IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10940

Equality of Opportunity for Well-Being

Daniel Gerszon Mahler
Xavier Ramos

AUGUST 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10940

Equality of Opportunity for Well-Being

Daniel Gerszon Mahler
University of Copenhagen

Xavier Ramos
Universitat Autonoma de Barcelona and IZA

AUGUST 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Equality of Opportunity for Well-Being*

A growing literature has tried to measure the extent to which individuals have equal opportunities to acquire income. At the same time, policy makers have doubled down on efforts to go beyond income when measuring well-being. We attempt to bridge these two areas by measuring the extent to which individuals have equal opportunities to achieve a high level of well-being. We use the German Socio-Economic Panel to measure well-being in four different ways including incomes. This makes it possible to determine if the way well-being is measured matters for identifying who the opportunity-deprived are and for tracking inequality of opportunity over time. We find that, regardless of how well-being is measured, the same people are opportunity-deprived and equality of opportunity has improved over the past 20 years. This suggests that going beyond income has little relevance if the objective is to provide equal opportunities.

JEL Classification: D3, D63, I31

Keywords: equality of opportunity, measurement, responsibility, effort, well-being

Corresponding author:
Xavier Ramos
Universitat Autònoma de Barcelona
Depart Econ Aplicada
Campus UAB
Bellaterra 08193
Spain
E-mail: xavi.ramos@uab.cat

* We are very grateful to Brice Magdalou for comments and advice on the use of norm-based inequality metrics, to Dirk Van de gaer for fruitful discussions and comments on several parts of the paper, to participants of the Seventh ECINEQ meeting in New York, the 2017 IARIW-Bank of Korea Conference \Beyond GDP: Past Experiences and Future Challenges in the Measurement of Economic Well-being", and the 2017 Winter School on Inequality and Social Welfare Theory.

1 Introduction

The notion that individuals ought to have equal opportunities in life is popular among politicians, the general public, and philosophers alike. A sizable number of empirical studies have been carried out analyzing the extent to which individuals have equal opportunities for income acquisition (see Ramos and Van de gaer (2016), Roemer and Trannoy (2015), and Ferreira and Peragine (2016) for excellent recent reviews). These studies are based on the idea that when evaluating the progress of societies, looking at the level and distribution of incomes provides an incomplete picture. A distinction has to be made between income differences arising from factors individuals ought to be held personally responsible for and income differences arising from factors outside the realm of personal responsibility.

Given the growing interest in going beyond income to measure individual well-being, it seems pertinent to apply this discussion to the equality of opportunity framework. Well-being (or welfare, we use the two interchangeably) is inherently multidimensional, and growth and income statistics fail to capture this multiplicity. For this reason, the philosophers who advocate for equality of opportunity do not advocate for equality of opportunity for income acquisition, but rather something broader than income such as welfare or advantage (Arneson, 1989; Cohen, 1990). If individuals ought to have equal opportunities for well-being, then using income as the acquisition variable in equality of opportunity studies could be problematic. Incomes ignore the disutility of effort, the well-being individuals receive from other dimensions of life, and the differences in preferences over income and these other dimensions.

Once it is acknowledged that incomes are not sufficient for measuring well-being, the door opens for many alternatives. Which other well-being dimensions are necessary? Should we measure these dimensions separately or somehow aggregate them to a single number? How can we incorporate the fact that individuals have different preferences over these various dimensions? Should we try to measure well-being directly by alluding to self-reported happiness levels?

We use 30 years of data from the German Socio-Economic Panel (GSOEP) to measure welfare in four ways; with incomes, life satisfaction, a multidimensional index, and equivalent incomes. We use incomes to facilitate comparisons with the generic way of measuring equality of opportunity. The other three measures have roots in different philosophical theories about well-being (Parfit, 1984; Griffin, 1986). Life satisfaction explicitly tries to measure mental states, the multidimensional index defines and aggregates an objective list of dimensions of importance for well-being, and equivalent incomes incorporate preference heterogeneity.

We will investigate if the measurement of welfare matters for 1) characterizing the opportunity-deprived and 2) tracking inequality of opportunity over time. In both cases, we first convert our welfare measures into welfare ranks. This minimizes our reliance on cardinality and assures that the different distributions the welfare measures follow do not drive the results. Next, we regress the welfare ranks on a set of effort and circumstance variables, or equivalently, variables we hold individuals responsible for and variables we do not hold individuals responsible for. Based on these regressions, we assign each individual a 'fair' rank and an 'unfair' rank. The fair ranks order individuals according to how much their effort variables contribute to their welfare, while the unfair ranks order individuals according to how much their circumstances contribute to their welfare. The latter ranks can be interpreted as an ordinal measure of individuals' opportunities, where the highest ranked possesses the best combination of circumstances and vice versa. With this procedure, for any given welfare measure, an individual will have a welfare rank, a fair rank, and an unfair rank.

In order to answer our first research question, if the measure of welfare matters for characterizing the opportunity-deprived, we compare the unfairness ranks across the four well-being measures. Broadly speaking, if individuals have similar unfairness ranks across the well-being measures, then a characterization of the opportunity-deprived will not depend on how we measure welfare.

In order to answer our second research question, if the measure of welfare matters for tracking inequality of opportunity over time, we use the norm-based approach (see Ramos and Van de gaer (2016) for more information on this approach). In our set-up, this implies using divergence measures to compare the divergence between the welfare ranks and either the fair ranks or the unfair ranks. If the welfare ranks and the fair ranks are highly related, then the individuals exerting the most effort also have the highest well-being, and there is a low presence of inequality of opportunity. If the welfare ranks and unfair ranks are highly related, then the individuals with the best circumstances also have the best outcomes, and there is a high presence of inequality of opportunity. We introduce a new summary statistic of equality of opportunity; the correlation between the welfare ranks and the unfairness ranks. If we consider income to be the outcome variable of interest and consider parental income as the only circumstance variable, then this summary statistic boils down to the Spearman's correlation between parents' and childrens' incomes. Hence, it generalizes a frequently used measure in mobility studies.

We find that the measure of well-being matters little for characterizing who the opportunity-deprived are. When tracking whether inequality of opportunity has changed over the past 30 years, we find that regardless of how welfare is measured, inequality of opportunity has decreased in the last two decades. These results are robust to using different divergence measures and, for the most part, to changing what we hold individuals responsible for. This is encouraging news for policy makers interested in providing equal opportunities while going beyond income, as they may broadly get things right if they proxy well-being with income.

To our knowledge, this is the first paper to explicitly address the Beyond GDP agenda in a Roemerian equality of opportunity framework. We are certainly not the first, however, to relate notions of fairness with welfare measurement. Fleurbaey and Maniquet (2011) summarizes extensive work on this topic. This prior work generally incorporates concerns about fairness directly into the well-being measure. Our approach, in contrast, first computes measures of welfare and then analyzes the extent to which factors beyond individual control are driving the welfare differences. A particularly relevant paper for our approach is Ravallion (2015), which incorporates the disutility of effort into estimates of inequality of opportunity. We go in a different direction by analyzing whether the concept of welfare matters for estimates of inequality of opportunity. In previous studies the measurement of welfare has been shown to matter for assessments of how welfare has developed over time (Blanchflower and Oswald, 2004), for how inequality in welfare has developed over time (Stevenson and Wolfers, 2008), and for identifying who the most welfare-deprived are (Decancq and Neumann, 2016).

Roemer (2012) explicitly argues against using welfare as the outcome variable in equality of opportunity estimations. He does so on the grounds that policy makers are interested in dimensions of well-being separately, such as health, income or education, rather than well-being itself. This may certainly be the case, but if the ultimate objective is to equalize opportunities for well-being, then equalizing opportunities for only one dimension of well-being might actually bring about the opposite result (Calsamiglia, 2009). To see this, consider a policy that targets people living outside of the main cities because they have fewer opportunities to acquire a high income. If these people simultaneously have better health, more leisure, or different preferences over the importance of income, they need not have less opportunities to acquire a high level of well-being. Our framework helps clarify if such examples have empirical leverage.

The rest of the paper is organized as follows. Section 2 explains both the philosophical and axiomatic theory behind measuring equality of opportunity for well-being. Section 3 details our data and measurement approach. Section 4 outlines the results and provides several robustness checks. Section 5 concludes.

2 Theory

2.1 Well-Being

Three overarching theories of well-being exist in the philosophical literature; objective list theory, preference satisfaction theory, and mental state theory (Parfit, 1984; Griffin, 1986). *Preference satisfaction theory* is the most commonly assumed in economics. It claims that an individual's welfare depends on the degree to which his preferences are satisfied. Often preference orderings are assumed to be revealed through choice behavior. The underlying tenet behind these revealed preferences is that if an agent chooses bundle A over bundle B, then the agent must prefer A over B, and the agent must be better off with A rather than B. *Mental state theory* takes its starting point in what goes on inside the mind of individuals rather than their observed choices. According to this theory, well-being is the degree to which individuals are happy or the extent to which they experience pleasure over pain. *Objective list theory* argues that individuals' lives go well to the degree that they are in possession of certain items on a list, which could be income, education, health, safety, etc.

In short, mental state theory care about what individuals feel, (revealed) preferences about what individuals choose, and objective list theory about external circumstances which could be independent of the choices or feelings of individuals. Each theory has its advantages and shortcomings. Preference satisfaction theory, at least in the revealed form, can be criticized when individuals' decision-making is subject to imperfect knowledge and behavioral biases. If individuals have mistaken beliefs about what is best for themselves or lack willpower to choose what is best for themselves, then there is little reason to believe that their choices are a good manifestation of their well-being. Mental state theory can be criticized for its 'physical condition neglect' (Sen, 1985), whereby individuals might feel well only because they have adapted to horrible conditions. In these scenarios, Sen argues, mental states are not an appropriate yardstick. Objective list theory can be criticized for being elitist in the sense that a set of indicators and weights are chosen somewhat independent of the preferences of individuals.[1]

This three-part division of well-being concepts is still very much in use today in both theoretical and empirical literature about well-being (see for example the chapter division in the Oxford Handbook of Well-Being and Public Policy (Adler and Fleurbaey, 2016) and the Stanford Encyclopedia of Philosophy on Well-Being (Crisp,

[1]All of these critiques can of course be counteracted, but doing so would be outside the scope of this paper.

2016)). We will operationalize a measure of well-being with roots in each of these theories and see if they lead to different conclusions about equality of opportunity. We are not attempting to argue in favor of any one of these welfare concepts. Rather, we will take some of the operationalizations of these concepts at face value and investigate if equality of opportunity estimations depend on which measure is used.

2.2 Distributive Justice

Until Rawls published his Theory of Justice (Rawls, 1971), the predominant view of justice was defined in utilitarian terms. Under this view, the best outcome is the one that maximizes total welfare or, equivalently, equalizes marginal utilities. This view is welfarist in the sense that if all individuals' welfare levels are known, then no additional information is needed to decide whether one scenario is more desirable than another. Rawls argued against this welfarist view, emphasizing that we should not seek to equalize marginal utilities, but rather primary goods, which is a broader notion that also encompasses rights and liberties.

A number of subsequent scholars proposed variations of what the right equalizandum ought to be, building on the work of Rawls. Sen (1980) argued that neither utilities nor primary goods were enough to judge outcomes. He concluded that we need to look at what individuals are capable of achieving with these goods, thus advocating for basic capability equality. Dworkin (1981) contended that resources is the right equalizandum, while Arneson (1989) and Cohen (1990) argued that the right equalizandum is, respectively, equality of opportunity for welfare and equal access to advantage (see Roemer and Trannoy (2016) for a more complete account on the developments in distributive justice since Rawls).

Although these philosophers differ in their preferred equalizandum, they all adhere to the point of view that knowing all individuals' welfare levels is not sufficient. We need to know how that welfare came about, whether it came from fortunate backgrounds or from factors we can hold individuals accountable for. As such, they agree on the need to go beyond welfarism and accept some degree of individual responsibility, and thereby some degree of just inequalities. Notably, none of the philosophers defined the equalizandum in terms of income. Rather, they considered broader notions than income such as welfare, advantage, or functionings. Our approach attempts to get a bit closer to these frameworks. In particular, our approach is closely related to that of Arneson (1989), who precisely argued for equalization of opportunities for welfare.[2]

[2]That being said, Arneson (1989) considered welfare to be preference satisfaction, thus differing from our take, where we will look at different theories of welfare.

2.3 Equality of Opportunity

The philosophical theories of distributive justice have been operationalized in economics through the works of Roemer (1993), Van de gaer (1993), and Fleurbaey (1994) amongst others. The starting point in many of these operationalizations is to consider a population, $\mathcal{N} = \{1, 2, \ldots, n\}$, and a distribution of an outcome variable for this population, $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$. Often $\boldsymbol{y}$ is considered to be income, but here we will take welfare/well-being as the outcome, such that y_i is the welfare of individual $i \in \mathcal{N}$. An individual's outcome is assumed to be a product of two sets of variables: circumstances, $\boldsymbol{a^C}$, and effort, $\boldsymbol{a^E}$. Circumstances are the factors outside the realm of control for the individual, the factors one ought not to hold an individual responsible for. These are often taken to be gender, region of birth, parental education, parental income etc. Effort variables are the factors one ought to hold an individual responsible for. The well-being of individual i is thus assumed to be given by $y_i = f(\boldsymbol{a_i^C}, \boldsymbol{a_i^E})$.

Inequality of opportunity is frequently measured using either the direct approach or the indirect approach. The *direct approach* obtains a counterfactual distribution of the outcome variable, which only depends on circumstances, and therefore only incorporates unfair variation. This can be obtained by fixing the effort variables at a certain level, $\boldsymbol{\tilde{a}^E}$. Inequality of opportunity is then measured as the inequality in the resulting counterfactual distribution, $I\big(f(\boldsymbol{a^C}, \boldsymbol{\tilde{a}^E})\big)$. The *indirect approach* obtains a counterfactual outcome distribution that only depends on effort, and therefore only incorporates fair variation. This can be done by fixing the circumstance variables at a given level, $\boldsymbol{\tilde{a}^C}$. Inequality of opportunity is then given as the difference between overall inequality and the inequality in this counterfactual distribution, $I(\boldsymbol{y}) - I\big(f(\boldsymbol{\tilde{a}^C}, \boldsymbol{a^E})\big)$.[3]

The axiomatic literature on fair allocations has put forward two criteria which inequality of opportunity estimates ideally ought to reflect, these being the compensation principle and the reward principle. The *compensation principle* states that differences in well-being due to differences in circumstances should be eliminated. The *reward principle* is concerned with the proper reward of effort for individuals with the same circumstances. Unfortunately, these two criteria are mutually incompatible (Bossert, 1995; Fleurbaey, 1995) unless one assumes that the outcome is linearly separable in circumstance and effort, such that $y_i = g(\boldsymbol{a_i^C}) + h(\boldsymbol{a_i^E})$. We will follow much of the literature and make that assumption.

We want to minimize our reliance on cardinality assumptions, and hence in-

[3]See Ramos and Van de gaer (2015) for more methods to obtain counterfactual distributions using the direct and the indirect approach.

stead of using the welfare levels as the outcome variable, we convert them into welfare ranks, $r_i^y = F(y_i)$, where F is the cumulative distribution function. We likewise convert our counterfactual distributions into ranks, such that $r_i^{unfair} = G(\boldsymbol{a_i^C})$ and $r_i^{fair} = H(\boldsymbol{a_i^E})$ where G and H are the cumulative distribution of $g(\boldsymbol{a^C})$ and $h(\boldsymbol{a^E})$, respectively. The fair ranks order individuals according to how much their effort variables contribute to their welfare, while the unfair ranks order individuals according to how much their circumstances contribute to their welfare. Within this set-up, neither the direct approach nor the indirect approach is feasible, as the inequalities in rank variables are uninteresting and identical. In other words, $I(\boldsymbol{r^y}) = I(\boldsymbol{r^{unfair}}) = I(\boldsymbol{r^{unfair}})$. Instead, we will use the norm-based approach.

2.4 Norm-Based Inequality Metrics

The norm-based approach (Ramos and Van de gaer, 2016), also called the *fairness gap* (Fleurbaey and Schokkaert, 2009), evaluates equality of opportunity by measuring the divergence between the actual outcomes and 'fair' outcomes. The fair outcomes only depend on effort variables and reflect how much each individual ideally is entitled to under some allocation principles. The more aligned these two distributions are, the lower inequality of opportunity. In our context, the fairness gap can be calculated by employing a divergence measure, $D(\boldsymbol{r^y} \| \boldsymbol{r^{fair}})$, which evaluates the divergence between the two distributions, $\boldsymbol{r^y}$ and $\boldsymbol{r^{fair}}$.

An intuitive measure of the (absence of) divergence between $\boldsymbol{r^y}$ and $\boldsymbol{r^{fair}}$ is simply their correlation. A correlation of 1 suggests complete equality of opportunity with respect to rewarding effort; individuals with higher effort levels also have higher levels of welfare. A correlation of 0, in contrast, suggests no equality of opportunity with respect to the rewarding effort.

Due to our rank conversion, this divergence measure will not satisfy the compensation principle. However, we can construct a parrallel measure which uses $\boldsymbol{r^{unfair}}$ instead of $\boldsymbol{r^{fair}}$ that does reflect this principle. Hence, we will likewise consider the correlation between $\boldsymbol{r^y}$ and $\boldsymbol{r^{unfair}}$. A high correlation between $\boldsymbol{r^y}$ and $\boldsymbol{r^{unfair}}$ suggests a high degree of inequality of opportunity with respect to the compensation principle; individuals with the best circumstances also have the highest welfare levels. A correlation of 0, in turn, reflects that the compensation principle is completely fulfilled; the quality of individuals' circumstances is not correlated with their welfare.

If we use income as the outcome variable and consider parental income as the only circumstance variable, then the correlation between $\boldsymbol{r^y}$ and $\boldsymbol{r^{unfair}}$ boils down to a frequently used measure of mobility; Spearman's correlation between parents' and children's income level (see for example Chetty *et al.* (2014)). We generalize this measure since we consider more than one circumstance when constructing the unfairness ranks and apply other outcome variables than income.

Magdalou and Nock (2011) (MN) provide a framework for more axiomatically grounded divergence measures, which we will use as robustness checks.[4] They build on function ϕ, for all $c \in \mathbb{R}_{++}$:

$$\phi(c) \coloneqq \begin{cases} \frac{1}{s(s-1)}\, c^r\,, & \text{if} \quad s \neq 0, 1\,, \\ c\, \ln c\,, & \text{if} \quad s = 1\,, \\ -\,\ln c\,, & \text{if} \quad s = 0\,. \end{cases} \tag{1}$$

to put forth the general class of divergence measures between an outcome distribution, $\boldsymbol{y}$, and a reference distribution, $\boldsymbol{z}$:

$$D_{MN}(\boldsymbol{y}\|\boldsymbol{z}) = \frac{1}{n}\sum_{i\in\mathcal{N}} \left[\phi(y_i) - \phi(z_i) - (y_i - z_i)\phi'(z_i)\right], \tag{2}$$

which satisfies partial symmetry along with other relevant properties. The class $D_{MN}(\boldsymbol{y}\|\boldsymbol{z})$ is suitable only for distributions with equal means. This is not a problem in our set-up since the rank-based measures that we will use in place of $\boldsymbol{y}$ and $\boldsymbol{z}$ by construction have equal means.[5] By using the function ϕ in (1), one obtains

$$D_{MN}(\boldsymbol{y}\|\boldsymbol{z}) = \begin{cases} \frac{1}{n}\frac{1}{s(s-1)} \sum_{i\in\mathcal{N}} \left[y_i^s + (s-1)z_i^s - s\, y_i\, z_i^{s-1}\right], & \text{if} \quad s \neq 0, 1\,, \\ \frac{1}{n}\sum_{i\in\mathcal{N}} \left[y_i \,\ln\left({}^{y_i}/_{z_i}\right)\right], & \text{if} \quad s = 1\,, \\ \frac{1}{n}\sum_{i\in\mathcal{N}} \left[{}^{y_i}/_{z_i} - \ln\left({}^{y_i}/_{z_i}\right) - 1\right], & \text{if} \quad s = 0\,. \end{cases} \tag{4}$$

Cowell (1985) suggests a different class of divergence measures, which he calls measures of distributional change, that satisfies different properties. The population

[4] We are heavily indebted for comments and advice from Brice Magdalou on the use and interpretation of appropriate divergence measures.

[5] Had not our actual and norm distributions had the same mean we could have normalized our divergence class further to obtain a (strong) scale invariant class:

$$D^m_{MN}(\boldsymbol{y}\|\boldsymbol{z}) = \frac{1}{n}\sum_{i\in\mathcal{N}} \left[\phi(\hat{y}_i) - \phi(\hat{z}_i) - (\hat{y}_i - \hat{z}_i)\phi'(\hat{z}_i)\right], \tag{3}$$

where $\hat{y}_i = y_i/\mu(\boldsymbol{y})$, $\hat{z}_i = z_i/\mu(\boldsymbol{z})$, and $\mu(\boldsymbol{z}) = \sum_{i=1}^n z_i/n$. It is worth noting that this class boils down to the generalized entropy class of standard inequality measures if the reference distribution is assumed to be the mean of the actual distribution.

invariant measure equivalent to (2) is:[6]

$$D_C(\boldsymbol{y}\|\boldsymbol{z}) = \frac{1}{n}\sum_{i\in\mathcal{N}} [z_i\, \phi(y_i/z_i)]\,. \tag{5}$$

By using the function ϕ in (1), $D_C(\boldsymbol{y}\|\boldsymbol{z})$ can be written as

$$D_C(\boldsymbol{y}\|\boldsymbol{z}) = \begin{cases} \frac{1}{n}\,\frac{1}{s(s-1)}\,\sum_{i\in\mathcal{N}} \left[y_i^s z_i^{1-s} - 1\right]\,, & \text{if} \quad s \neq 0, 1\,, \\ \frac{1}{n}\,\sum_{i\in\mathcal{N}} y_i\,\ln\left[{}^{y_i}/_{z_i}\right]\,, & \text{if} \quad s = 1\,, \\ \frac{1}{n}\,\sum_{i\in\mathcal{N}} z_i\,\ln\left[{}^{z_i}/_{y_i}\right]\,, & \text{if} \quad s = 0\,. \end{cases} \tag{6}$$

The two different classes of divergence measures, $D_{MN}(\boldsymbol{y}\|\boldsymbol{z})$ and $D_C(\boldsymbol{y}\|\boldsymbol{z})$, coincide for one – and unique – parameter value, $s = 1$. For this reason we are going to use $s = 1$ as one of our robustness checks. Parameters $s = 0$ and $s = 2$ with $D_{MN}(\boldsymbol{y}\|\boldsymbol{z})$ will be used as other robustness checks. Why these two values? One of the features of D_{MN} is that a progressive transfer in the actual distribution $\boldsymbol{y}$ reduces the divergence between $\boldsymbol{y}$ and $\boldsymbol{z}$ as long as the individuals involved in the transfer have the same reference, $\boldsymbol{z}$. It is a kind of priority given to the worse-off individuals when the individuals involved in the transfer share the same $\boldsymbol{z}$. Moreover, if (and only if) $s < 2$, the further down the distribution $\boldsymbol{y}$ such transfer takes place, the more the divergence between $\boldsymbol{z}$ and $\boldsymbol{y}$ is reduced. This property resembles the principle of diminishing transfers in the context of inequality measurement, which holds for the class of entropy indices when $s < 2$. When $s = 2$ the measure is ordinally equivalent to the Euclidian distance, and it is thus insensitive to the position on the distribution where the progressive transfer (among individuals with the same reference) takes place. Thus, the parameter value $s = 2$ can be seen as a threshold. Contrary to this, the parameter value $s = 0$ yields a measure that is more sensitive to transfers lower down the distribution than our baseline measure with $s = 1$.

As another robustness check we will also use a divergence measure which is a generalization of the standard Gini coefficient developed by Almås *et al.* (2011), called "the Unfairness Gini", $D_{Gini}(\boldsymbol{y}\|\boldsymbol{z})$:

$$D_{Gini}(\boldsymbol{y}\|\boldsymbol{z}) = \frac{1}{2n(n-1)\mu(\boldsymbol{y})}\sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} |(y_i - z_i) - (y_j - z_j)| \tag{7}$$

[6] A scale invariant measure can be obtained by replacing y and z by $\hat{y}$ and $\hat{z}$ in (5). Devooght (2008) provides an empirical application of this measure to equality of opportunity in Belgium.

3 Data & Measurement

We use data from the German Socio-Economic Panel (GSOEP), which is a yearly panel that started in 1984. The panel contains detailed questions on household income, life satisfaction, other well-being dimensions, as well as biographical and historical data that we use to construct circumstance variables. We use data from 1984-2014 and include all working and unemployed individuals but drop individuals outside the labor market and observations with missing values. In total, we have 176,196 person-year observations meeting our baseline specification. These are spread around 21,038 individuals in 15,067 different households.

Our baseline analysis will use the following circumstance variables: gender, father's education (3 categories), mother's education (3 categories), father's occupation (6 categories), polynomial of age, height, place of birth (West Germany, East Germany, abroad), degree of urbanization at place of birth (4 categories), and number of siblings. As baseline effort variables we use years of education, work hours, a dummy for whether the respondent is self-employed, and a dummy for whether the respondent works in the public sector. 'Effort' may be a slightly misleading term here. The point of these four variables is that they plausibly lie within individual control and hence constitute factors that we may hold individuals accountable for. Summary statistics of the circumstance and effort variables are given in Table 1.

[Table 1 around here]

3.1 Constructing Welfare Variables

We use four different welfare variables in the analysis. First, we use incomes. This is the most frequently used outcome variable in equality of opportunity studies. We use it as a baseline for comparison to the other well-being measures. We use annual net household income expressed in 2010 constant EUR. The other three welfare variables are rooted in the three concepts of well-being that Parfit (1984) and Griffin (1986) put forward.

The second welfare measure we use is life satisfaction, which has roots in mental state theory. Life satisfaction is the answer to the question, 'How satisfied are you with your life, all things considered?' The answer categories range from 0 (completely dissatisfied) to 10 (completely satisfied). For the purpose of this study, we consider the answers interpersonally comparable. This is not meant as an endorsement of this particular account of well-being but rather as an inquiry into how inequality of opportunity estimates would look if one accepted these assumptions.

The third welfare measure we use is a multidimensional welfare measure, which has roots in objective list theory. To construct the measure of multidimensional welfare we partly follow Decancq and Neumann (2016). We consider four dimensions; income, health, leisure, and unemployment.[7] Income is measured the same way as above but converted into logs. Unemployment is a binary variable taking the value of 1 if the respondent had a job at the time of the survey. Leisure is measured as the amount of daily hours spent on leisure (capped at 6 hours). Health is itself a composite index composed of 1) an indicator for whether the individual is disabled, 2) the number of doctor appointments the respondent had last year and 3) the number of inpatient nights in hospitals the respondent had last year. To aggregate these sub-dimensions into one health dimension we regress a health satisfaction question on the three variables and use the coefficients as weights. The health satisfaction variable is composed of answers to how satisfied individuals are with their health on a scale from 0 (not at all satisfied) to 10 (completely satisfied). For the income, leisure, and health dimension, we standardize the values such that the highest possible level is 1 and the lowest possible level is 0. Now we have four dimensions each bounded between 0 and 1. To arrive at the final multidimensional index, we simply add these four together.

The fourth welfare measure, equivalent incomes, is based on preference satisfaction theory. In short, equivalent incomes are the incomes individuals need together with a reference bundle, to make them indifferent with their actual bundle. Although preferences are often estimated from choice behavior, this is difficult if the arguments cover a wide array of dimensions of well-being. An alternative method to recover preferences used by Decancq *et al.* (2015), is to regress life satisfaction on the dimensions of well-being and interpret the weights as marginal rates of substitution. The resulting utility functions seem to be highly correlated with the utility functions one would recover from choice behavior (Akay *et al.*, 2015). This approach easily accommodates preference heterogeneity by simply allowing for interactions between sociodemographic characteristics, w, and the various dimensions, dim. We follow this approach and use the following subset of the circumstance and effort variables as preference heterogeneity parameters, w: birth location, sex, age, age^2, education, work hours, self-employed, public sector worker. As the non-income dimensions, dim, we consider health, employment, and leisure, like in the multidimensional in-

[7]Although we would like to include more dimensions, such as education, we run into estimation problems since this also is considered an effort variable. As we will regress the welfare variable on circumstance and effort variables, and since we do not want to have the same variables on each side of the regression, we omit this dimension.

dex. Our life satisfaction regression looks as follows:

$$lifesat_{it} = [\beta^{inc} + \gamma^{inc} w_{it}] ln(inc_{it}) + [\beta^{dim} + \gamma^{dim} w_{it}] dim_{it} + \mu_t + \alpha_i + \varepsilon_{it} \tag{8}$$

In order to calculate the equivalent incomes, we first select a reference vector, $\tilde{dim}$, of all other dimensions than income. Here we choose the mean outcome (mode for categorical variables), since this avoids favoring any extreme marginal rates of substitution. Then we calculate the income needed together with $\tilde{dim}$ for this joint bundle to make individuals indifferent with their actual bundle. That is, we isolate inc_{it}^{eq} below:

$$\begin{aligned} lifesat(inc_{it}, dim_{it}) &= lifesat(inc_{it}^{eq}, \tilde{dim}) \\ \Leftrightarrow inc_{it}^{eq} &= exp\Big(ln(inc_{it}) + \frac{\beta^{dim} + \gamma^{dim} w_{it}}{\beta^{inc} + \gamma^{inc} w_{it}} (dim_{it} - \tilde{dim})\Big) \end{aligned} \tag{9}$$

The final result is an interpersonally comparable measure of individual welfare that takes differences in preferences into account.

By employing equivalent incomes as a welfare measure in our analysis, we are implicitly taking sides in a rich philosophical debate about whether individuals should be held responsible for their preferences. Our baseline approach deems differences in well-being arising from preference heterogeneity unfair if they stem from circumstance variables, but fair if they stem from effort variables. This is in contrast to most applications of equivalent incomes, which hold that individuals should be responsible for all their preferences. We can amend our approach such that individuals are held responsible for all their preferences by decomposing the equivalent income measure into a part that is due to preference heterogeneity and a part that does not incorporate preferences. We will do so as a robustness check. A detailed discussion of this method along with an illustration is given in Appendix A.1.

Histograms of the four final welfare measures (with the incomes and equivalent incomes expressed in logs) are presented in Figure 1. We turn these welfare measures into welfare ranks, $\boldsymbol{r^y}$, and break ties by adding a small amount of random noise to the well-being levels.

[Figure 1 around here]

3.2 Estimating Equality of Opportunity

For our empirical specification, we consider the well-being ranks to be linear in effort and circumstance variables:

$$r^y_{it} = \beta^C a^C_{it} + \beta^E a^E_{it} + \epsilon_{it} \tag{10}$$

Two important issues remain unsettled. First is the issue of how to interpret the error term, ϵ_{it}. The error contains omitted effort variables, omitted circumstance variables, measurement error, and general uncertainty. It is unclear whether this should be considered within individual control. This is an important decision as it accounts for most of the variation in the welfare ranks. In our baseline specification, we consider it an effort variable, but as a robustness check we shift it to the other side of the responsibility cut.

The other unsettled issue is what to do with the correlation between the effort and circumstance variables. Individuals' effort levels are partly determined by their own choices and partly by their circumstances. Years of education, for example, is partly influenced by individuals' social background and partly by an individual's own choices. We follow Roemer's approach and consider this correlation to be outside the realm of individual responsibility (Roemer, 1998).[8] In practice this means that prior to estimating the impact of circumstances and efforts on well-being we perform an auxiliary regression of the following form:

$$a^E_{it} = \gamma a^C_{it} + \eta_{it} \tag{11}$$

We perform such a regression for each effort variable and use the residuals from these regressions as our effort variables in our main regression, which then becomes:

$$r^y_{it} = (\beta^C + \gamma\beta^E) a^C_{it} + \beta^E \eta_{it} + \epsilon_{it} \tag{12}$$

Due to the Frisch-Waugh-Lowell theorem, the coefficients on the effort variables will be the same in (10) and (12). The coefficients on the circumstance variables will be different as they in (12) also incorporate the indirect effect of circumstances on effort. We will later report specifications where we omit this auxiliary regression.

To compare who the opportunity-deprived are across the four well-being measures, we rank individuals according to their opportunity profile, that is, according to the quality of their circumstances. We calculate each person's yearly unfairness

[8] Jusot *et al.* (2013) likewise call this Roemer's view, while not correcting for this correlation is termed Barry's view (Barry, 2005). A final possibility, where the correlation between effort and circumstances is considered effort, is called Swift's view (Swift, 2005).

rank as

$$r_{it}^{unfair} = G_t\big[(\beta^C + \gamma\beta^E)a_{it}^C\big], \tag{13}$$

where G_t is the yearly cumulative distribution of individuals' unfair advantage. This allows us to compare the opportunity-deprived across the four well-being measures.

To estimate equality of opportunity over time using the norm-based approach, we also compute 'fair' well-being ranks. These are given by

$$r_{it}^{fair} = H_t\big[\beta^E\eta_{it} + \epsilon_{it}\big], \tag{14}$$

where H_t is yearly the cumulative distribution of individuals' fair advantage. Using a divergence measure, we are now able to compute the level of inequality of opportunity for each of the welfare variables for each year of the survey.

4 Results

4.1 Who are the Opportunity-Deprived?

Table A.1 in Appendix A.2 shows the results of the regressions from equation (12). Based on this output and equation (13), we calculate each person's unfairness rank. Table 2 shows the correlations between these unfairness ranks for the four welfare measures. The correlations reveal the extent to which the same people are opportunity-deprived across the four measures. In the table - and throughout the paper - we bootstrap confidence intervals in order to take all derived uncertainty into account, including the uncertainty when constructing the welfare measures. We bootstrap 500 resamples at individual-level clusters.

[Table 2 around here]

The unfairness ranks display rather high correlations, suggesting that the same people are opportunity-deprived regardless of how we measure welfare. The welfare measures we have constructed are of course partly contained within each other; the income variable is included in both the multidimensional index and the equivalent income measure, and the latter two use the same four dimensions but aggregate them differently. We can analyze the extent to which this is driving the high unfairness correlations by comparing the unfairness correlations with the correlation between the welfare ranks, $\boldsymbol{r^y}$, which are shown in Table 3. In all cases, the correlations are higher when we look at $\boldsymbol{r^{unfair}}$. This indicates that the high unfairness correlations are not driven just by the measures' interrelatedness. This also suggests that the way welfare is measured matters more if we target the welfare-deprived than if we target the opportunity-deprived.

[Table 3 around here]

Since individuals' opportunities are unobservable, policy makers may have to assist the opportunity-deprived indirectly. One way of doing so is by targeting individuals with circumstance profiles that are highly correlated with having low opportunities. We can use the unfairness ranks to test if the characteristics of the most opportunity-deprived are similar across the four welfare measures. To do so, we calculate the average unfairness rank for individuals with a given circumstance. Results are shown in Figure 2. The lower the average unfairness rank, the less opportunities individuals with the given circumstance have, and the more this circumstance is a potential factor policy makers can use to target the opportunity-deprived. If the average rank is less than 50, the particular group has less than average opportunities.

[Figure 2 around here]

There are many similarities across the welfare measures. Individuals with low educated parents and individuals whose father was a blue-collar worker or not employed have low opportunities. The same applies to individuals who grew up in the countryside, individuals born in East Germany, short individuals, females, and individuals with many siblings.

Important differences emerge only in two places, for people born abroad and for different age groups. People born abroad are more opportunity-deprived in all measures but life satisfaction. A possible explanation for this is that people born abroad perceive the life satisfaction scale differently than Germans, in which case this difference has little to do with differences in opportunity sets. Indeed, the regression output in Table A.1 in Appendix A.2 reveal that the only time where being born abroad, ceteris paribus, is not associated with lower welfare is for life satisfaction. If we exclude this circumstance variable from the regression, calculate new unfairness ranks, and again compute the average unfairness rank of individuals born abroad, then it falls below 50. Hence, we cannot exclude that this effect solely is due to scaling effects.

With respect to age, young people are opportunity-deprived in income but not in the multidimensional index. This is hardly surprising as the multidimensional index includes health, and young people on average are more healthy. It is questionable whether resources should be allocated such that individuals have equal opportunities in every part of their life. For this reason, we will later on place age on the other side of the responsibility cut. This may seem counterintuitive but it amounts to saying that individuals should have equal opportunities on expectation over their lifecycle rather than in every point of their life (see Almås *et al.* (2011) for a similar approach).

In sum, there seems to be relatively large agreement about whom the opportunity-deprived are across the four measures. Hence, if a policy maker strives to target the individuals with low opportunities, it matters relatively little how welfare is measured.

4.2 Equality of Opportunity over Time

To compare the development in equality of opportunity over time across the four welfare measures, we calculate the yearly correlations between $\boldsymbol{r_t^y}$ and, respectively, $\boldsymbol{r_t^{unfair}}$ and $\boldsymbol{r_t^{fair}}$. Results are displayed in Figure 3

[Figure 3 around here]

Before commenting on the trends over time, one remark is necessary. The *level* of the correlations cannot be compared across the different welfare measures. The reason for this is that some of the welfare ranks, particularly the life satisfaction ranks, by construction are noisier. This means that the residuals of the regression take up most of the variation. In our baseline set-up the residual is considered an effort variable that gets incorporated into the fairness ranks. The fairness ranks will therefore be very highly correlated with the actual ranks for a measure such as life satisfaction (and vice versa for the unfairness ranks). It cannot be concluded that there is more equality of opportunity with life satisfaction than, say, income, as the difference in levels may be entirely due to the degree of random noise inherent in the measures. This explains why the correlations in panel (b) of Figure 3 are much greater than in panel (a). To circumvent this problem, we will compare the trends over time.

With regards to the trends over time, all measures undergo a sharp increase in inequality of opportunity in the early '90s. This is caused by the introduction of East

Germany into the sample. In panel (a), there is a gradual improvement in inequality of opportunity over the past 15 years. In other words, the quality of individuals' circumstances is becoming less correlated with their welfare levels - regardless of how welfare is measured. At the same time panel (b) shows that for both the income and the equivalent income measure, the effort ranks are becoming more associated with the welfare ranks, suggesting that effort is being more rewarded. Only for life satisfaction and the multidimensional index do we not see improvements in this aspect over the past 15 years.

The figures suggest that regardless of what concept of welfare we use to measure inequality of opportunity, the trends over broadly similar; inequality of opportunity increased with the introduction of East Germany, and overall inequality of opportunity - understood as some combination of the trends of panel (a) and panel (b) - has improved over the past 15 years.

4.3 Altering the Responsibility Cut

The analysis thus far was based on important normative assumptions regarding what individuals were held responsible for. We assumed that individuals were responsible for four variables (*4var*); their education, work hours, and whether they are self-employed or work in the public sector. We also assumed that individuals should not be held responsible for the part of these variables that could be accounted for by circumstance variables. That is, the correlation (*cor*) between circumstance and efforts was itself considered outside the control of individuals. We further assumed that the part of individual well-being that was unaccounted for by circumstance or effort variables (*residual*) was within individual control. Next, we implicitly considered well-being differences across different age groups (*age*) as unfair. Finally, for the equivalent income measure, we assumed that individuals were not responsible for well-being differences due to preference heterogeneity arising from circumstances (*pref*).

Our baseline effort set contained *4var* and *residual*. In this section, we try to shift the responsibility cut by alterting what goes into the effort set. First, we look at whether the characteristics of the opportunity-deprived change, as we change the effort set. This is analyzed in Figures 4 and 5. Figure 4 uses our most narrow effort set, only *4var*, while Figure 5 uses the widest possible effort set, *4var, residual, cor, age, pref*.

[Figures 4 and 5 around here]

The figures show the same overall pattern as our main results; for the most part the opportunity-deprived share the same characteristics across all four measures. The only disputes are, once again, for individuals born abroad, and for different age groups. The responsibility cut does matter, however, for quantifying the degree to which a particular group is opportunity-deprived. For example, individuals born in East Germany have an average opportunity rank of about 35 with the smallest effort set and 15 with the largest effort set.

Next, we look at whether the developments over time depend on where we place the responsibility cut. We try six different specifications, which gradually expand the effort set. Results are displayed in Figure 6. Keep in mind that our primary interest is not whether the trends change as we change the responsibility cut, but rather whether the four well-being measures follow similar trends regardless of where we place the cut.

[Figure 6 around here]

Our baseline results are displayed in panel (c). If we add more to the responsibility side (panel (d) to (f)), then the pattern remains the same; inequality of opportunity increased with the inclusion of East Germany but since then gradually fell. This applies to all four welfare variables. Hence, broadening our effort set does not influence our baseline results. The picture looks similar but less pronounced if we only consider the residual in the effort set (panel (b)). There are slightly diverging patterns in the '90s, but over the last fifteen years all four well-being display improvements in equality of opportunity. If we only hold individuals responsible for the four choice variables and not the residual (panel (a)), then the picture looks very different, particularly for life satisfaction. As already argued, the life satisfaction ranks carry a lot of noise, implying that the four choice variables are only able to explain a very little part of the variance. When we rank individuals according to their unfairness and let the residual be part of the unfairness ranks, these will be almost perfectly aligned with the welfare ranks. Other studies that have tried to switch the residual to the other side of the responsibility cut likewise found this to have a great impact on the results (see for example Almås *et al.* (2011)and Devooght (2008)). Still, the trends over time are broadly the same across the other three well-being measures.

In Figure A.2 in Appendix A.3, we show the impact of changing the responsibility cut for the correlations between $\boldsymbol{r^y}$ and $\boldsymbol{r^{fair}}$. As long as we do not change the residual to the other side, the figures display a modest decrease in inequality of opportunity for incomes and equivalent incomes since 2000 and a flat trend for life satisfaction and the multidimensional index. The picture looks different, but the trends are somewhat similar, if we switch the residual to the other side of the responsibility cut. This is little different from our baseline results.

In sum, we find that when characterizing whom the opportunity-deprived are, neither the measure of welfare nor the precise location of the responsibility cut is of great importance. When analyzing developments in equality of opportunity over time, where we place the residual matters quite a bit, while enlarging the effort set further has few implications. For all well-being variables, we find that over the past fifteen years, the correlation between welfare ranks and unfairness ranks have decreased, while the correlation between welfare ranks and fairness ranks in no case has become worse.

4.4 Robustness Checks

The main analysis was based on a specific choice of a divergence measure. In this section, we test whether our findings are sensitive to using other divergence measures. We use the Magdalou-Nock divergence measures with $s = 0$, $s = 1$, and $s = 2$, as well as the fairness Gini. We display the results both when calculating the divergence between $\boldsymbol{r^y}$ and $\boldsymbol{r^{fair}}$ (Figure A.3 in Appendix A.3) and when calculating the divergence between $\boldsymbol{r^y}$ and $\boldsymbol{r^{unfair}}$ (Figure A.4 in Appendix A.3). Recall that for these measures a large divergence between $\boldsymbol{r^y}$ and $\boldsymbol{r^{unfair}}$ suggests *greater* equality of opportunity. The reverse applies to the divergence between $\boldsymbol{r^y}$ and $\boldsymbol{r^{fair}}$.

The results are qualitatively unchanged when we use the Magdalou-Nock divergence measures with $s = 1$, $s = 2$, or the fairness Gini. In all three cases, the unfairness ranks and the welfare ranks are growing more apart in the past fifteen years. In contrast, the fairness ranks and the welfare ranks are growing more alike for equivalent incomes and incomes while no trend is visible for life satisfaction and the multidimensional index. This mimics exactly our baseline results, suggesting that they are not driven by our particular choice of divergence measure. Only the Magdalou-Nock divergence measure with $s = 0$ displays different developments. In particular, the developments are quite erratic with large confidence bans making it hard to extract any meaningful changes. This divergence measure puts great emphasis on divergences at the bottom of the distribution and is particularly sensitive to values close to zero. If we add 1 to all ranks (such that the ranks go from 1 to 101), then the Magdalou-Nock index with $s = 0$ shows the same trends as the other divergence measures.

5 Conclusion

We have investigated if equality of opportunity estimates depend on what, precisely, it is that we seek to equalize opportunities for. Based on philosophical literature on well-being, we constructed four measures of welfare that are candidates for what we ought to equalize opportunities for. Upon constructing these, we analyzed if the way welfare is measured matters for 1) characterizing the opportunity-deprived and 2) tracking inequality of opportunity over time. We found that for the most part, neither depend greatly on what measure of well-being we use. These results are robust to most alternative measurement assumptions and changes to the responsibility cut. This is encouraging news for researchers and policy makers interested in going beyond GDP. Whereas previous research has shown that going beyond GDP matters greatly for defining the welfare-deprived and for tracking growth in welfare over time, our findings suggest that going beyond income is less important if the object of interest is equality of opportunity. Circumstances beyond individual control influence welfare in a similar fashion regardless of how welfare is measured. Hence, for matters of distributive justice, alternative measures of GDP seem to have less importance, and a good picture may be achieved by simply using incomes as a proxy variable for welfare.

References

Adler, M. D. and Fleurbaey, M. (2016). *The Oxford Handbook of Well-Being and Public Policy*. Oxford University Press. (page 5)

Akay, A., Bargain, O. and Jara, H. X. (2015). Back to Bentham: Should we? Large-scale comparison of decision versus experienced utility for income-leisure preferences. *Working paper*. (page 12)

Almås, I., Cappelen, A. W., Lind, J. T., Sørensen, E. Ø. and Tungodden, B. (2011). Measuring unfair (in)equality. *Journal of Public Economics*, **95** (7), 488–499. (page 10, 17, 19)

Arneson, R. J. (1989). Equality and equal opportunity for welfare. *Philosophical studies*, **56** (1), 77–93. (page 2, 6)

Barry, B. (2005). *Why social justice matters*. Polity. (page 14)

Blanchflower, D. G. and Oswald, A. J. (2004). Well-being over time in Britain and the USA. *Journal of Public Economics*, **88** (7), 1359–1386. (page 4)

Bossert, W. (1995). Redistribution mechanisms based on individual characteristics. *Mathematical Social Sciences*, **1** (31), 51. (page 7)

Calsamiglia, C. (2009). Decentralizing equality of opportunity. *International Economic Review*, **50** (1), 273–290. (page 4)

Chetty, R., Hendren, N., Kline, P., Saez, E. and Turner, N. (2014). Is the United States still a land of opportunity? Recent trends in intergenerational mobility. *The American Economic Review*, **104** (5), 141–147. (page 9)

Cohen, G. A. (1990). Equality of what? On welfare, goods and capabilities. *Recherches Économiques de Louvain/Louvain Economic Review*, pp. 357–382. (page 2, 6)

Cowell, F. A. (1985). Measures of distributional change: An axiomatic approach. *The Review of Economic Studies*, **52** (1), 135–151. (page 9)

Crisp, R. (2016). Well-being. In E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, summer 2016 edn., Stanford University. (page 5)

Decancq, K., Fleurbaey, M. and Schokkaert, E. (2015). Happiness, equivalent incomes and respect for individual preferences. *Economica*, **82** (s1), 1082–1106. (page 12)

— and Neumann, D. (2016). Does the choice of well-being measure matter empirically? In M. D. Adler and M. Fleurbaey (eds.), *Oxford Handbook of Well-Being and Public Policy*, Oxford University Press, pp. 553–587. (page 4, 12)

Devooght, K. (2008). To each the same and to each his own: A proposal to measure responsibility-sensitive income inequality. *Economica*, **75**, 280–295. (page 10, 19)

Dworkin, R. (1981). What is equality? Part 2: Equality of resources. *Philosophy & Public Affairs*, pp. 283–345. (page 6)

Ferreira, F. H. and Peragine, V. (2016). Individual responsibility and equality of opportunity. In M. D. Adler and M. Fleurbaey (eds.), *Oxford Handbook of Well-Being and Public Policy*, Oxford University Press, pp. 746–784. (page 2)

Fleurbaey, M. (1994). On fair compensation. *Theory and decision*, **36** (3), 277–307. (page 7)

— (1995). The requisites of equal opportunity. *Social choice, welfare and ethics*, pp. 37–53. (page 7)

— and Maniquet, F. (2011). *A theory of fairness and social welfare*, vol. 48. Cambridge University Press. (page 4)

— and Schokkaert, E. (2009). Unfair inequalities in health and health care. *Journal of health economics*, **28** (1), 73–90. (page 8)

Griffin, J. (1986). *Well-being: Its meaning, measurement, and moral importance.* Clarendon Press. (page 2, 5, 11)

Jusot, F., Tubeuf, S. and Trannoy, A. (2013). Circumstances and efforts: How important is their correlation for the measurement of inequality of opportunity in health? *Health economics*, **22** (12), 1470–1495. (page 14)

Magdalou, B. and Nock, R. (2011). Income distributions and decomposable divergence measures. *Journal of Economic Theory*, **146** (6), 2440–2454. (page 9)

Parfit, D. (1984). *Reasons and persons.* Oxford University Press. (page 2, 5, 11)

Ramos, X. and Van de gaer, D. (2015). Measurement of inequality of opportunity based on counterfactuals. (page 7)

— and — (2016). Empirical approaches to inequality of opportunity: Principles, measures, and evidence. *Journal of Economic Surveys*, **30** (5), 855–883. (page 2, 3, 8)

Ravallion, M. (2015). Inequality when effort matters. (page 4)

Rawls, J. (1971). *A Theory of Justice*. Harvard University Press. (page 6)

Roemer, J. E. (1993). A pragmatic theory of responsibility for the egalitarian planner. *Philosophy & Public Affairs*, pp. 146–166. (page 7)

— (1998). *Equality of Opportunity*. Harvard University Press. (page 14)

— (2012). On several approaches to equality of opportunity. *Economics and Philosophy*, **28** (02), 165–200. (page 4)

— and Trannoy, A. (2015). Equality of opportunity. In A. B. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution, Volume 2A-2B*, North Holland, pp. 217–300. (page 2)

— and — (2016). Equality of opportunity: Theory and measurement. *Journal of Economic Literature*, **54**, 1288–1332. (page 6)

Sen, A. (1980). Equality of what? *The Tanner Lecture on Human Values*, pp. 197–220. (page 6)

— (1985). *Commodities and Capabilities*. North-Holland. (page 5)

Stevenson, B. and Wolfers, J. (2008). Happiness inequality in the United States. *Journal of Legal Studies*, **37**. (page 4)

Swift, A. (2005). *Justice, luck and the family*. Princeton University Press, Princeton, NJ.[361]. (page 14)

Van de gaer, D. (1993). *Equality of opportunity and investment in human capital*. Ph.D. thesis, Leuven: Katholieke Universiteit Leuven. (page 7)

Tables

Table 1: Summary Statistics

	mean	sd	min	max
Circumstance Variables				
Father's Education: Primary school	0.68	0.47	0	1
Father's Education: Secondary school	0.19	0.40	0	1
Father's Education: Tertiary school	0.13	0.33	0	1
Mother's Education: Primary school	0.71	0.45	0	1
Mother's Education: Secondary school	0.22	0.42	0	1
Mother's Education: Tertiary school	0.07	0.25	0	1
Father's Occupation: Blue-collar (untrained)	0.14	0.35	0	1
Father's Occupation: Blue-collar (trained)	0.34	0.47	0	1
Father's Occupation: Not employed	0.06	0.24	0	1
Father's Occupation: White-collar	0.26	0.44	0	1
Father's Occupation: Self-employed	0.12	0.32	0	1
Father's Occupation: Civil servant	0.08	0.28	0	1
Place of Upbringing: Large city	0.22	0.41	0	1
Place of Upbringing: Medium city	0.18	0.38	0	1
Place of Upbringing: Small city	0.23	0.42	0	1
Place of Upbringing: Countryside	0.37	0.48	0	1
Place of Birth: West Germany	0.66	0.47	0	1
Place of Birth: East Germany	0.27	0.44	0	1
Place of Birth: Abroad	0.07	0.26	0	1
Height (cm)	173	9.13	80	210
Female	0.47	0.50	0	1
Number of siblings	1.94	1.68	0	17
Age	42.0	12.1	17	91
Effort Variables				
Years of education	12.6	2.72	7	18
Weekly working time	35.9	15.9	0	80
Self-employed	0.09	0.29	0	1
Works in public sector	0.18	0.39	0	1

Notes: Summary statistics of circumstance and effort variables. $n = 176,196$.

Table 2: Correlation Between Unfairness Ranks

	Income	Life Sat.	Multidim. Index	Equivalent Inc.
Income	-	-	-	-
Life Satisfaction	0.60	-	-	-
	(0.55, 0.66)			
Multidim. Index	0.64	0.90	-	-
	(0.61, 0.68)	(0.87, 0.93)		
Equivalent Inc.	0.88	0.72	0.82	-
	(0.79, 0.92)	(0.66, 0.79)	(0.76, 0.88)	

Notes: Correlations between $\boldsymbol{r^{unfair}}$ for the four welfare measures. Bootstrapped 95^{th} percentile confidence intervals in parenthesis.

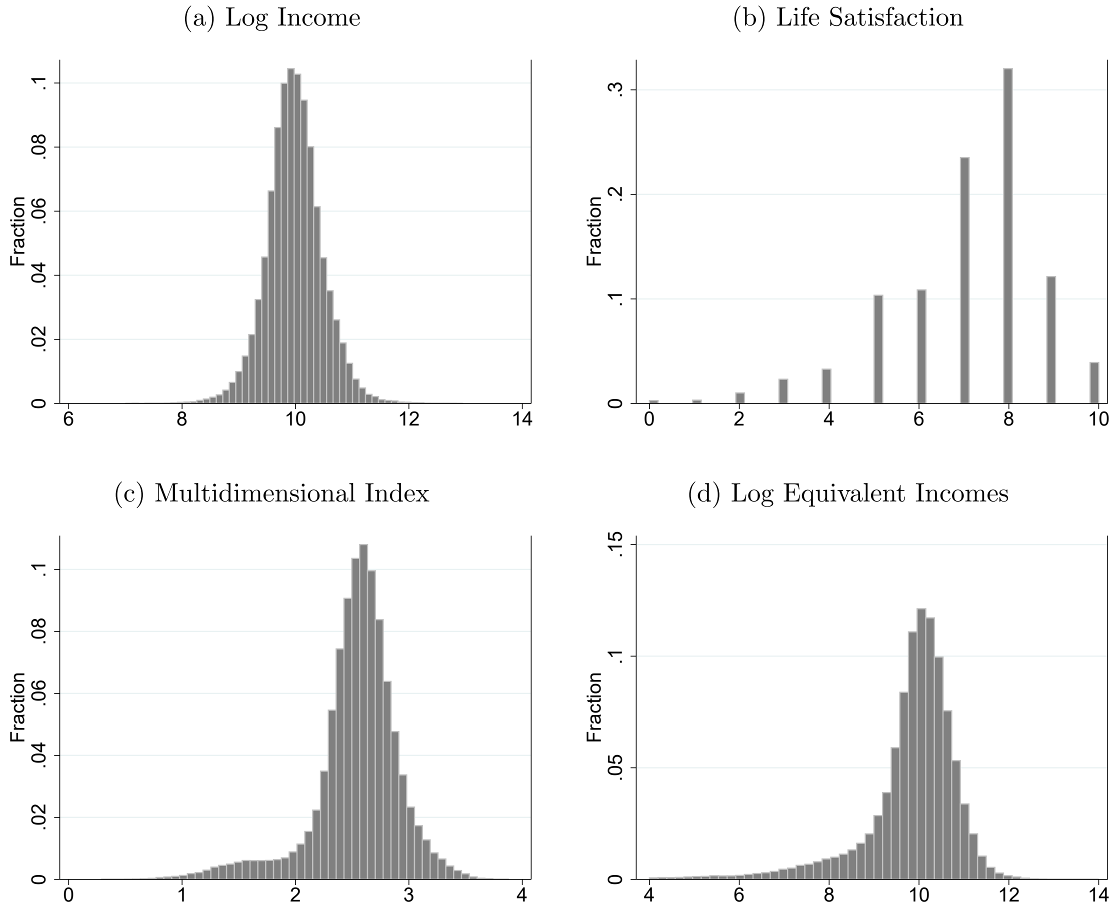
Table 3: Correlation Between Welfare Ranks

	Income	Life Sat.	Multidim. Index	Equivalent Inc.
Income	-	-	-	-
Life Satisfaction	0.21	-	-	-
	(0.20, 0.22)			
Multidim. Index	0.41	0.23	-	-
	(0.40, 0.43)	(0.22, 0.24)		
Equivalent Inc.	0.66	0.27	0.77	-
	(0.61, 0.71)	(0.26, 0.28)	(0.74, 0.80)	

Notes: Correlation between welfare ranks, $\boldsymbol{r^{y}}$. Bootstrapped 95^{th} percentile confidence intervals in parenthesis.

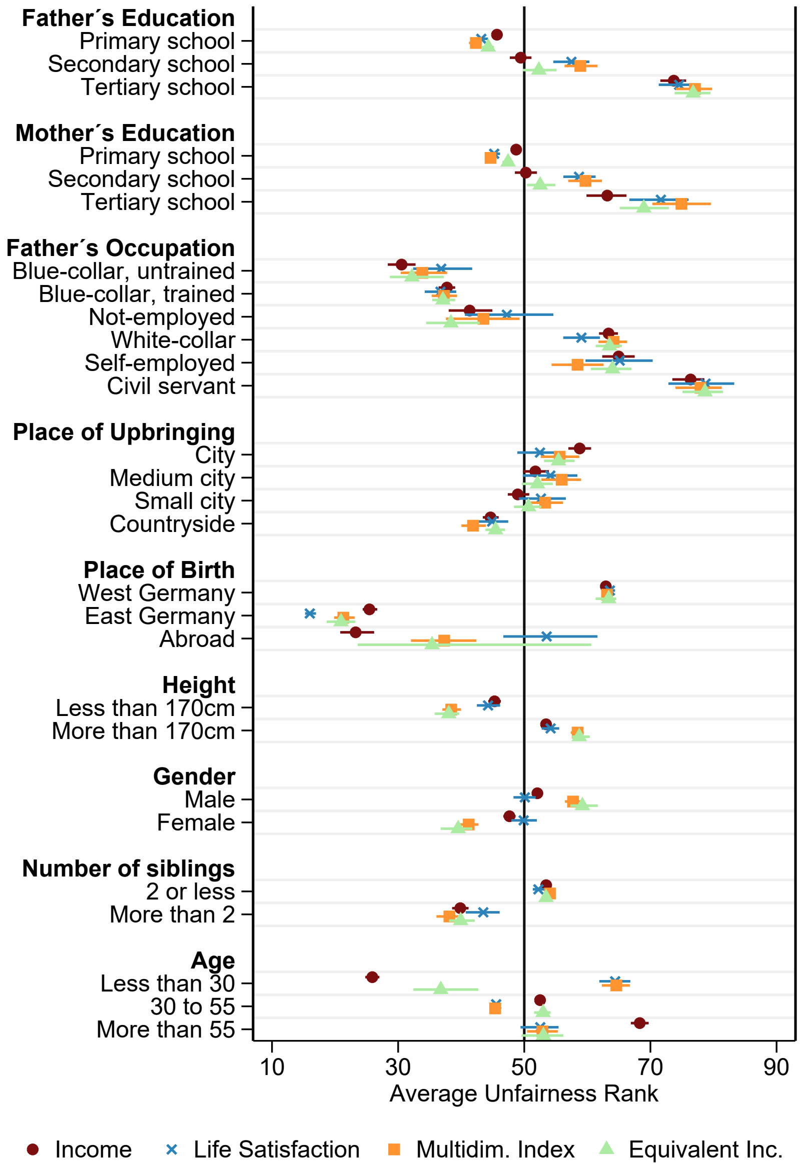
Figures

Figure 1: Histograms



Notes: Histograms of the four welfare measures.

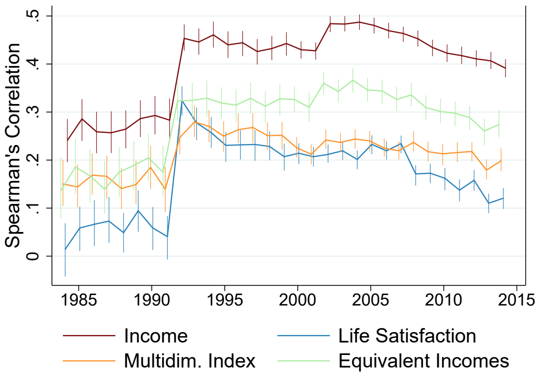
Figure 2: Who are the Opportunity-deprived?



Notes: The figure shows the average unfairness rank for individuals that share a given circumstance. If the points are to the left of the line at 50, then individuals with this circumstance are more than average opportunity-deprived and vice versa. Bars indicate bootstrapped 95^{th} percentile confidence bans.

Figure 3: Equality of Opportunity over Time

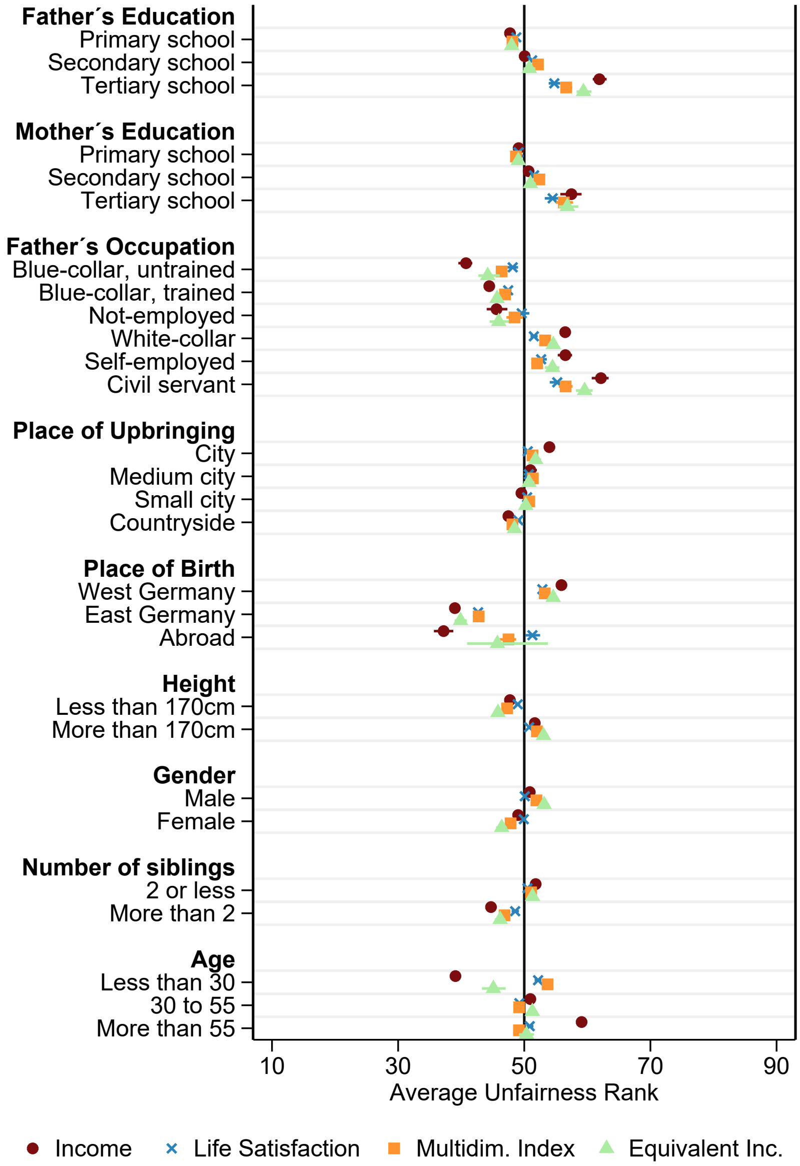
(a) Correlation between r^y & r^{unfair}



(b) Correlation between r^y & r^{fair}

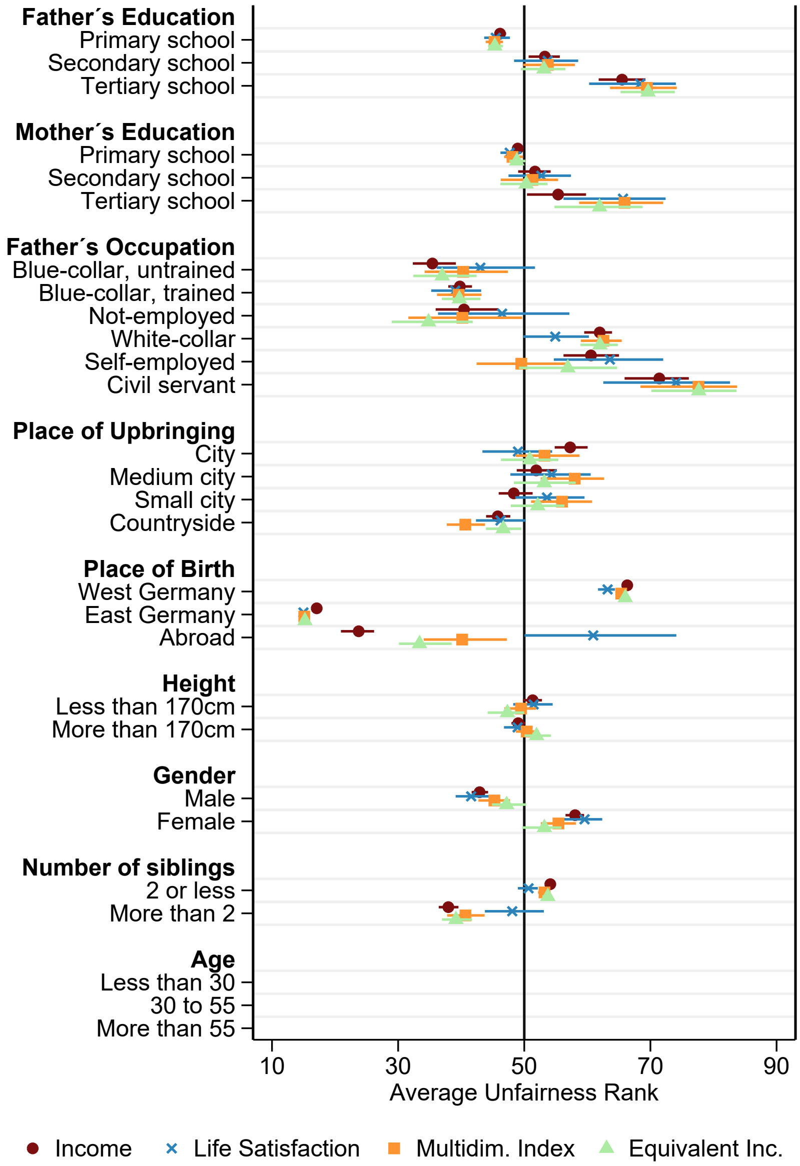
Notes: Development in inequality of opportunity in each well-being variable from 1984-2014. In panel (a) large correlations suggest a high presence of inequality of opportunity, while in panel (b) high correlations suggest a low presence of inequality of opportunity. Bars indicate bootstrapped 95^{th} percentile confidence bans.

Figure 4: Changing the Responsibility Cut: Effort = {4var}



Notes: The figure shows the average unfairness rank for people sharing a particular circumstance using a narrow effort set. Bars indicate bootstrapped 95^{th} percentile confidence bans.

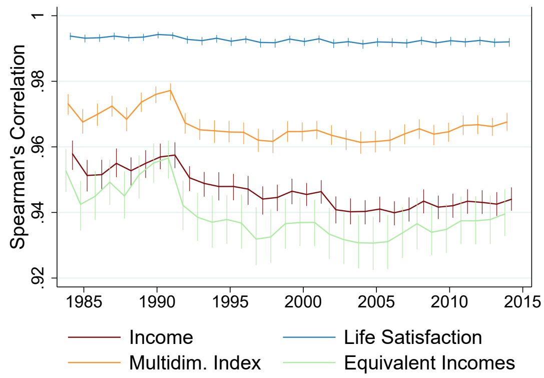
Figure 5: Changing the Responsibility Cut: Effort = {4var, residual, cor, age, pref}



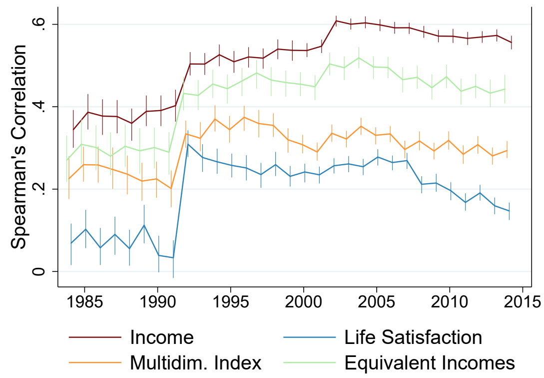
Notes: The figure shows the average unfairness rank for people sharing a particular circumstance using a wide effort set. We no longer report differences in the average unfairness rank by age groups, as the objective now is to equalize lifetime opportunities. Bars indicate bootstrapped 95^{th} percentile confidence bans.

Figure 6: Altering the Responsibility Cut

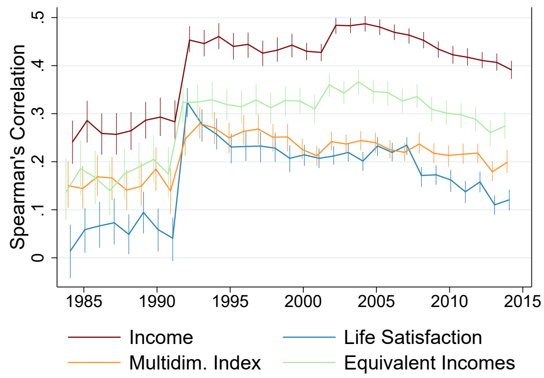
(a) Effort: 4var



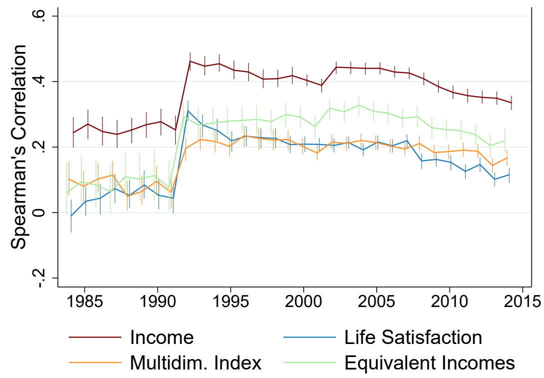
(b) Effort: Residual



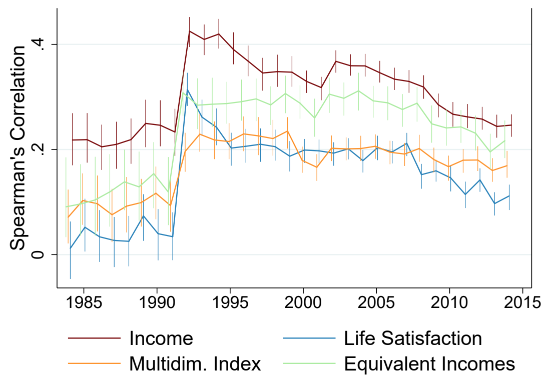
(c) Effort: 4var, residual



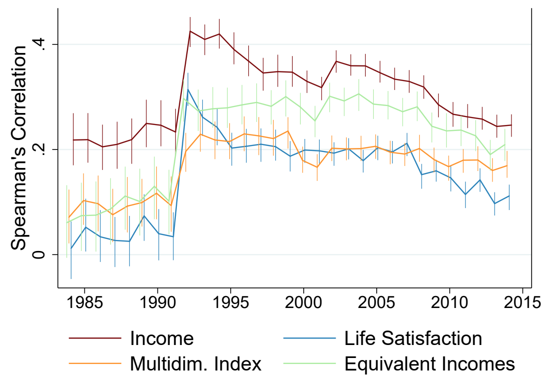
(d) Effort: 4var, residual, cor



(e) Effort: 4var, residual, cor, age



(f) Effort: 4var, residual, cor, age, pref



Notes: Development in inequality of opportunity in each well-being variable from 1984-2014 for different responsibility cuts. *4var:* The four variables work hours, education, self-employed, and works in public sector are considered effort. *Residual:* The residuals from the regressions of the well-being variables on circumstance and effort variables are considered effort. *Age:* Age is considered effort (implying we are equalizing lifetime opportunities). *Cor:* The correlation between effort and circumstance variables is not considered a circumstance. *Pref:* Individuals are fully held responsible for their preferences (only applies to the equivalent income measure). Our baseline specification used *4var* and *residual* as effort.

A Appendix

A.1 Decomposition of Equivalent Incomes

The equivalent incomes are calculated from the following equation:

$$inc_{it}^{eq} = exp\Big(ln(inc_{it}) + \frac{\beta^{dim} + \gamma^{dim} w_{it}}{\beta^{inc} + \gamma^{inc} w_{it}}(dim_{it} - \tilde{dim})\Big) \tag{15}$$

We want to decompose this into a part that reflects the welfare individuals derive from their specific preferences, and a part that reflects the welfare individuals receive independent of their particular preferences. To do so we define a set of norm preferences by fixing the preference heterogeneity variables at a given level, $\tilde{w}$, and computing individuals' equivalent incomes with these norm preferences:

$$\tilde{inc}_{it}^{eq} = exp\Big(ln(inc_{it}) + \frac{\beta^{dim} + \gamma^{dim} \tilde{w}}{\beta^{inc} + \gamma^{inc} \tilde{w}}(dim_{it} - \tilde{dim})\Big) \tag{16}$$

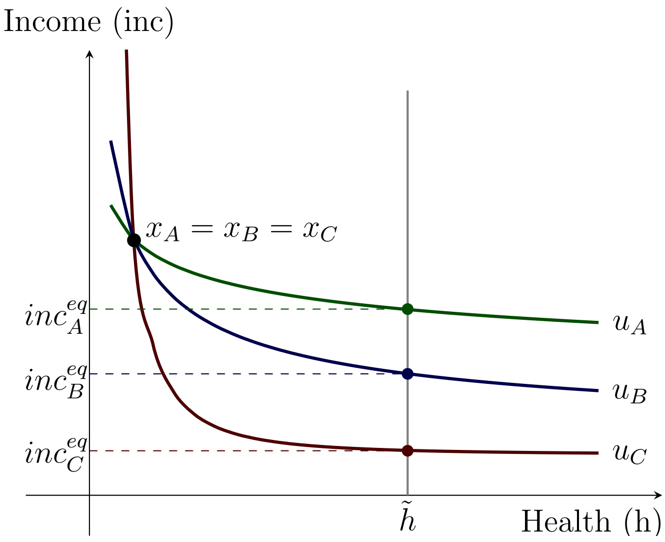
We choose the mean (mode for categorical variables) of the preference heterogeneity variables as the norm preferences. $\tilde{inc}_{it}^{eq}$ does not depend on an individual's preferences, and has thus taken the preference heterogeneity out of the original equivalent incomes. We can now express an individual's equivalent income as:

$$inc_{it}^{eq} = \tilde{inc}_{it}^{eq} + \phi_{it}, \tag{17}$$

where ϕ_{it} is the contribution to equivalent incomes individuals receive from the match between their particular bundle and their preferences. An illustrative example is given in Figure A.1. We consider three individuals, A, B, and C, and two dimensions of well-being, income and health. A, B, and C all consume the same bundle, which contains a lot of income but only little health. They have different preferences, though, with A putting the largest preference on income relative to health compared to B and C. It seems fitting that A should derive the most welfare from this income-heavy bundle, which also is the case when we calculate the three individuals' equivalent incomes.

Suppose that A, B, and C have different preferences because of variation in a circumstance variable, for example their parents' level of education. If we use the equivalent income ranks as the welfare variable that is regressed on effort and circumstance variables, which include parental education, then the coefficient on parents' education will eat up the differences in equivalent incomes. When we construct the unfairness ranks, A will have a higher unfairness rank than B and C, because

Figure A.1: Equivalent Income Illustration



A's circumstances - through his preferences - are yielding a higher well-being rank. This means that well-being differences due to preference heterogeneity from circumstances are not considered fair, and individuals are not held fully responsible for their preferences.

Our decomposition can circumvent this problem. In this example, we choose individual B's preferences as the norm preferences. Our decomposition asks what level of equivalent incomes each individual would have had with B's preferences. In this case, all three individuals would have $\tilde{inc}_i^{eq} = inc_B^{eq}$ for $i = A, B, C$. This means that $\phi_A = inc_A^{eq} - inc_B^{eq} > 0$ and $\phi_C = inc_C^{eq} - inc_B^{eq} < 0$, implying that A gets a positive boost from the match between his bundle of goods and his preferences, while the reverse applies to C.

In essence, we want to hold individuals responsible for their ϕ-term. Suppose we use ranks of $\tilde{inc}_{it}^{eq}$ instead of inc_{it}^{eq} to regress on circumstance and effort variables. Since the ϕ-terms are not part of these ranks, the circumstance variables can no longer pick up well-being differences due to preference heterogeneity. Hence, when we calculate the unfair ranks, preference heterogeneity does not enter, and individuals are held responsible for their preferences. We still use the baseline equivalent incomes, inc_{it}^{eq}, to calculate the welfare ranks we use for the divergence measures, but the unfairness ranks will be based on a regression with $\tilde{inc}^{eq}$.

Unfortunately, this method does not allow us to calculate fairness ranks that hold individuals responsible for their preferences. If we follow the same approach as above, and hence regress the ranks of $\tilde{inc}_{it}^{eq}$ on circumstance and effort variables, we face the problem that the fairness ranks should be composed of both the contribution from the effort variables and the ϕ_{it}-terms. As these two terms are on different scales, it is not entirely clear how to add them and turn them into a ranked-based measure. Hence, we will only use the decomposition of equivalent incomes for divergences between $\boldsymbol{r^{y}}$ and $\boldsymbol{r^{unfair}}$.

A.2 Regression Output

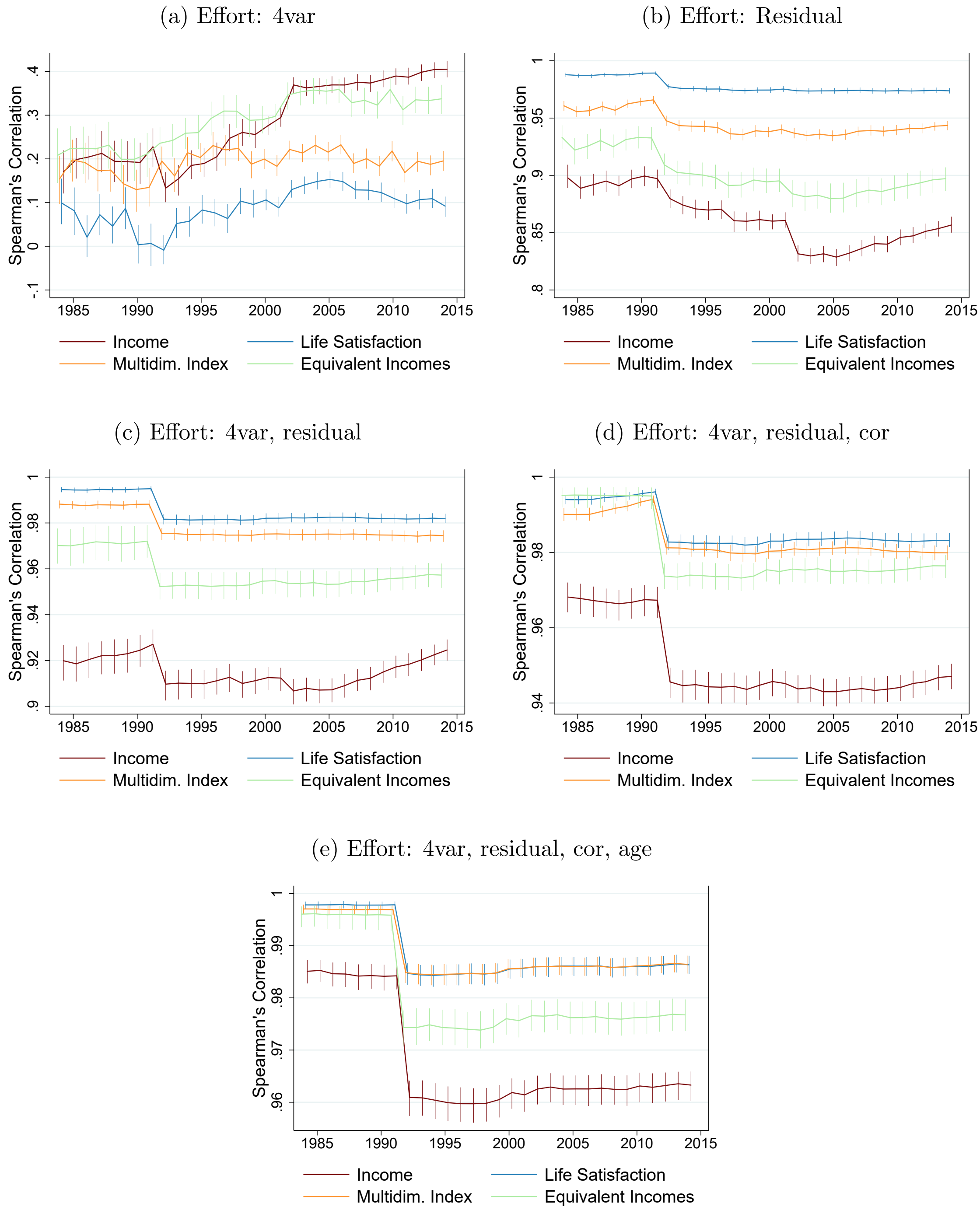
Table A.1: Regressing Welfare Ranks on Circumstances and Effort

	Income		Life Satisfaction		Multidim. Index		Equivalent Income	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Circumstance Variables								
Father's education (ref: primary school)								
Secondary school	4.37***	(0.48)	1.79***	(0.51)	2.56***	(0.48)	3.44***	(0.46)
Tertiary school	7.81***	(0.64)	2.98***	(0.66)	4.14***	(0.63)	6.76***	(0.62)
Mother's education (ref: primary school)								
Secondary school	2.38***	(0.46)	1.10**	(0.49)	1.13**	(0.46)	1.15**	(0.45)
Tertiary school	2.86***	(0.76)	2.11***	(0.76)	2.56***	(0.77)	2.33***	(0.73)
Father's occup. (ref: blue collar, untrained)								
Blue-collar, trained	3.08***	(0.53)	0.73	(0.57)	1.20**	(0.53)	2.30***	(0.49)
Not employed	2.44***	(0.76)	1.15	(0.81)	0.68	(0.76)	1.02	(0.70)
White-collar	8.98***	(0.59)	2.50***	(0.63)	4.17***	(0.58)	6.71***	(0.55)
Self-employed	7.81***	(0.68)	2.65***	(0.70)	2.58***	(0.66)	5.73***	(0.65)
Civil servant	9.13***	(0.76)	3.43***	(0.81)	4.51***	(0.74)	7.19***	(0.73)
Place of upbringing (ref: large city)								
Medium city	-0.90*	(0.52)	0.79	(0.54)	0.90*	(0.51)	0.35	(0.50)
Small city	-1.11**	(0.49)	0.93*	(0.50)	0.87*	(0.48)	0.64	(0.47)
Countryside	-1.16***	(0.45)	0.32	(0.47)	-0.50	(0.44)	0.55	(0.43)
Place of Birth (ref: West Germany)								
East Germany	-15.53***	(0.38)	-10.08***	(0.39)	-10.04***	(0.37)	-13.45***	(0.33)
Abroad	-12.56***	(0.65)	-0.60	(0.69)	-4.10***	(0.64)	-5.07***	(0.61)
Height (cm)	0.29***	(0.02)	0.18***	(0.03)	0.16***	(0.02)	0.24***	(0.02)
Female	2.75***	(0.45)	2.49***	(0.47)	-1.33***	(0.45)	-2.67***	(0.43)
Number of siblings	-1.71***	(0.10)	-0.26**	(0.11)	-0.75***	(0.10)	-1.18***	(0.09)
Age	1.31***	(0.07)	-0.64***	(0.07)	-0.85***	(0.07)	1.07***	(0.06)
Age squared	-0.01***	(0.00)	0.01***	(0.00)	0.01***	(0.00)	-0.01***	(0.00)
Effort Variables								
Years of education	2.77***	(0.07)	0.79***	(0.07)	1.07***	(0.07)	2.01***	(0.07)
Weekly working time	0.42***	(0.01)	0.18***	(0.01)	0.46***	(0.01)	0.57***	(0.01)
Self-employed	5.55***	(0.56)	-0.21	(0.53)	1.72***	(0.53)	5.38***	(0.51)
Works in public sector	2.01***	(0.32)	2.71***	(0.35)	1.88***	(0.34)	1.77***	(0.35)
r^2	0.30		0.05		0.12		0.23	

Notes: * p< 0.10, ** p< 0.05, *** p< 0.01. Based on a pooled OLS with individual level clustered standard errors. $n = 176,196$.

A.3 Robustness Checks

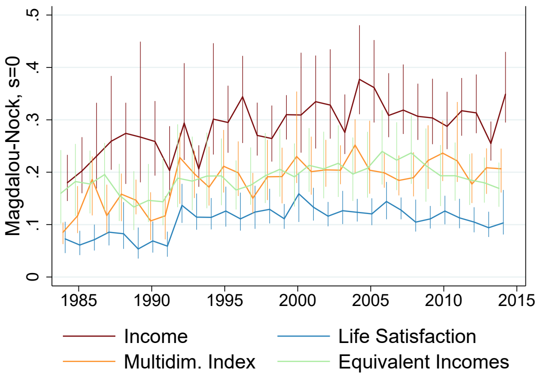
Figure A.2: Altering the Responsibility Cut, $D(\boldsymbol{r^y} \| \boldsymbol{r^{fair}})$



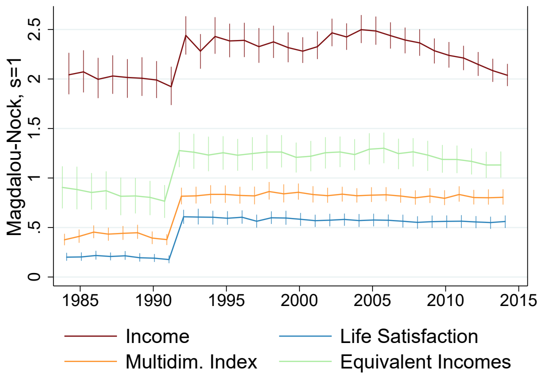
Notes: Development in inequality of opportunity from 1984-2014 for different responsibility cuts. *4var:* The four variables work hours, education, self-employed, and works in public sector are considered effort. *Residual:* The residuals from the regressions of the well-being variables on circumstance and effort variables are considered effort. *Age:* Age is considered effort (implying we are equalizing lifetime opportunities). *Cor:* The correlation between effort and circumstance variables is not considered a circumstance. Our method for decomposing the equivalent incomes cannot be used for generating r^{fair} ranks, hence we do not show a figure where individuals are held responsible for their preferences (see Appendix A.1 for details).

Figure A.3: Changing the Divergence Measure, $D(\boldsymbol{r^y} \| \boldsymbol{r^{fair}})$

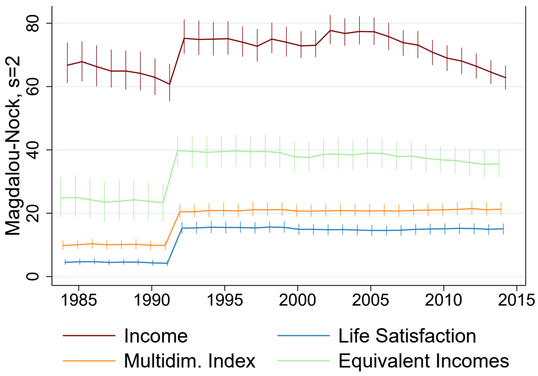
(a) Magdalou-Nock, $s = 0$



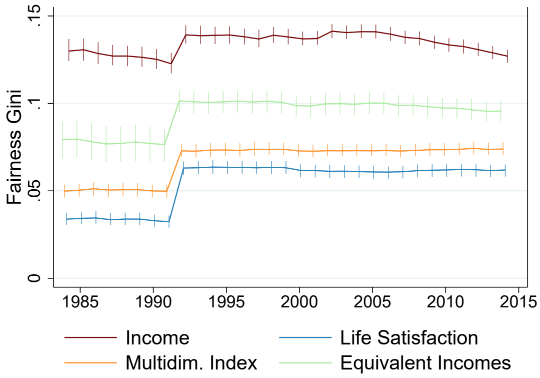
(b) Magdalou-Nock, $s = 1$



(c) Magdalou-Nock, $s = 2$

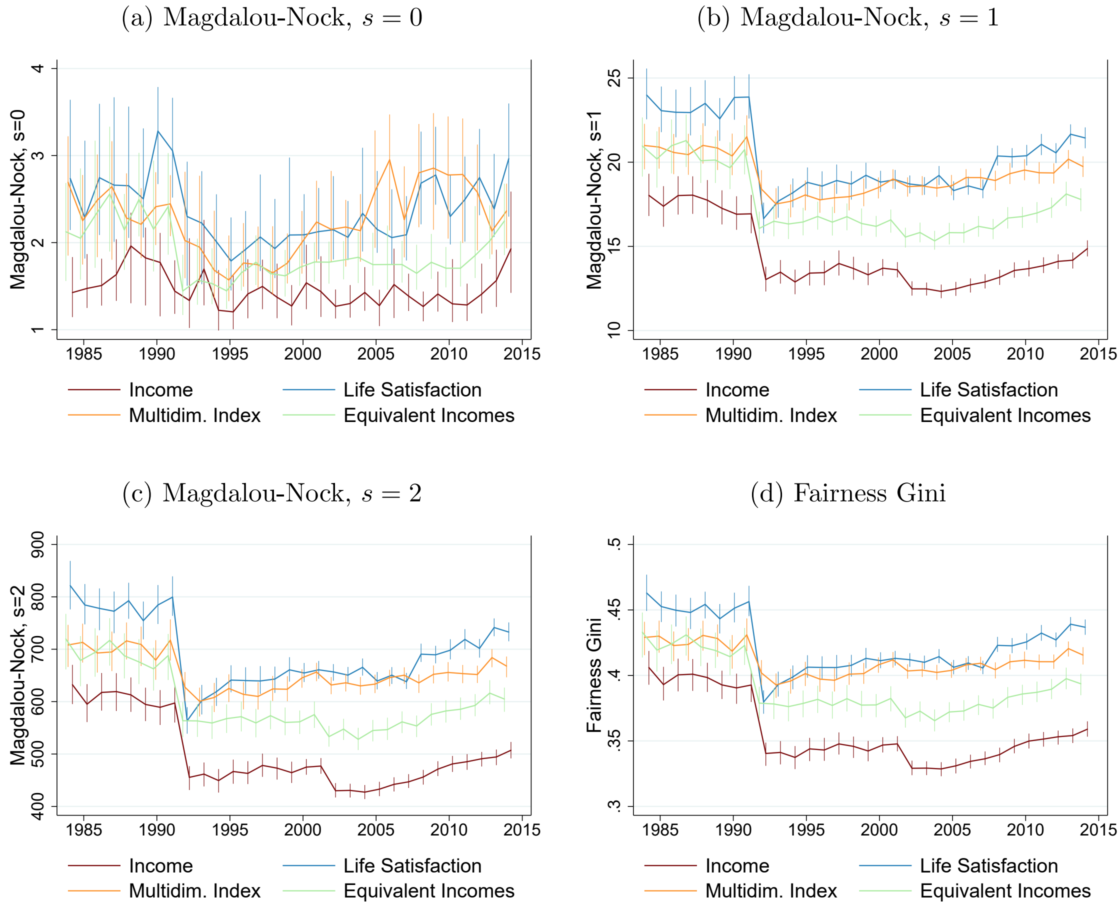


(d) Fairness Gini



Notes: Development in inequality of opportunity in each well-being variable from 1984-2014 using different divergence measures.

Figure A.4: Changing the Divergence Measure, $D(\boldsymbol{r^y} \| \boldsymbol{r^{unfair}})$



Notes: Development in inequality of opportunity in each well-being variable from 1984-2014 using different divergence measures.