personal care; *(ii)* $NVisitDoc_i$ that takes values from 1 to 5 and is constructed taking quintile positions in the distribution of the sum of yearly visits to general practitioner, specialist, dentist, physiotherapist and home nursing care.[7] When computing $NVisitDoc_i$ we account for two confounding factors: (i) younger individuals tend to have better health and may not go to doctor simply because they do not need to; (ii) the choice of visiting a doctor may be constrained by its cost, and not all households can afford it. We define three age classes (below 30, 31-59, above 60) and three groups based on the answer to the following question: *In general, do you find that the money you have to pay in personal contributions to health costs easily/it is difficult to/it is impossible to fit in your budget?* Hence, when assigning $NVisitDoc_i$ we consider individuals within the same age class and that give the same answer to the previous question. This allows us to divide the population in 25 groups of individuals. Within each of these groups, any difference in health score should be considered unacceptable since it can be due to genes or bad luck, for example. Assuming that these circumstances can affect health both positively and negatively, we define the fair health score of each individual as the average score in the group he belongs to.

Individuals with chronic diseases or disability (34.49% of our sample) are excluded from the procedure above described. This a special category

[7]To limit the role of outliers, we set each component of the sum to be maximum 50 (that is more than 4 visits per month to the same type of doctor).

of individuals who we assume to be not responsible for their condition so that, independently of their effort, they should be entitled to a minimum level of health score. For those individuals, the fair health score is the one from the POV distribution.

Online Appendix A reports the summary statistics for the norm consumption and health score computed in each of the three definitions of the fair distribution.

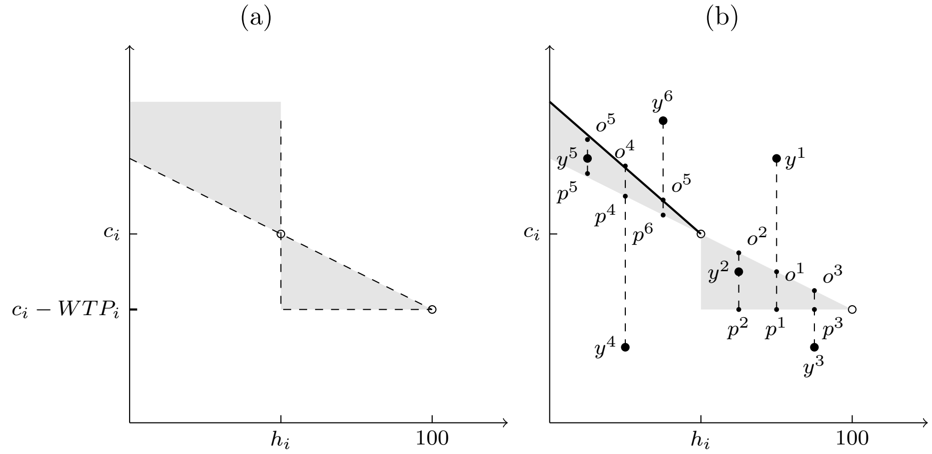
# 4 Set identification of Equivalent Advantage

An essential step for implementing our framework is the identification of the individual indifference curve passing through the current outcome. We will therefore introduce a nonparametric method that can be used whenever we have information on all dimensions (i.e. income, health, etc.) and willingness to pay (WTP) for the non-income dimensions. For simplicity we will restrict our exposition below to the consumption-health space setting that we use in our application. The MEqIn database contains for each respondent the WTP in terms of consumption in order to go from the observed health score $h_i$ to a perfect health score of 100. We denote this information by $WTP_i$.

**Set identification.** As we illustrate in Figure 2a, we have for each individual two points on the indifference curve we want to identify. The observed outcome $(c_i, h_i)$ and the indifferent bundle with perfect health $(c_i - WTP_i, 100)$. The depicted gray area defines the region containing all the possible convex indifference curves passing through these two bundles. Since we do not want to impose any extra (parametric) structure on preferences, we consider the shaded area in Figure 2a as a thick indifference curve containing the 'true' individual indifference curve. This thick indifference curve can be used to set identify our EA, simply by computing a lower and upper bound on the EA. By construction this closed interval will contain the true EA.

In Figure 2b we illustrate our set identification. Let us start by considering the cases where the norm bundle is situated to the right of the observed bundle (i.e. $y^1$, $y^2$ and $y^3$). If $y^1$ is the fair outcome, then the in-

Figure 2: Set identification of Equivalent Advantage.




dividual clearly prefers this over the observed outcome $(c_i, h_i)$. To quantify the EA, we therefore need to find a bundle with the same value for health as $y^1$ but a lower level of consumption. Given the thick indifference curve we can not exclude any bundle between $o^1$ and $p^1$. We say that $o^1$ is the optimistic scenario, since it allows to compute a upper bound on the EA, which is the minimal welfare loss (or maximum advantage). While $p^1$ is the pessimistic scenario since it gives us the lower bound, that coincides with the maximal welfare loss (or minimum advantage). Clearly both lower and upper bound are in this situation negative since the individual prefers $y^1$ over $(c_i, h_i)$. That is, the individual is disadvantaged which explains our terminology.

The same reasoning applies to the situations with bundles $y^2$ and $y^3$. The only difference being that with $y^2$, the lower bound is negative and the upper bound is positive. Meaning that we can not exclude that the true EA is zero. While for $y^3$ both lower and upper bound are positive, which implies that whatever the true preferences are we always conclude that the EA is positive. That is, the individual prefers the current outcome over the fair outcome and is therefore advantaged.

**Restricting individual preferences.** Unfortunately, due to data limitation, we can not readily apply the same reasoning when the norm bundle is situated to the left of the observed bundle (i.e. $y^4$, $y^5$ and $y^6$). In these situations, the grey area in Figure 2a does not exclude Leontief preferences. This implies that our set identification approach does not result in informative bounds since the upper bound will always be plus infinity. Better data in the future, containing information on willingness to accept, say, a health score equal to 10, would easily allow us to circumvent this issue.[8] For now, we we will introduce a restrictive assumption to obtain informative upper bounds. This is not essential for our method and only serves to ease the illustration of our results in he next section.

Note that an alternative would be to drop upper bounds all together. Such a strategy is normatively appealing under a prioritarian perspective which emphasizes the situation of the worse off individuals. Indeed, when an individual is disadvantaged, his EA's lower-bound is the highest welfare loss he may experience; while, if an individual is advantaged, then the lower-bound is the lowest welfare gain he may experience. In other words, at the EA's lower-bound we exacerbate welfare losses and alleviate welfare gains. The main technical advantages of this solution are: (i) the absence of additional assumptions about individual preferences and (ii) the point identification of the EA. In Section 5.3 we implement this conservative solution to compare the intensity of EA across population's subgroups.

To introduce our assumption on the shape of the individual preferences, recall that for an individual with $h_i < 100$, moving along the indifference curve from the point $(c_i, h_i)$ to the point $(c_i - WTP_i, 100)$ is equivalent to buying $100 - h_i$ units of health at a price $uWTP_i = WTP_i/(100 - h_i)$. Conversely, if the same individual moves, along his indifference curve, towards a bundle $(c_i', h_i')$ such that $c_i' \geq c_i$ and $h_i' < h_i$, then it is as if she was selling $h_i - h_i'$ units of health at a price $uWTA_i = (c_i' - c_i)/(h_i - h_i')$, which corresponds to her willingness-to-accept ($\text{WTA}_i$) for one unit of health.

The literature has shown a persistent disparity between WTP and WTA (see, for instance, Kahneman & Tversky, 1980; Knetsch & Sinden, 1984; Knetsch, 1989; Kahneman *et al.*, 1990; Huck *et al.*, 2005). Early experimental evidence concludes that WTA can be four times larger than WTP because of the so called endowment effect (Knetsch & Sinden, 1984; Knetsch, 1989). In line with this literature, we set $uWTA_i = 4\,uWTP_i$ and ob-

[8] In our setting, willingness to accept a health score equal to 40 would suffice, as no fair health score falls below the poverty line (health score equal to 42).

tain a well-defined upper-bound. This solution preserves the heterogeneity in individual preferences, since it depend on $WTP_i$, and implies that the solid (dark) line in Figure 2b will have different slopes depending on the considered individual. Admittedly, it relies on an ad-hoc assumption that also leaves open the problem of defining $uWTA$ for individuals with zero WTP.[9]

A more conservative approach is to set $uWTA_i$ equal to the maximum (observed) $uWTP$; i.e. $uWTA_i = Max\ uWTP$ for all $i$. This solution is data driven and focuses on the individual with the most intense preferences for health. However, it removes part of the heterogeneity in individual preferences because the solid line in Figure 2b will have the same slope independently of the considered individual, and may produce overoptimistic assessments of EA. We recomputed all our results for this more conservative approach, which resulted in a limited impact on the respective upper bounds. To streamline our discussion we will therefore not focus on this scenario.

**Empirical results.** Table 1 shows, for each norm, some descriptive statistics of the EA distributions. As expected there is, for both lower and upper bounds, quite some heterogeneity in the EA. Interpreting the EA as the welfare gain or loss caused by the existing unfairness, we can use the average EA as a synthetic measure of social justice. Focusing on the lower-bound (or implementing a prioritarian approach), we always observe that the average welfare gains are not sufficient to compensate for the average losses. Except for the POV norm, which by construction can not result in a positive EA, this conclusion does not carry over to the upper bound (or using an optimistic approach). The latter is probably driven by our partial identification results, which can be significantly improved by having better data.

Next, Table 2 shows the distribution of the difference between upper and lower-bounds of EA, normalized by the average consumption. Interestingly, despite the weak assumptions on individual preferences, our partial identification of EA is fairly precise. In particular, for about a quarter of

[9] In what follows, when $WTP_i = 0$ (48% of our sample) we set $uWTA_i$ equal to four times the average $uWTP$.

Table 1: Summary statistics of the EA distributions

| | Mean | p10 | p25 | p50 | p75 | p90 |
|---|---|---|---|---|---|---|
| *EO* | | | | | | |
| Lower-bound | -39.10 | -485.62 | -317.62 | -97.43 | 177.71 | 499.25 |
| Upper-bound | 147.28 | -396.74 | -224.65 | 22.05 | 335.19 | 730.99 |
| *EOp* | | | | | | |
| Lower-bound | -54.11 | -447.41 | -278.66 | -83.53 | 128.06 | 379.65 |
| Upper-bound | 73.90 | -360.94 | -201.85 | -1.24 | 229.96 | 548.66 |
| *POV* | | | | | | |
| Lower-bound | -45.05 | -166.20 | -1.20 | 0.00 | 0.00 | 0.00 |
| Upper-bound | -23.49 | -98.20 | -0.20 | 0.00 | 0.00 | 0.00 |

our sample we can point identify EA in all the three norms, and for the majority of the sample the difference is below 100 €. The higher differences are registered when we implement multidimensional equality (EO), while the EOp norm seems to be less sensitive to the change from upper to lower-bound EA. Finally, note that the POV norm induces very small differences. This is partially due to the fact that poor individuals tend to express lower WTP, but is mainly by construction.

Table 2: Difference between EA's upper and lower-bounds.

| | Mean | p10 | p25 | p50 | p75 | p90 |
|---|---|---|---|---|---|---|
| EO | 0.24 | 0.00 | 0.03 | 0.13 | 0.29 | 0.44 |
| EOp | 0.16 | 0.00 | 0.00 | 0.07 | 0.19 | 0.33 |
| POV | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.06 |

**Description**: The table shows summary information on the distribution of the difference between upper and lower-bound of EA, as fraction of the average consumption (777.62 €).

# 5 Social justice in Belgium

In this section we want to illustrate that, despite our nonparametric and general set-up, we can draw meaningful conclusions that respect both heterogeneity in individual preferences and the multidimensionality of fairness.

We start by discussing the fraction of disadvantaged individuals in the population and their characteristics. Subsequently we focus on inequality of disadvantage by comparing different subgroups. We end by a basic exercise to explore the importance of individual preferences for our empirical conclusions.

## 5.1 The fraction of disadvantaged individuals

To compare our method to more traditional measures we present in Table 3 the percentage of disadvantaged individuals computed according to different scenarios. The first two columns refer to the unidimensional approach which counts the number of individuals that have a current consumption (column (1)) or health (column (2)) below the fair level induced by the given norm. All the other columns implement a multidimensional approach. In columns (3) and (4) we do not consider individual heterogeneous preferences, but instead use some common weighting of both dimensions. Column (3) reports the percentage of individuals that are disadvantaged in terms of both consumption and health, which is called the counting approach in the poverty literature. Column (4) instead contains the percentage of individuals that have a negative average standardized difference (ASD). This is a multidimensional measure of the divergence between the current and fair outcome, constructed as the average of the standardized consumption advantage and the standardized health advantage.[10] The ASD is an arbitrary measure, implementing the same weights for both dimensions for all individuals, motivated by the well-known Human Development Index (HDI). Finally, columns (5) and (6) report the fraction of individuals labeled disadvantaged by respectively our lower bound on EA and our upper bound.

Unsurprisingly the different approaches lead to very different conclusions. Whatever norm we consider, columns (3) and (4) confirm once more that it is important to respect the multidimensionality of social justice and that the counting approach gives a too optimistic summary. Next, columns (5) and (6) show that allowing for heterogeneity in the weighting has quite

[10] Using the notation from Section 3, $ASD_i = 0.5 \cdot S(c_i - c_i^{\star}) + 0.5 \cdot S(h_i - h_i^{\star})$. $S(x) = (x - \bar{x})/\sigma(x)$, where $\bar{x}$ is the average and $\sigma(x)$ is the standard deviation of $x$.

Table 3: Percentage of disadvantaged individuals for unidimensional and multidimensional approaches.

| | (1) Cons. | (2) Health | (3) Count | (4) ASD | (5) AEA-lb | (6) AEA-ub |
|---|---|---|---|---|---|---|
| EO | 0.578 | 0.486 | 0.291 | 0.494 | 0.604 | 0.481 |
| EOp | 0.562 | 0.398 | 0.238 | 0.503 | 0.592 | 0.501 |
| POV | 0.173 | 0.131 | 0.031 | 0.213 | 0.251 | 0.251 |

**Description**: The table shows the percentage of individuals with: (1) consumption below the fair one; (2) health below the fair one; (3) both consumption and health below the fair level; (4) negative average standardized difference; negative EA's lower (5) or upper (6) bound.

some impact on concluding who is disadvantaged. By construction, the share of disadvantaged people is much higher for the pessimistic scenario (i.e. the lower bound) compared to the optimistic scenario (i.e. the upper bound). Nevertheless for both cases the numbers are suggesting that column (4) depicts a too optimistic conclusion.

When zooming in on the different norms we considered, our results for the POV norm show the importance of individual preferences for measuring poverty. To recall, these results can be interpreted as a measure of multidimensional poverty with heterogeneous weights. According to our EA measure, one quarter of the population is poor, which is a much higher number than the more traditional approaches in the other columns, including our measure that we introduce to mimic the Human Development Index. Next, we conclude that compared to the average in society (the EO norm) between 48% and 60% of society is deprived. This conclusion remains when we correct for effort (the EOp norm). Interestingly, we conclude from column (6) that our multidimensional approach should not by default always result in more deprived people compared to an unidimensional approach. Indeed the fraction of disadvantaged people in column (1) is well above those in column (6). Better data, improving our upper bound estimates could further perfectly demonstrate this.

Finally, in Online Appendix B we present some more insight in the overlap between our method and the unidimensional approach. That is, we present the number of individuals that is labeled disadvantaged by more than one method or by just one. From this we conclude that there is no individual having a negative EA, without being deprived in either consumption or health. This confirms that allowing for individual preferences does not necessarily lead to extreme results. On the contrary, some individuals are deprived according to the unidimensional measure, but are still having a positive EA. Meaning that according to their own preferences they do not consider themselves disadvantaged.

## 5.2 Characteristics of the disadvantaged people

The previous section demonstrated that our method allows for a much richer analysis of social justice at the aggregate level. As a further illustration we complement this by a regression analysis of the characteristics of the disadvantaged population. Among other things, this shows that our set identification results are not preventing us from this more detailed analysis. For brevity, we only discuss the results for the EO norm. The corresponding results for the other two norms, leading to similar conclusions, can be found in Online Appendix C.

We investigate how the probability of being disadvantaged varies with age, gender, employment status and region of residence. The dependent variable of our logistic regression takes a value of one when the upper bound is negative, which means that the whole range of possible EAs is for sure negative and thus that we conclude that the individual is disadvantaged.

In Table 4 we consider gender, age, employment status and region of residence separately. Column (1) suggests the presence of a gender gap in EA, with women being more likely than men to prefer the average outcome to the current one. Looking at the age partition of our sample we observe a slightly U-shaped pattern for the probability of being disadvantaged, along the life cycle. 40-49 years old individuals are the least likely to be disadvantaged, while the over 80 years old are the most likely to be disadvantaged. Not surprisingly, employment status - column (3) - is extremely correlated with the risk of disadvantage, with the unemployed individuals performing worse that the retired ones. This may come as a surprise, given that retired people are older and likely to be in worse health conditions. Finally,

Table 4: Risk of being disadvantaged in the EO norm.

| | (1) Gender | (2) Age | (3) Employment status | (4) Region of residence |
|---|---|---|---|---|
| Woman | 0.167** | | | |
| | (0.076) | | | |
| 18-29 | | 0.256* | | |
| | | (0.147) | | |
| 30-39 | | 0.269** | | |
| | | (0.128) | | |
| 50-59 | | 0.192 | | |
| | | (0.121) | | |
| 60-69 | | 0.041 | | |
| | | (0.127) | | |
| 70-79 | | 0.324** | | |
| | | (0.144) | | |
| 80+ | | 1.231*** | | |
| | | (0.204) | | |
| Unemployed | | | 1.120*** | |
| | | | (0.172) | |
| Retired | | | 0.667*** | |
| | | | (0.081) | |
| Flanders | | | | -0.124 |
| | | | | (0.145) |
| Wallonia | | | | -0.146 |
| | | | | (0.152) |
| Constant | -0.164*** | -0.291*** | -0.400*** | 0.047 |
| | (0.056) | (0.088) | (0.052) | (0.136) |

Standard errors in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Description**: The table shows the results of a logistic regression where the dependent variable takes value 1 if the EA upper bound is lower than zero. We compute robust standard errors. The reference categories are (1) Gender: Male; (2) Age: 40-49; (3) Employment status: Employed; (4) Region of residence: Brussels.

Table 5: Risk of being disadvantaged in the EO norm. Interaction effects

| | (1) All | (2) Bruxelles | (3) Flanders | (4) Wallonia |
|---|---|---|---|---|
| Male Young Employed | -0.252 | -1.625 | -0.460* | 0.488 |
| | (0.218) | (1.121) | (0.279) | (0.401) |
| Female Young Employed | 0.593*** | -0.121 | 0.778*** | 0.355 |
| | (0.209) | (0.819) | (0.263) | (0.390) |
| Male Young Non-Employed | 0.987*** | 0.573 | 1.053* | 1.007* |
| | (0.371) | (0.960) | (0.568) | (0.578) |
| Female Young Non-Employed | 1.161*** | 0.000 | 1.852** | 0.601 |
| | (0.361) | (.) | (0.798) | (0.454) |
| Female Adult Employed | 0.048 | -0.219 | 0.159 | -0.114 |
| | (0.116) | (0.429) | (0.146) | (0.212) |
| Male Adult Non-Employed | 1.208*** | 0.860 | 1.592*** | 0.909*** |
| | (0.226) | (0.915) | (0.349) | (0.328) |
| Female Adult Non-Employed | 1.406*** | 1.614** | 1.471*** | 1.261*** |
| | (0.174) | (0.628) | (0.238) | (0.283) |
| Male Old Employed | -0.266 | -0.526 | -0.282 | -0.169 |
| | (0.327) | (1.262) | (0.418) | (0.578) |
| Female Old Employed | -0.489 | 0.167 | -0.697 | -0.428 |
| | (0.383) | (1.044) | (0.523) | (0.707) |
| Male Old Non-Employed | 0.555*** | 0.619 | 0.732*** | 0.121 |
| | (0.131) | (0.449) | (0.166) | (0.246) |
| Female Old Non-Employed | 0.498*** | -0.158 | 0.505*** | 0.615*** |
| | (0.126) | (0.466) | (0.164) | (0.218) |
| Constant | -0.427*** | -0.167 | -0.466*** | -0.419*** |
| | (0.083) | (0.290) | (0.106) | (0.151) |
| Observations | 2778 | 210 | 1662 | 901 |

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Description**: The table shows the results of a logistic regression where the dependent variable takes value 1 if the EA upper bound is lower than zero. We compute robust standard errors. Model (1) considers the entire sample. Models (2) - (4) focus on specific regions of residence. The baseline category is Male Adult Employed. **Note**: Female Young Non-Employed omitted in model (2) because all are disadvantaged.

the region of residence in column (4) seems to play no role on the risk of disadvantage.

Next, we focus on the interaction effects of our demographic information. We distinguish between employed and non-employed (retired and unemployed), and consider three age groups: young (below 30 years old), adult (between 30 and 60) and old (above 60). We then define a list of binary variables for each combination of these categories, distinguishing between male and female. Table 5 reports the results, for the EO norm, taking adult employed males as baseline category.

The first column confirms the hypothesis of a gender gap in the risk of disadvantage. Young females are more likely to have negative EA, independently of the employment status. Not surprisingly, at all age groups, non-employed male are worse off than adult employed. Interestingly, the gender gap fades out among old individuals. While these results are in line with Table 4, the same conclusion does not hold for the regional analysis. Comparing columns (3) and (4), Wallonia seems to be a region with weaker between-groups disparities in terms of risk of disadvantage. Indeed, almost all the significant coefficients in column (3) are either of smaller magnitude or not significant in (4). Because of the smaller sample size, many of the dynamics highlighted before loose significance when we focus on the Brussels region - column (2). A relevant exception in this case is the risk of disadvantage for young and adult non-employed females, which is much higher.

## 5.3 Equivalent Advantage across groups

In the first two sections, we focused on the incidence of disadvantage. Another relevant aspect that emerges from Table 1 is the presence of high inequality in EA. To further investigate this, we partition the population in groups - based on age, gender, employment status or region of residence - and compare their EA distributions. This exercise complements the previous ones in two ways. First, it considers EA as a continuous variable, rather than a binary one that distinguishes advantaged from disadvantaged individuals. In particular, it helps us understanding if the probability of being disadvantaged goes hand to hand with the intensity of such disadvantage, which is in line with poverty focusing on headcount and poverty gap ratios. Second, while our identification of the disadvantaged individuals is based on the EA upper-bound, in the following section we implement a conservative and prioritarian approach that gives more weight to individuals with negative EA. As discussed in the previous section, such an approach is consistent with analyzing the distribution of the EA's lower-bounds. Once more this shows that the generality of our approach does not exclude empirical bite.

The figures in Online Appendix D contain the generalized Lorenz curves of EA for our population subgroups. Following the discussion in Section 2.4, comparing EA for, say men and women, in terms of their generalized Lorenz curves is a robust method to establish whether one group is unambiguously more advantaged than the other. Table 6 summarizes the subgroup comparisons based on their generalized Lorenz curves. The sign $>$ indicates that the first group dominates the second one while $<$ denotes the opposite case. If the curves are crossing, the ranking is indeterminate; we indicate this situation with the letter I and we perform another comparison between the generalized Lorenz curves for $p$ in eq. (1) between 0.1 and 0.9. The results, if different, are reported in parenthesis.

The first row of Table 6 confirms the presence of a clear gender gap, that remain consistent across norm distributions. Interestingly, for two norms we observe an instance of indeterminate ranking; this is because at the very bottom of the distributions the gender differences in terms of EA shrink and sometimes invert in such a way that a leximin order may consider women more advantaged than men. Nevertheless, if we remove the bottom and top decile of the EA distributions, the gender comparison shows that independently of individual preferences (which we account for with the EA) women struggle more than men in seeing their effort fairly remunerated, suffer more because of relative poverty and tend to be more disadvantaged than men compared to the average individual in the society.

To investigate the age distribution of EA, we consider the three groups (young, adult and old) introduced above. A stable result emerges here: adult individuals are better off with respect to the old ones. This is intuitive since we can expect adults to have on average both a better health and a higher consumption given life-cycle dynamics. The multidimension-

Table 6: Generalized Lorenz dominance of EA distributions by population subgroups.

| | EO | EOp | POV |
|---|---|---|---|
| **Gender** | | | |
| Men - Women | I (>) | > | I (>) |
| **Age** | | | |
| Adult - Young | I | I | < |
| Adult - Old | > | > | > |
| Young - Old | I | > | > |
| **Region** | | | |
| Bxl - Fla | < | < | < |
| Bxl - Wal | < | < | < |
| Fla - Wal | I (<) | < | < |
| **Employment status** | | | |
| Empl - Unempl | I (>) | I | > |
| Empl - Retired | > | > | > |
| Unempl - Retired | I | I | I |

**Description**: The table shows the results of the generalized Lorenz ranking between EA distributions. The sign $>$ indicates that the first group dominates the second one while $<$ denotes the opposite case. If the curves are crossing we put the letter I and we perform another comparison between the generalized Lorenz curves for $p$ in eq. (1) between 0.1 and 0.9. The results (if different) are reported in parenthesis.

ality of our measure plays a relevant role when we compare young and old individuals. Indeed, it appears that the better health enjoyed by the former compensates for the average higher consumption of the latter, often inducing dominance of young over old individuals.

The comparison between regions of residence approximates the geographical distribution of Equivalent Advantages. It is striking to observe that Wallonia performs better than Brussels and Flanders independently of the norm. Since the fair outcomes are computed at the national level, this is a surprising result given the better economic conditions of Flanders and Brussels, and confirms that economic development does not necessarily go hand in hand with social justice.[11] Next, the fact that Brussels is dominated by the two other regions, is most likely a result of the high inequalities in the biggest metropolitan area of the country.

Our last comparison is between employed, unemployed and retired indi-

[11] Among those individual that expressed positive WTP, the average unitary WTP is 5.04 in Brussels, 5.93 in Flanders and 5.17 in Wallonia. These differences are too small to fully explain the observed dominance.

viduals. This does not result in surprising results. To better understand the indeterminacy results for the employed-unemployed comparison, we want to note that under the EO and EOp norm, there is a share of extremely disadvantaged unemployed individuals that still has better EA than the employed ones. After excluding the lower part of the distributions, we can be quite confident in concluding that employed individuals have better EA than the other group.

## 5.4 The role of preferences

As a final exercise, we want to do a back of the envelope analysis, to investigate how much of the previous conclusions is driven by allowing for heterogeneity in the weighting of the dimensions (i.e by taking the individual preferences into account). We therefore first run the following OLS regression

$$EA(n)_i^e = \beta_0 + \beta_1 c_i + \beta_2 h_i + \beta_3 c^\star(n)_i + \beta_4 h^\star(n)_i + \epsilon_i,$$

where $EA(n)_i^e$ is the EA's upper or lower bound - $e \in \{lb, ub\}$ - of individual $i$ under the fair distribution $n \in \{EO, EOp, POV\}$, $c_i$ and $h_i$ are the actual outcome, and $c^\star(n)_i$ and $h^\star(n)_i$ are the respective norm outcomes. The only element of the construction of EA that is not included in this regression is the individual preferences (based on the WTP). The unexplained variability, $1 - R^2$, is therefore indicative of the role of individual preferences in determining Equivalent Advantages.

The results are presented in Online Appendix E. All coefficients are strongly significant and they all have the expected signs. The unexplained variation ranges from 7% to 56%, which indicates that is very important to allow for this heterogeneity. At the same time it is not driving all our conclusions. Preferences seem to be more important for the lower bound estimates under the POV norm and the upper bound estimates for the EO and EOp norm.

As a final exercise, we investigate to what extent the results of Table 6 can be obtained using an objective measure of advantage like the ASD introduced before (i.e. using the same weights for both dimensions for all individuals). These results are also included in Online Appendix E. From that table we can conclude that no less than 70% of our results change, mainly from dominance towards indeterminacy. This holds in particular for our age and regional comparisons. Once more we therefore conclude that it is very important to evaluate social justice by respecting individual preferences.

# 6 Conclusion

It is by now well accepted that to analyze distributive fairness it is essential to take into account both its multidimensional nature and the related heterogeneity of individual preferences. However, this general set-up should still allow practitioners to readily apply the suggested methods and, more importantly, to easily draw interesting empirical conclusions.

We have therefore introduced, and axiomatised, the concept of Equivalent Advantage as a measurement of multidimensional fairness. Since our measure fully allows for heterogeneity in the individual preferences over the observed outcome and in the fair outcome, we have introduced a nonparametric and easy to implement set identification method of the individual indifference curve. This in turns allows for set identifying our Equivalent Advantage measure.

To illustrate the versatility and attractivity of our new approach, we have used the MEqIn database to analyze fairness in the distribution of consumption and health in Belgium. Our results clearly demonstrate that the generality of our approach does not limit our empirical bite, both in terms of incidence (at the aggregate and individual level), as well as for a distributive analysis. Importantly, the multidimensional nature of our approach does not simply lead to the appearance of new disadvantaged individuals, but provides a different picture of distributive justice, which can be hardly replicated when disregarding individual preferences.

The simplicity and flexibility of our framework should motivate its implementation for policy evaluations. If income and health are the targeted outcomes of a structural policy, we can assess its distributive impact by looking at the Equivalent Advantage distribution computed before and after the intervention. More in general, although we refrained from directly doing so, Equivalent Advantage can also be the argument of a social evaluation function. For example, the average EA under the POV norm can be used to measure relative poverty. Finally, the value of our theoretical contribution is strengthened by the methodological one. We have shown that much can be said by set identifying EA. However, much more can be done with better data that improve our partial identification. In this sense, collecting and using information on willingness to accept is of crucial interest. We hope that our paper also stimulates this research agenda.

# References


Almås, Ingvild, Cappelen, Alexander W, Lind, Jo Thori, Sørensen, Erik Ø, & Tungodden, Bertil. 2011. Measuring unfair (in) equality. *Journal of Public Economics*, **95**(7-8), 488–499.

Alvaredo, Facundo, Chancel, Lucas, Piketty, Thomas, Saez, Emmanuel, & Zucman, Gabriel. 2017. Global inequality dynamics: New findings from WID. world. *American Economic Review*, **107**(5), 404–09.

Andreoli, Francesco, & Zoli, Claudio. 2020. From unidimensional to multidimensional inequality: a review. *Metron*, **78**(1), 5–42.

Atkinson, Anthony B, & Bourguignon, Francois. 1987. Income distribution and differences in needs. *Pages 350–370 of: Arrow and the Foundations of the Theory of Economic Policy*. Springer.

Banerjee, Abhijit, Deaton, Angus, & Duflo, Esther. 2004. Wealth, health, and health services in rural Rajasthan. *American Economic Review*, **94**(2), 326–330.

Blundell, Richard, Browning, Martin, & Crawford, Ian. 2008. Best nonparametric bounds on demand responses. *Econometrica*, **76**(6), 1227–1262.

Bosmans, Kristof, Lauwers, Luc, & Ooghe, Erwin. 2009. A consistent multidimensional Pigou–Dalton transfer principle. *Journal of Economic Theory*, **144**(3), 1358–1371.

Brunori, Paolo, Peragine, Vito, & Serlenga, Laura. 2019. Upward and downward bias when measuring inequality of opportunity. *Social Choice and Welfare*, **52**(4), 635–661.

Capéau, Bart, Cherchye, Laurens, Decancq, Koen, Decoster, André, De Rock, Bram, Maniquet, François, Nys, Annemie, Périlleux, Guillaume, Ramaekers, Eve, Rongé, Zoé, *et al.* 2020. *Well-being in Belgium*. Springer.

Chetty, Raj, Hendren, Nathaniel, Kline, Patrick, Saez, Emmanuel, & Turner, Nicholas. 2014. Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *American Economic Review*, **104**(5), 141–47.

Cowell, Frank A, Flachaire, Emmanuel, & Bandyopadhyay, Sanghamitra. 2013. Reference distributions and inequality measurement. *The Journal of Economic Inequality*, **11**(4), 421–437.

Crawford, Ian, & De Rock, Bram. 2014. Empirical revealed preference. *Annu. Rev. Econ.*, **6**(1), 503–524.

de Frahan, Lancelot Henry, & Maniquet, François. 2021. Preference responsibility versus poverty reduction in the taxation of labor incomes. *Journal of Public Economics*, **197**, 104386.

Deaton, Angus. 2008. Height, health, and inequality: the distribution of adult heights in India. *American Economic Review*, **98**(2), 468–74.

Decancq, Koen, & Lugo, Maria Ana. 2012. Inequality of wellbeing: A multidimensional approach. *Economica*, **79**(316), 721–746.

Decancq, Koen, & Lugo, María Ana. 2013. Weights in multidimensional indices of wellbeing: An overview. *Econometric Reviews*, **32**(1), 7–34.

Decancq, Koen, & Nys, Annemie. 2021. Non-parametric well-being comparisons. *European Economic Review*, 103666.

Decancq, Koen, Fleurbaey, Marc, & Schokkaert, Erik. 2017. Wellbeing inequality and preference heterogeneity. *Economica*, **84**(334), 210–238.

Decancq, Koen, Fleurbaey, Marc, & Maniquet, François. 2019. Multidimensional poverty measurement with individual preferences. *The Journal of Economic Inequality*, **17**(1), 29–49.

Dworkin, Ronald. 1981a. What is equality? Part 1: Equality of welfare. *Philosophy & public affairs*, 185–246.

Dworkin, Ronald. 1981b. What is equality? Part 2: Equality of resources. *Philosophy & public affairs*, 283–345.

Fleurbaey, Marc. 2008. *Fairness, responsibility, and welfare.* Oxford University Press.

Fleurbaey, Marc, & Blanchet, Didier. 2013. *Beyond GDP: Measuring welfare and assessing sustainability.* Oxford University Press.

Fleurbaey, Marc, & Tadenuma, Koichi. 2014. Universal social orderings: An integrated theory of policy evaluation, inter-society comparisons, and interpersonal comparisons. *Review of Economic Studies*, **81**(3), 1071–1101.

Fleurbaey, Marc, & Van der Linden, Martin. 2021. Fair social ordering, egalitarianism, and animal welfare. *American Economic Journal: Microeconomics*, **13**(4), 466–91.

Fleurbaey, Marc, & Zuber, Stéphane. 2021. Fair utilitarianism. *American Economic Journal: Microeconomics*, **13**(2), 370–401.

Huck, Steffen, Kirchsteiger, Georg, & Oechssler, Jörg. 2005. Learning to like what you have–explaining the endowment effect. *The Economic Journal*, **115**(505), 689–702.

Jenkins, Stephen P, & Lambert, Peter J. 1997. Three 'I's of poverty curves, with an analysis of UK poverty trends. *Oxford economic papers*, **49**(3), 317–327.

Kahneman, Daniel, & Tversky, Amos. 1980. Prospect theory. *Econometrica*, **12**.

Kahneman, Daniel, Knetsch, Jack L, & Thaler, Richard H. 1990. Experimental tests of the endowment effect and the Coase theorem. *Journal of political Economy*, **98**(6), 1325–1348.

Knetsch, Jack L. 1989. The endowment effect and evidence of nonreversible indifference curves. *The American Economic Review*, **79**(5), 1277–1284.

Knetsch, Jack L, & Sinden, John A. 1984. Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *The Quarterly Journal of Economics*, **99**(3), 507–521.

Maasoumi, Esfandiar. 1986. The measurement and decomposition of multidimensional inequality. *Econometrica: Journal of the Econometric Society*, 991–997.

Mahler, Daniel Gerszon, & Ramos, Xavier. 2019. Equality of Opportunity in Four Measures of Well-Being. *Review of Income and Wealth*.

Maniquet, François, & Neumann, Dirk. 2021. Well-Being, Poverty, and Labor Income Taxation: Theory and Application to Europe and the United States. *American Economic Journal: Microeconomics*, **13**(2), 276–310.

Manski, Charles F. 2003. *Partial identification of probability distributions*. Vol. 5. Springer.

Piacquadio, Paolo Giovanni. 2017. A fairness justification of utilitarianism. *Econometrica*, **85**(4), 1261–1276.

Piketty, Thomas, & Saez, Emmanuel. 2006. The evolution of top incomes: a historical and international perspective. *American Economic Review*, **96**(2), 200–205.

Piketty, Thomas, Postel-Vinay, Gilles, & Rosenthal, Jean-Laurent. 2006. Wealth concentration in a developing economy: Paris and France, 1807-1994. *American Economic Review*, **96**(1), 236–256.

Piketty, Thomas, Yang, Li, & Zucman, Gabriel. 2019. Capital accumulation, private property, and rising inequality in China, 1978–2015. *American Economic Review*, **109**(7), 2469–96.

Ramos, Xavier, & Van de Gaer, Dirk. 2016. Approaches to inequality of opportunity: Principles, measures and evidence. *Journal of Economic Surveys*, **30**(5), 855–883.

Samuelson, Paul A. 1974. Complementarity: An essay on the 40th anniversary of the Hicks-Allen revolution in demand theory. *Journal of Economic Literature*, **12**(4), 1255–1289.

Samuelson, Paul A, & Swamy, Subramanian. 1974. Invariant economic index numbers and canonical duality: survey and synthesis. *American Economic Review*, **64**(4), 566–593.

Shorrocks, Anthony F. 1983. Ranking income distributions. *Economica*, **50**(197), 3–17.

Stiglitz, Joseph E, Sen, Amartya, Fitoussi, Jean-Paul, *et al.* 2009. *Report by the commission on the measurement of economic performance and social progress.*

Tsui, Kai-Yuen. 1995. Multidimensional generalizations of the relative and absolute inequality indices: the Atkinson-Kolm-Sen approach. *Journal of Economic Theory*, **67**(1), 251–265.

Tsui, Kai-yuen. 1999. Multidimensional inequality and multidimensional generalized entropy measures: An axiomatic derivation. *Social Choice and Welfare*, **16**(1), 145–157.

UN, United Nations Development Programme. 2018. Human Development Indices and Indicators. *New York: UNDP.*

# Online Appendix

## Appendix A Summary statistics and main variables

We provide here more detailed information on the variables from the MEqIn database that are most important for our analysis. For more details on all variables, see *https://sites.google.com/view/meqin/data*. Subsequently we present some descriptive statistics.

**Health score**: a 0-100 score computed for all dimensions of health; it summarizes the answers to more than 40 questions related to physical and emotional health, including the consequences on everyday life. In the data source, this variable is coded *V40052101* and is used to elicit WTP.

**Personal consumption**: it is the sum of individual monthly spending on restaurants, cigarettes/tobacco, clothes, hygiene/care products, leisure activities, education, transport, food and other things, plus the monthly household spending on utilities, insurance, holidays and other things, divided by the household size. In the data source, this variable is coded *V40124801* and is used to elicit WTP.

**Willingness to pay:** Consider an imaginary situation and how it could be equivalent for you to your current situation. In this new scenario, you have perfect health for the next twelve months. To be equivalent to your current situation, you have to reduce your monthly consumption. How much would you have to reduce your monthly consumption in the next twelve months for the new scenario with perfect health to be equivalent for you to your current situation? In the data source, this variable is coded *WTP_ health*.

Table 7 describes the composition of our sample. Table 8 reports the summary statistics of the observed and fair distributions, and willingness to pay for perfect health.

Table 7: Sample composition

| | Population Share |
|---|---|
| Female | 0.51 |
| Age 18-29 | 0.10 |
| Age 30-59 | 0.57 |
| Age 60 + | 0.33 |
| Living in Bruxelles | 0.07 |
| Living in Flanders | 0.60 |
| Living in Wallonia | 0.33 |
| Employed | 0.56 |
| Unemployed | 0.06 |
| Retired | 0.38 |
| Observations | 2778 |

**Description**: The table reports the weighted frequencies for various individual characteristic in our sample.

Table 8: Summary statistics .

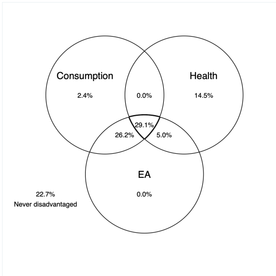
| | Mean | SD | Min | p25 | p50 | p75 | Max |
|---|---|---|---|---|---|---|---|
| Consumption | 777.62 | 371.88 | 70.00 | 515.00 | 717.00 | 987.00 | 2,265.00 |
| EO Consumption | 777.62 | 0.00 | 777.62 | 777.62 | 777.62 | 777.62 | 777.62 |
| EOp Consumption | 777.62 | 187.61 | 184.11 | 677.32 | 755.67 | 886.90 | 1,516.72 |
| POV Consumption | 797.70 | 346.75 | 430.20 | 515.00 | 717.00 | 987.00 | 2,265.00 |
| Health score | 65.21 | 18.73 | 15.00 | 55.00 | 70.00 | 80.00 | 95.00 |
| EO Health score | 65.21 | 0.00 | 65.21 | 65.21 | 65.21 | 65.21 | 65.21 |
| EOp Health score | 66.51 | 11.78 | 42.00 | 63.12 | 70.63 | 73.93 | 95.00 |
| POV Health score | 66.69 | 15.96 | 42.00 | 55.00 | 70.00 | 80.00 | 95.00 |
| WTP perfect health | 88.18 | 170.41 | 0.00 | 0.00 | 25.00 | 100.00 | 2,000.00 |

**Description**: The table reports the descriptive statistics of the information used to compute Equivalent Advantages

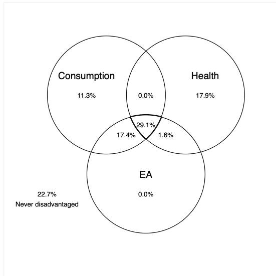
# Appendix B Share of disadvantaged individuals

Figure 3: Share of disadvantaged individuals

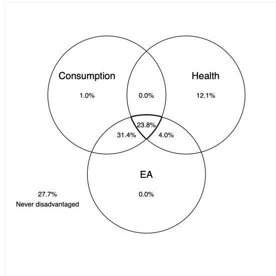
(a) EO - LB



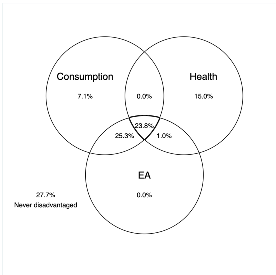

(b) EO - UB



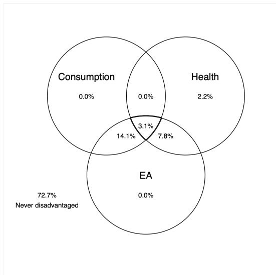

(c) EOp - LB



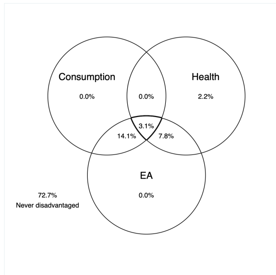

(d) EOp - UB




(e) POV - LB




(f) POV - UB




**Description**: The Venn diagrams show the intersection between the set of individuals with consumption and/or health below the fair one, and the set of individuals with negative EA. We present the results for both the lower (LB) and the upper (UB) bounds of Equivalent Advantage for each norm distribution.

# Appendix C Risk of disadvantage

Table 9: Risk of being disadvantaged in the EOp norm.

| | (1) Gender | (2) Age | (3) Employment status | (4) Region of residence |
|---|---|---|---|---|
| Woman | 0.185** | | | |
| | (0.076) | | | |
| 18-29 | | 0.152 | | |
| | | (0.146) | | |
| 30-39 | | 0.240* | | |
| | | (0.127) | | |
| 50-59 | | 0.054 | | |
| | | (0.121) | | |
| 60-69 | | -0.025 | | |
| | | (0.127) | | |
| 70-79 | | 0.231 | | |
| | | (0.144) | | |
| 80+ | | 0.972*** | | |
| | | (0.200) | | |
| Unemployed | | | 0.671*** | |
| | | | (0.167) | |
| Retired | | | 0.459*** | |
| | | | (0.080) | |
| Flanders | | | | -0.139 |
| | | | | (0.145) |
| Wallonia | | | | -0.214 |
| | | | | (0.152) |
| Constant | -0.093* | -0.132 | -0.211*** | 0.158 |
| | (0.056) | (0.087) | (0.051) | (0.137) |

Standard errors in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Description**: The table shows the results of a logistic regression where the dependent variable takes value 1 if the EA upper bound is lower than zero. We compute robust standard errors. The reference categories are (1) Gender: Male; (2) Age: 40-49; (3) Employment status: Employed; (4) Region of residence: Brussels.

Table 10: Risk of being disadvantaged in the EOp norm. Interaction effects.

| | (1) All | (2) Bruxelles | (3) Flanders | (4) Wallonia |
|---|---|---|---|---|
| Male Young Employed | -0.022 | -1.540 | -0.004 | 0.216 |
| | (0.209) | (1.121) | (0.257) | (0.401) |
| Female Young Employed | 0.548*** | -0.665 | 0.756*** | 0.349 |
| | (0.210) | (0.888) | (0.266) | (0.389) |
| Male Young Non-Employed | -0.159 | 0.657 | 0.268 | -1.015 |
| | (0.366) | (0.960) | (0.545) | (0.669) |
| Female Young Non-Employed | 0.884** | 0.000 | 1.654** | 0.285 |
| | (0.354) | (.) | (0.798) | (0.452) |
| Female Adult Employed | 0.091 | 0.442 | 0.143 | -0.095 |
| | (0.114) | (0.426) | (0.144) | (0.209) |
| Male Adult Non-Employed | 0.637*** | 0.944 | 0.991*** | 0.285 |
| | (0.215) | (0.916) | (0.322) | (0.320) |
| Female Adult Non-Employed | 1.136*** | 1.414** | 0.974*** | 1.304*** |
| | (0.170) | (0.591) | (0.226) | (0.290) |
| Male Old Employed | -0.422 | -0.442 | -0.320 | -0.632 |
| | (0.327) | (1.262) | (0.408) | (0.610) |
| Female Old Employed | -0.134 | 0.251 | -0.020 | -0.563 |
| | (0.355) | (1.044) | (0.453) | (0.706) |
| Male Old Non-Employed | 0.481*** | 0.944** | 0.567*** | 0.136 |
| | (0.130) | (0.459) | (0.165) | (0.244) |
| Female Old Non-Employed | 0.351*** | 0.057 | 0.244 | 0.579*** |
| | (0.125) | (0.465) | (0.163) | (0.217) |
| Constant | -0.271*** | -0.251 | -0.268** | -0.285* |
| | (0.082) | (0.292) | (0.104) | (0.149) |
| Observations | 2778 | 210 | 1662 | 901 |

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**Description**: The table shows the results of a logistic regression where the dependent variable takes value 1 if the EA upper bound is lower than zero. We compute robust standard errors. Model (1) considers the entire sample. Models (2) - (4) focus on specific regions of residence. The baseline category is Male Adult Employed. **Note**: Female Young Non-Employed omitted in model (2) because all are disadvantaged.

Table 11: Risk of being disadvantaged in the POV norm.

| | (1) Gender | (2) Age | (3) Employment status | (4) Region of residence |
|---|---|---|---|---|
| Woman | 0.347*** | | | |
| | (0.089) | | | |
| 18-29 | | 0.134 | | |
| | | (0.169) | | |
| 30-39 | | 0.155 | | |
| | | (0.147) | | |
| 50-59 | | 0.008 | | |
| | | (0.142) | | |
| 60-69 | | -0.114 | | |
| | | (0.151) | | |
| 70-79 | | 0.192 | | |
| | | (0.165) | | |
| 80+ | | 0.708*** | | |
| | | (0.199) | | |
| Unemployed | | | 1.344*** | |
| | | | (0.169) | |
| Retired | | | 0.857*** | |
| | | | (0.094) | |
| Flanders | | | | -0.261 |
| | | | | (0.160) |
| Wallonia | | | | -0.248 |
| | | | | (0.168) |
| Constant | -1.287*** | -1.183*** | -1.559*** | -0.858*** |
| | (0.068) | (0.102) | (0.067) | (0.149) |

Standard errors in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Description**: The table shows the results of a logistic regression where the dependent variable takes value 1 if the EA upper bound is lower than zero. We compute robust standard errors. The reference categories are (1) Gender: Male; (2) Age: 40-49; (3) Employment status: Employed; (4) Region of residence: Brussels.

Table 12: Risk of being disadvantaged in the POV norm. Interaction effects.

| | (1) All | (2) Bruxelles | (3) Flanders | (4) Wallonia |
|---|---|---|---|---|
| Male Young Employed | 0.119 | 0.000 | -0.170 | 1.052** |
| | (0.282) | (.) | (0.385) | (0.456) |
| Female Young Employed | 0.431* | -0.693 | 0.652** | 0.201 |
| | (0.261) | (1.133) | (0.315) | (0.534) |
| Male Young Non-Employed | 1.568*** | 0.693 | 1.220** | 2.138*** |
| | (0.368) | (0.974) | (0.577) | (0.582) |
| Female Young Non-Employed | 1.804*** | 2.485** | 2.213*** | 1.482*** |
| | (0.348) | (1.169) | (0.662) | (0.484) |
| Female Adult Employed | 0.383** | -0.201 | 0.433** | 0.483* |
| | (0.151) | (0.504) | (0.194) | (0.280) |
| Male Adult Non-Employed | 1.655*** | 1.099 | 1.767*** | 1.690*** |
| | (0.226) | (0.884) | (0.322) | (0.356) |
| Female Adult Non-Employed | 1.824*** | 1.386** | 1.879*** | 1.874*** |
| | (0.178) | (0.554) | (0.240) | (0.308) |
| Male Old Employed | -0.329 | 0.405 | -0.313 | -0.715 |
| | (0.488) | (1.272) | (0.629) | (1.060) |
| Female Old Employed | -0.617 | 0.000 | -1.188 | -0.347 |
| | (0.615) | (1.205) | (1.036) | (1.076) |
| Male Old Non-Employed | 0.702*** | 0.278 | 0.947*** | 0.172 |
| | (0.162) | (0.493) | (0.203) | (0.341) |
| Female Old Non-Employed | 0.854*** | 0.043 | 0.977*** | 0.878*** |
| | (0.154) | (0.530) | (0.201) | (0.277) |
| Constant | -1.750*** | -1.099*** | -1.807*** | -1.850*** |
| | (0.114) | (0.334) | (0.148) | (0.215) |
| Observations | 2778 | 208 | 1662 | 901 |

Standard errors in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Description**: The table shows the results of a logistic regression where the dependent variable takes value 1 if the EA upper bound is lower than zero. We compute robust standard errors. Model (1) considers the entire sample. Models (2) - (4) focus on specific regions of residence. The baseline category is Male Adult Employed. **Note**: Male Young Employed omitted in model (2) because all are advantaged.

# Appendix D Generalized Lorenz curves

Table 13: Generalized Lorenz dominance of EA distributions by population subgroups. Both lower and upper bounds.

| | | EO | EOp | POV |
|---|---|---|---|---|
| Gender | | | | |
| Men - Women | Lower-bound | I (>) | > | I (>) |
| | Upper-bound | > | I (>) | > |
| Age | | | | |
| Adult - Young | Lower-bound | I | I | < |
| | Upper-bound | I | I | I (>) |
| Adult - Old | Lower-bound | > | > | > |
| | Upper-bound | > | > | > |
| Young - Old | Lower-bound | I | > | > |
| | Upper-bound | > | > | I |
| Region | | | | |
| Bxl - Fla | Lower-bound | < | < | < |
| | Upper-bound | I (<) | I (<) | I (<) |
| Bxl - Wal | Lower-bound | < | < | < |
| | Upper-bound | < | < | < |
| Fla - Wal | Lower-bound | I (<) | < | < |
| | Upper-bound | I (<) | I | I |
| Employment status | | | | |
| Empl - Unempl | Lower-bound | I (>) | I | > |
| | Upper-bound | I (>) | I (>) | > |
| Empl - Retired | Lower-bound | > | > | > |
| | Upper-bound | > | > | > |
| Unempl - Retired | Lower-bound | I | I | I |
| | Upper-bound | I | I | I (<) |

**Description**: The table shows the results of the generalized Lorenz ranking between EA distributions. The sign > indicates that the fist group dominates the second one while < denotes the opposite case. If the curves are crossing we put the letter I and we perform another comparison between the generalized Lorenz curves for $p$ in eq. (1) between 0.1 and 0.9. The results (if different) are reported in parenthesis.

Figure 4: Generalized Lorenz curves of EA by gender.

(a) EO (b) EOp (c) POV

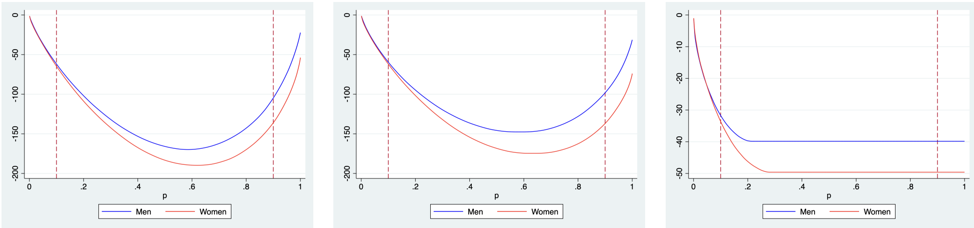



Figure 5: Generalized Lorenz curves of EA by age class.

(a) EO (b) EOp (c) POV

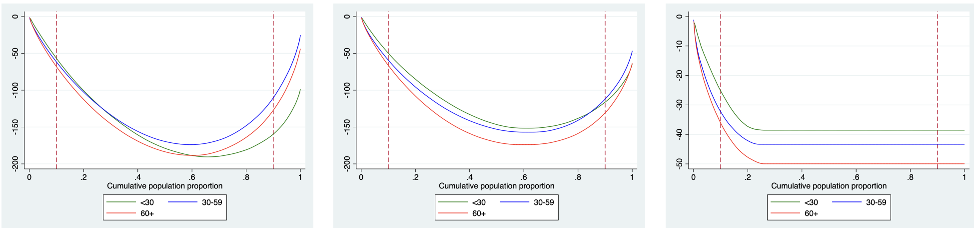



Figure 6: Generalized Lorenz curves of EA by region.

(a) EO (b) EOp (c) POV

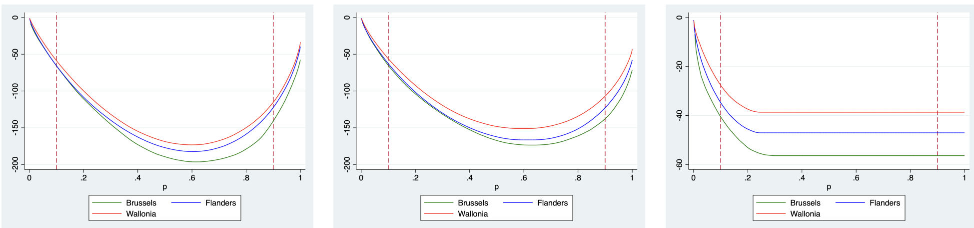
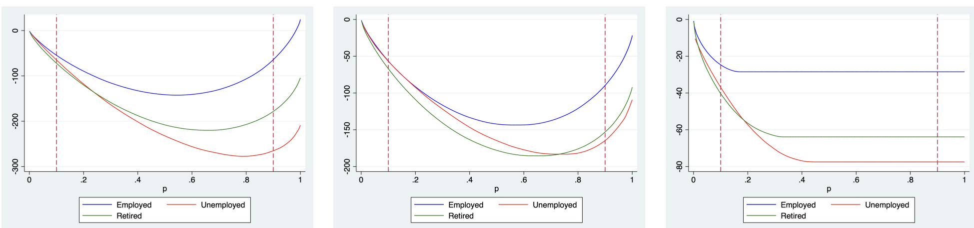



Figure 7: Generalized Lorenz curves of EA by employment status.

(a) EO (b) EOp (c) POV



# Appendix E The role of preferences

Table 14: The role of preferences.

| | (1) EO(lb) | (2) EO(ub) | (3) EOp(lb) | (4) EOp(ub) | (5) POV(lb) | (6) POV(ub) |
|---|---|---|---|---|---|---|
| Consumption | 0.962*** | 1.083*** | 0.928*** | 1.038*** | 1.137*** | 0.987*** |
| | (0.010) | (0.025) | (0.010) | (0.019) | (0.035) | (0.006) |
| Health | 5.024*** | 11.988*** | 5.382*** | 12.236*** | 8.619*** | 2.184*** |
| | (0.195) | (0.488) | (0.262) | (0.512) | (0.405) | (0.071) |
| Norm consumption | -1.419*** | -1.884*** | -1.069*** | -1.016*** | -1.172*** | -0.991*** |
| | (0.019) | (0.048) | (0.018) | (0.035) | (0.036) | (0.006) |
| Norm Health | 0.000 | 0.000 | -4.285*** | -11.016*** | -8.328*** | -2.142*** |
| | (.) | (.) | (0.323) | (0.631) | (0.420) | (0.073) |
| Observations | 2778 | 2778 | 2778 | 2778 | 2778 | 2778 |
| $R^2$ | 0.792 | 0.514 | 0.778 | 0.572 | 0.432 | 0.927 |

Standard errors in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

**Description**: The table shows the results of an OLS regression with EA as dependent variable, for each norm distribution and both lower and bound.

Table 15: Generalized Lorenz dominance of ASD distributions by population subgroups.

| | EO | EOp | POV |
|---|---|---|---|
| **Gender** | | | |
| Men - Women | > | > | > |
| **Age** | | | |
| Adult - Young | I ( <) | I (>) | < |
| Adult - Old | I (>) | > | I (>) |
| Young - Old | > | I | I (>) |
| **Region** | | | |
| Bxl - Fla | I | I (<) | I |
| Bxl - Wal | I | I (<) | I |
| Fla - Wal | I | I | > |
| **Employment status** | | | |
| Empl - Unempl | I (>) | > | > |
| Empl - Retired | > | > | > |
| Unempl - Retired | I (>) | I | I (<) |

**Description**: The table replicates the comparison in Table 6 using ASD rather than EA. **Note:** $ASD_i = 0.5 \cdot S(c_i - c_i^\star) + 0.5 \cdot S(h_i - h_i^\star)$, with $S(x) = (x - \bar{x})/\sigma(x)$, where $\bar{x}$ is the average and $\sigma(x)$ is the standard deviation of $x$.