

## EQUALITY OF OPPORTUNITY IN FOUR MEASURES OF WELL-BEING

BY DANIEL GERSZON MAHLER\*

*Development Research Group, The World Bank*

AND

XAVIER RAMOS

*Department of Applied Economics, Universitat Autònoma de Barcelona*

A growing literature has tried to measure the extent to which individuals have equal opportunities to acquire income. At the same time, policymakers have doubled down on efforts to go beyond income when designing policies to enhance 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 in which 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 10 years. This suggests that going beyond income has little relevance if the objective is to provide equal opportunities.

**JEL Codes:** D3, D63, I31

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

### 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 analyzed the extent to which individuals have equal opportunities for income acquisition (for recent reviews, see Roemer and Trannoy, 2015; Ramos and Van de gaer, 2016; Ferreira and Peragine, 2016). 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 for which individuals ought to be held personally responsible and income differences arising from

*Note:* We are very grateful to Brice Magdalou for comments and advice on the use of norm-based inequality metrics, to Dirk Van de gaer, Paul Hufe, Erik Schokkaert, and Niels Johannesen 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. Xavier Ramos acknowledges financial support of projects ECO2016-76506-C4-4-R (Spanish Ministry of Science, Innovation and Universities), and 2017SGR-1571 (Generalitat de Catalunya).

\*Correspondence to: Daniel Gerszon Mahler, The World Bank, 1818 H Street, NW Washington, DC 20433, USA (dmahler@worldbank.org).

factors outside the realm of personal responsibility. Whereas the former income differences are considered fair, the latter are considered unfair, and ought to be minimized.

In recent years, there has been a growing interest in going beyond income to measure individual well-being (Fitoussi *et al.*, 2010). Well-being (or welfare—we use the two interchangeably) is inherently multidimensional, and growth and income statistics fail to capture this multiplicity. Given the growing interest in going beyond income, it seems pertinent to apply this discussion to the equality-of-opportunity framework. If individuals ought to have equal opportunities for well-being, then the use of 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. Indeed, 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).

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 into 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 25 years of data from the German Socio-Economic Panel (SOEP) 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 measure of welfare matters for (1) characterizing the opportunity-deprived and (2) tracking inequality of opportunity over time.<sup>1</sup> In both cases, we first regress the welfare measures on circumstance and effort variables or, equivalently, variables for which we hold individuals responsible and variables for which we do not hold individuals responsible. In order to answer our first research question, whether the measure of welfare matters for characterizing the opportunity-deprived, we assign each individual an “opportunity” rank. The opportunity ranks order individuals according to how much welfare they derive from their circumstances. They can be interpreted as an ordinal measure of individuals’ opportunities, where the highest ranked possesses the best combination of circumstances and vice versa. We compare the determinants of the opportunity ranks across the four well-being measures. Broadly speaking, if individuals have

<sup>1</sup>Although Ramos and Van de gaer (2017) also test the sensitivity of inequality-of-opportunity estimates, their empirical analysis is very different. While we look at whether the measure of well-being matters for estimates of inequality of opportunity, Ramos and Van de gaer (2017) analyzes whether the method used to measure inequality of opportunity matters for estimates of inequality of opportunity.

similar opportunity ranks across the four 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, whether the measure of welfare matters for tracking inequality of opportunity over time, we use the norm-based approach. Upon regressing the welfare measures on circumstance and effort variables, this entails assigning a fair welfare level to each individual, which only depends on the individual's effort variables. Next, norm-based inequality metrics that calculate the divergence between the actual welfare levels and these fair welfare levels are computed. The more the two vectors diverge, the more circumstances rather than effort are determining welfare levels, and the more inequality of opportunity there is. For more information on the norm-based approach, see Ramos and Van de gaer (2016).

Since the four welfare measures follow different distributions, we are not able to compare the *extent* of inequality of opportunity across the different measures. This is little different from the fact that it is not possible to compare the level of welfare or inequality of welfare across the different measures. We will deal with this problem by indexing the extent of equality opportunity and tracking the development over time. The development over time is comparable across the different measures.

We find that the measure of well-being matters little, both with regard to characterizing the opportunity-deprived and with regard to tracking inequality of opportunity over time. In particular, we find that regardless of how welfare is measured, inequality of opportunity in Germany has decreased in the past 10 years. 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 policymakers 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) summarize 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 (2017), 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 in a variety of contexts. Average incomes in the United Kingdom and the United States (U.S.) have increased over recent decades while average happiness levels have stayed flat or decreased (Blanchflower and Oswald, 2004), income inequality in the U.S. has increased at the same time as inequality in happiness has decreased (Stevenson and Wolfers, 2008), and measures of welfare close to the ones we are adopting have been shown to matter for identifying the most welfare-deprived (Decancq and Neumann, 2016).

Roemer (2012) explicitly argues against using welfare as the outcome variable in equality-of-opportunity estimations. He does so on two grounds. First, policymakers 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 born in a certain region of a country 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 fewer opportunities to acquire a high level of well-being. Our framework helps to clarify if such examples have empirical leverage. The second reason why Roemer argues against using welfare as the outcome variable is grounded in the difficulty of measuring well-being in a cardinal way. Although this certainly complicates the exercise, we believe that useful lessons can still be drawn.

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 or her 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, and so on.

In short, mental state theory cares about what individuals feel, (revealed) preferences about what individuals choose, and objective list theory about external factors, which could be independent of the choices and 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 or 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. 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 and for not accounting for spillovers in well-being levels. All of these critiques can, of course, be counteracted, but doing so would be outside the scope of this paper.

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, 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 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 just outcome 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 are the right equalizandum, while Arneson (1989) and Cohen (1990) argued that we ought to pursue equality of opportunity for welfare and equal access to advantage (for a more complete account on the developments in distributive justice since Rawls, see Roemer and Trannoy, 2016).

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 for which we can hold individuals accountable. 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 have defined the equalizandum in terms of income. Rather, they have 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.<sup>2</sup>

### 2.3. Equality of Opportunity with the Norm-Based Approach

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,  $\mathcal{N} = \{1, 2, \dots, n\}$ , and a distribution of an outcome variable for this population,  $\mathbf{y} = (y_1, y_2, \dots, y_n)$ . Often,  $\mathbf{y}$  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,  $\mathbf{a}^C$ , and effort,  $\mathbf{a}^E$ . Circumstances are the factors outside the realm of control for the individual, the factors for which one ought not to hold an individual responsible. These are often taken to be gender, region of birth, parental education, parental income, and so on. Effort variables are the factors for which one ought to hold an individual responsible. The well-being of individual  $i$  is thus assumed to be given by  $y_i = f(\mathbf{a}_i^C, \mathbf{a}_i^E)$ . We consider the well-being levels to be cardinal and interpersonally comparable, but our setting also works in an ordinal framework, where we convert the welfare levels into welfare ranks.

Based on this setup, the literature proceeds by measuring the extent to which the outcome variable is driven by circumstance or effort variables. To do so, we will rely on the norm-based approach (Ramos and Van de gaer, 2016), also called the *fairness gap* in Fleurbaey and Schokkaert (2009). This approach uses fair allocation rules to assign each individual a fair outcome, which only depends on his or her effort level, and calculates inequality of opportunity as the divergence between the actual outcomes and the fair outcomes. The more the two distributions diverge, the less individuals' effort is associated with their well-being and the more circumstances are shaping individuals' outcomes.

The axiomatic literature on fair allocations has put forward two criteria that fair allocation rules ideally ought to satisfy, 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). The literature has proposed allocation rules that weaken these two principles in order to make them compatible. Two such rules are the Egalitarian Equivalent principle (Bossert and Fleurbaey, 1996) and the Generalized Proportionality principle (Cappelen and Tungodden, 2017).

Both principles start by defining a norm vector of circumstance variables,  $\tilde{\mathbf{a}}^C$ , and calculate the well-being that individuals would have had, if they had this norm level of circumstances,  $f(\mathbf{a}_i^E, \tilde{\mathbf{a}}^C)$ . This neutralizes the effect of circumstances on well-being and the differences in  $f(\mathbf{a}_i^E, \tilde{\mathbf{a}}^C)$  can therefore be thought of as expressing fair advantage. As the mean of  $f(\mathbf{a}_i^E, \tilde{\mathbf{a}}^C)$  may differ from the mean of the

<sup>2</sup>That being said, Arneson (1989) considered welfare to be preference satisfaction, thus differing from our take, in which we will look at different theories of welfare.



actual outcomes, and since resources should not be wasted and cannot come out of nothing, some adjustment is necessary. The two fair allocation rules deal with this in different ways. To estimate the distribution of fair outcomes, which we call  $\mathbf{z}$ , the Egalitarian Equivalent (EE) principle distributes the difference between the size of the two distributions equally across individuals:

$$(1) \quad z_i^{EE} = f(\mathbf{a}_i^E, \bar{\mathbf{a}}^C) + \left( \mu(\mathbf{y}) - \frac{\sum_{j \in \mathcal{N}} f(\mathbf{a}_j^E, \bar{\mathbf{a}}^C)}{n} \right),$$

where  $\mu(\mathbf{y}) = \frac{1}{n} \sum_{j \in \mathcal{N}} y_j$ . This principle has the advantage that if the well-being function is linearly separable in circumstances and effort, such that  $y_i = g(\mathbf{a}_i^E) + h(\mathbf{a}_i^C)$ , then the fair well-being levels do not depend on which norm vector of circumstances is chosen and the expression boils down to

$$(2) \quad z_i^{EE} = g(\mathbf{a}_i^E) + \frac{\sum_{j \in \mathcal{N}} h(\mathbf{a}_j^C)}{n}.$$

The Generalized Proportionality (GP) principle, instead, respects the well-being shares of the fair well-being allocation  $f(\mathbf{a}_i^E, \bar{\mathbf{a}}^C)$  by scaling them by the same factor:

$$(3) \quad z_i^{GP} = f(\mathbf{a}_i^E, \bar{\mathbf{a}}^C) \frac{\mu(\mathbf{y})}{\sum_{j \in \mathcal{N}} f(\mathbf{a}_j^E, \bar{\mathbf{a}}^C)/n}.$$

An individual’s fair outcome only depends on his or her effort variables and reflects how much the individual ideally is entitled to under the given allocation principles.<sup>3</sup> The more aligned the fair outcomes and the actual outcomes are, the lower is the inequality of opportunity. Conversely, the more they diverge, the greater is the inequality of opportunity. We can measure the extent to which the two distributions diverge from each other by employing a divergence measure,  $D(\mathbf{y} \parallel \mathbf{z})$ , which evaluates the divergence between the two distributions,  $\mathbf{y}$  and  $\mathbf{z}$ .

Magdalou and Nock (2011) (MN) provide a framework for axiomatically grounded divergence measures.<sup>4</sup> They put forth a class of divergence measures between an outcome distribution,  $\mathbf{y}$ , and a reference distribution,  $\mathbf{z}$ :

$$(4) \quad D_{MN}(\mathbf{y} \parallel \mathbf{z}) = \frac{1}{n} \sum_{i \in \mathcal{N}} [\phi(y_i) - \phi(z_i) - (y_i - z_i)\phi'(z_i)],$$

where, for all  $c \in \mathbb{R}_{++}$ , the function  $\phi(c)$  is given by the following:

$$(5) \quad \phi(c) := \begin{cases} \frac{1}{s(s-1)} c^s, & \text{if } s \neq 0, 1, \\ c \ln c, & \text{if } s = 1, \\ -\ln c, & \text{if } s = 0. \end{cases}$$

<sup>3</sup>Here, we neglect the minor impact that each individual’s circumstances have on the mean welfare level.

<sup>4</sup>We are heavily indebted for comments and advice from Brice Magdalou on the use and interpretation of appropriate divergence measures.

$s$  is a sensitivity parameter, similar to the one we find in standard inequality measurement. This class boils down to the generalized entropy class of standard inequality measures if  $\mathbf{z}$  is assumed to be the mean of the actual distribution. The class  $D_{MN}(\mathbf{y}||\mathbf{z})$  satisfies partial symmetry (i.e. it is invariant to permutations of  $(y_i, z_i)$  pairs) along with other relevant properties. By using the function  $\phi$  from equation (5) in equation (4), one obtains

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

Cowell (1985) suggests a different class of divergence measures, which he calls measures of distributional change, that satisfies different properties.<sup>5</sup> The measure equivalent to (4) is as follows:

$$(7) \quad D_C(\mathbf{y}||\mathbf{z}) = \frac{1}{n} \sum_{i \in \mathcal{N}} [z_i \phi(y_i/z_i)].$$

By using the function  $\phi$  from equation (5) in equation (7),  $D_C(\mathbf{y}||\mathbf{z})$  can be written as follows:

$$(8) \quad D_C(\mathbf{y}||\mathbf{z}) = \begin{cases} \frac{1}{n} \frac{1}{s(s-1)} \sum_{i \in \mathcal{N}} (y_i^s z_i^{1-s} - 1), & \text{if } s \neq 0, 1, \\ \frac{1}{n} \sum_{i \in \mathcal{N}} y_i \ln \left( \frac{y_i}{z_i} \right), & \text{if } s = 1, \\ \frac{1}{n} \sum_{i \in \mathcal{N}} z_i \ln \left( \frac{z_i}{y_i} \right), & \text{if } s = 0, \end{cases}$$

where “-1” is added when  $s \neq 0, 1$  to obtain a proper generalization of the generalized class of inequality indices. Devooght (2008) provides an empirical application of this class to equality of opportunity in Belgium.

The two different classes of divergence measures,  $D_{MN}(\mathbf{y}||\mathbf{z})$  and  $D_C(\mathbf{y}||\mathbf{z})$ , coincide when  $s = 1$  (the last two terms in  $D_{MN}(\mathbf{y}||\mathbf{z})$  when  $s = 1$  cancel each other out in our case). For this reason, we are going to use  $s=1$  as our main specification. Parameters  $s = 0$  and  $s = 2$  with  $D_{MN}(\mathbf{y}||\mathbf{z})$  will be used as robustness checks. One of the features of  $D_{MN}$  is that a progressive transfer in the actual distribution,  $\mathbf{y}$ , reduces the divergence between  $\mathbf{y}$  and  $\mathbf{z}$  as long as the individuals involved in the transfer have the same reference,  $\mathbf{z}$ . It is a kind of priority given to the worse-off individuals when the individuals involved in the transfer share the same  $\mathbf{z}$ .

<sup>5</sup>The measures proposed by Magdalou and Nock (2011) and Cowell (1985) differ in two important aspects. Both divergence measures generalize the principle of transfers, a cornerstone of inequality measurement, but Magdalou and Nock (2011) impose a weaker generalization of the principle of transfers than Cowell (1985). In addition, and unlike Cowell (1985), Magdalou and Nock (2011) impose judgment separability, which allows measures to be additively decomposed into two components when the reference distribution is egalitarian, but the mean of the reference distribution differs from the mean of the actual distribution. The first component evaluates the divergence between the actual distribution and a hypothetical distribution in which everyone has mean income, while the second one evaluates the efficiency loss brought about by the divergence between the hypothetical distribution in which everyone has mean income and the egalitarian reference distribution.



Moreover, if (and only if)  $s < 2$ , the further down the distribution of  $\mathbf{y}$  such a transfer takes place, the more the divergence between  $\mathbf{z}$  and  $\mathbf{y}$  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 use a divergence measure that is a generalization of the standard Gini coefficient developed by Almås *et al.* (2011), called the Unfairness Gini,  $D_{Gini}(\mathbf{y}||\mathbf{z})$ :

$$(9) \quad D_{Gini}(\mathbf{y}||\mathbf{z}) = \frac{1}{2n(n-1)\mu(\mathbf{y})} \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} |(y_i - z_i) - (y_j - z_j)|.$$

As a final robustness check, we will use a very simple measure of measure of the divergence; the rank correlation between welfare levels and  $\mathbf{h}(\mathbf{a}^C)$ . A large correlation suggests a high degree of inequality of opportunity, since individuals with the best circumstances also have the highest welfare levels. A correlation of 0, in turn, reflects that 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 this measure boils down to a frequently used measure of immobility; Spearman's correlation between parents' and children's income level (see, e.g., Chetty *et al.*, 2014).

### 3. DATA AND MEASUREMENT

We use data from the German Socio-Economic Panel (SOEP), which is a yearly panel that started in 1984. The panel contains detailed questions on household income, life satisfaction, and other well-being dimensions, as well as biographical and historical data that we use to construct circumstance variables. We use data from 1992, the first year in which East Germany can be included in our sample, to 2016, and include all working and unemployed individuals but drop individuals outside the labor market, since several of our effort variables do not extend easily to individuals outside the labor market. We also drop observations with missing values. In total, we have 170,142 person-year observations meeting our baseline specification. These are spread around 21,838 individuals in 15,452 different households.

Our baseline analysis will use the following circumstance variables: gender, father's education (three categories), mother's education (three categories), father's occupation (six categories), a polynomial of age, height, and place of birth (West Germany, East Germany, abroad), degree of urbanization at place of birth (four 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 for which we may hold individuals accountable. One could easily argue that, say, number of hours worked does not lie within individual control. For this reason, we will provide robustness checks where we move these four variables to the other side of the responsibility cut. Summary statistics of the circumstance and effort variables are given in Table A.1 of the Appendix (in the Online Supporting Information).

There are certainly other circumstance and effort variables that could be considered, such as the respondent's ethnicity (circumstance) and the respondent's marital status and number of children (potentially effort variables). We keep the set of circumstance and effort variables somewhat limited in an attempt to avoid inflating the variance of our estimates (Brunori et al., forthcoming).

### 3.1. *Constructing Welfare Variables*

We use four different welfare variables in the analysis. First, we use log 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 adjusted for family size using the OECD equivalence scales. 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 to be 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 Decanq and Neumann (2016). We consider four dimensions; income, health, leisure, and employment.<sup>6</sup> Income is measured the same way as above. Employment 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 number 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's appointments the respondent had last year, and (3) the number of inpatient nights in hospitals that the respondent had last year. To aggregate these sub-dimensions into one health dimension, we regress a health satisfaction

<sup>6</sup>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. Although a similar point can be made with leisure and work hours, which we have on each side of the regression in two of the measures, there is not a direct mapping between leisure and work hours, as individuals can engage in several other activities.

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 dimensions, we standardize the values such that the highest observed level is 1 and the lowest observed 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 that individuals need together with a reference bundle, to make them indifferent to their actual bundle. Although preferences are often estimated from choice behavior, this is difficult when 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.*, 2017).<sup>7</sup> This approach easily accommodates preference heterogeneity by 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, work hours, self-employed, and public sector worker. As the non-income dimensions,  $dim$ , we consider health, employment, and leisure, as in the multidimensional index. Our life satisfaction regression looks as follows:

$$(10) \quad lifesat_{it} = [\beta^{inc} + \gamma^{inc} w_{it}] \ln(inc_{it}) + [\beta^{dim} + \gamma^{dim} w_{it}] dim_{it} + w_{it} + \mu_t + \alpha_i + \varepsilon_{it}.$$

Output from this regression is shown in Table A.2 in the Appendix. In order to calculate the equivalent incomes, we first select a reference vector,  $\tilde{dim}$ , of all dimensions other than income. Here, we choose the mean outcome (the mode for categorical variables), since this avoids favoring any extreme marginal rates of substitution.<sup>8</sup> Then, we calculate the income needed together with  $\tilde{dim}$  for this joint bundle to make individuals indifferent to their actual bundle. That is, we isolate  $\ln(inc_{it}^{eq})$  below:

$$(11) \quad \begin{aligned} lifesat(\ln(inc_{it}), dim_{it}) &= lifesat(\ln(inc_{it}^{eq}), \tilde{dim}) \\ \Leftrightarrow \ln(inc_{it}^{eq}) &= \ln(inc_{it}) + \frac{\beta^{dim} + \gamma^{dim} w_{it}}{\beta^{inc} + \gamma^{inc} w_{it}} (dim_{it} - \tilde{dim}). \end{aligned}$$

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

<sup>7</sup>To the contrary, Benjamin *et al.* (2014) find large and significant differences when comparing the tradeoffs arising from such life satisfaction regressions with the tradeoffs revealed through actual choice behavior.

<sup>8</sup>To see this, suppose that we were to choose the maximum value rather than the mean. Then, we would give individuals with a high preference for income a relatively high level of well-being, as these individuals need a relatively high income together with this reference bundle to be indifferent to their actual bundle.

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 empirical approach implicitly 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 preference heterogeneity. We will do so as a robustness check. A detailed discussion of this method along with an illustration is given in Appendix A.3. Still, our usage of equivalent incomes as a welfare measure that is regressed on income and circumstance variables is in contrast to the typical use of equivalent incomes, which incorporate issues of unfairness directly into the measure. We simply interpret equivalent incomes as another measure of well-being to be used in our analysis.<sup>9</sup>

Histograms of the four final welfare measures are presented in Figure 1.

### 3.2. Estimating Equality of Opportunity

For our empirical specification, we consider well-being to be a linearly separable function of effort and circumstance variables:

$$(12) \quad y_{it} = \beta^C a_{it}^C + \beta^E a_{it}^E + \epsilon_{it}.$$

Two important issues remain unsettled. The first is the issue of how to interpret the error term,  $\epsilon_{it}$ . The error contains omitted effort variables, omitted circumstance variables, measurement error, and general uncertainty. It is unclear whether this should be considered to be within individual control. This is an important decision, as it accounts for most of the variation in the welfare levels. In our baseline specification, we follow the inequality-of-opportunity literature and 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).<sup>10</sup> 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:

$$(13) \quad a_{it}^E = \gamma a_{it}^C + \eta_{it}.$$

<sup>9</sup>A challenge with our use of equivalent incomes is the endogeneity problems generated by having circumstance and effort variables influence the dependent variable as well as entering on the right-hand side of the regressions. Measurement error in one of the variables that we use as preference shifters would create a spurious relationship between the fair outcomes and the actual outcomes. For the preference shifters that we deal with, we conjecture this is a relatively minor concern.

<sup>10</sup>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).

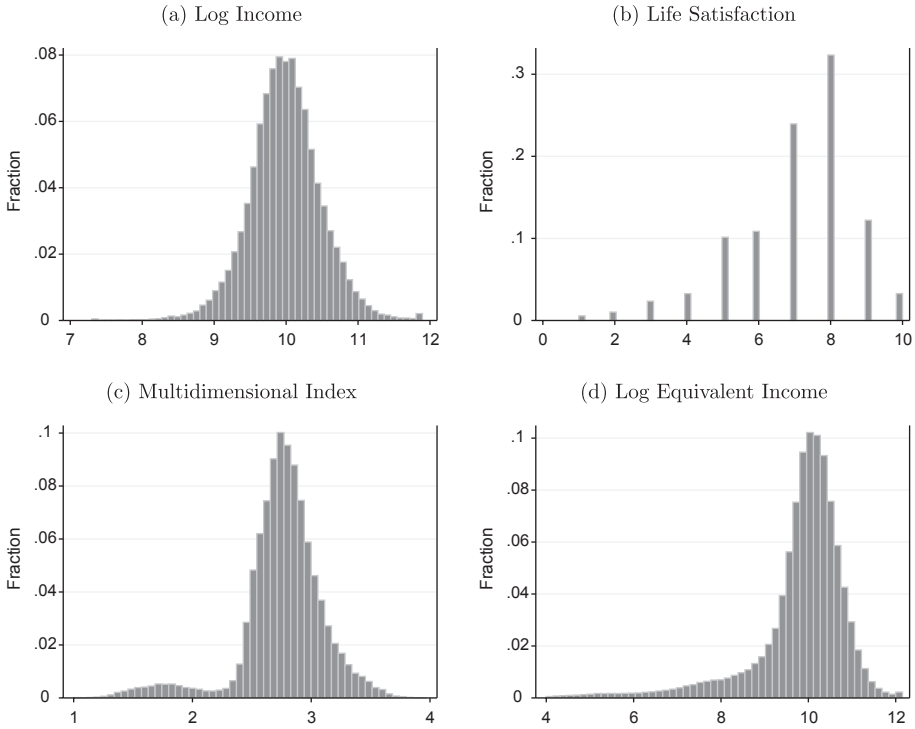


Figure 1. Histograms of Welfare Measures

Notes: Histograms of the four welfare measures. The income and equivalent income distribution is bottom (top) coded at the 0.1th (99.9th) percentile. Life satisfaction is bottom-coded at 1, since values of 0 do not work with some of our divergence measures, which contain the log of welfare.

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

$$(14) \quad y_{it} = (\beta^C + \gamma \beta^E) a_{it}^C + \beta^E \eta_{it} + \epsilon_{it}.$$

Due to the Frisch–Waugh–Lowell theorem, the coefficients on the effort variables will be the same in (12) and (14). The coefficients on the circumstance variables will be different, as they in (14) also incorporate the indirect effect of circumstances on effort.<sup>11</sup> 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

<sup>11</sup>This regression obviously suffers from omitted variable bias, which makes the interpretation of the coefficients from the regression problematic. As long as the regression does not overfit the data, an implication of the omitted variables is that the regression provides a lower-bound estimate of the amount of inequality attributable to circumstances (Ferreira and Gignoux, 2011; Brunori *et al.*, 2018).

to the quality of their circumstances. We calculate each person’s yearly opportunity rank as follows:

$$(15) \quad r_{it}^{opportunity} = H_t[(\beta^C + \gamma\beta^E)a_{it}^C],$$

where  $H_t$  is the yearly cumulative distribution of individuals’ unfair advantage; that is, the part of their well-being due to circumstances. 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 first compute each individual’s fair outcome, using equation (2):

$$(16) \quad z_{it}^{EE} = g(\mathbf{a}_{it}^E) + \frac{\sum_{j \in \mathcal{N}} h(\mathbf{a}_{jt}^C)}{n_t} = \beta^E \eta_{it} + \epsilon_{it} + \frac{\sum_{j \in \mathcal{N}} (\beta^C + \gamma\beta^E)a_{jt}^C}{n_t}.$$

That is, we compute each individual’s fair outcome as the contribution to well-being from the individual’s effort and the average contribution coming from circumstances for all individuals in the sample for the year in question. Second, we utilize our battery of divergence measures to compute the divergence between the fair outcomes and the actual outcomes.

#### 4. RESULTS

##### 4.1. Who Are the Opportunity-Deprived?

Table A.3 in the Appendix shows the results of the regressions from equation (14). Based on this output and equation (15), we calculate each person’s yearly opportunity rank. Table 1 shows the correlations between these opportunity 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.

The opportunity 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 rank correlations by comparing them with the rank correlation between the welfare levels, which are shown in Table 2. In all cases, the correlations are higher when we look at  $r^{opportunity}$ . This indicates that the high-opportunity rank correlations are not driven solely by the measures’ interrelatedness. This also suggests that the way in which welfare is measured matters more if we target the welfare-deprived than if we target the opportunity-deprived.

Since individuals’ opportunities are unobservable, policymakers 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



TABLE 1  
THE CORRELATION BETWEEN OPPORTUNITY RANKS

	Log Income	Life Satisfaction	Multidimensional Index	Log Equivalent Income
Log income	0.53	–	–	–
Life satisfaction	(0.47, 0.57)	–	–	–
Multidimensional index	0.73	0.89	–	–
	(0.70, 0.76)	(0.86, 0.92)		
Log equivalent income	0.89	0.59	0.74	–
	(0.81, 0.93)	(0.50, 0.66)	(0.66, 0.79)	

Notes: Correlations between  $r_{opportunity}$  for the four welfare measures. Bootstrapped 95th percentile confidence intervals in parentheses.

TABLE 2  
THE SPEARMAN CORRELATION BETWEEN WELFARE MEASURES

	Log Income	Life Satisfaction	Multidimensional Index	Log Equivalent Income
Log income	—	—	—	—
Life satisfaction	0.20 (0.18, 0.22)	—	—	—
Multidimensional index	0.42 (0.41, 0.43)	0.20 (0.19, 0.22)	—	—
Log equivalent income	0.68 (0.61, 0.73)	0.27 (0.25, 0.28)	0.63 (0.58, 0.68)	—

*Notes:* Spearman correlation between welfare levels. Bootstrapped 95th percentile confidence intervals in parentheses.

opportunities. We can use the opportunity ranks to test if the characteristics of the opportunity-deprived are similar across the four welfare measures. To do so, we calculate the average opportunity rank for individuals with a given circumstance. The results are shown in Figure 2. The lower the average opportunity rank, the fewer opportunities individuals with the given circumstance have, and the more this circumstance is a potential factor that policymakers can use to target the opportunity-deprived. If the average rank is less than 50, then the particular group has less than average opportunities.

There are many similarities across the welfare measures. Individuals with low-educated parents and individuals whose father was a blue-collar worker or not

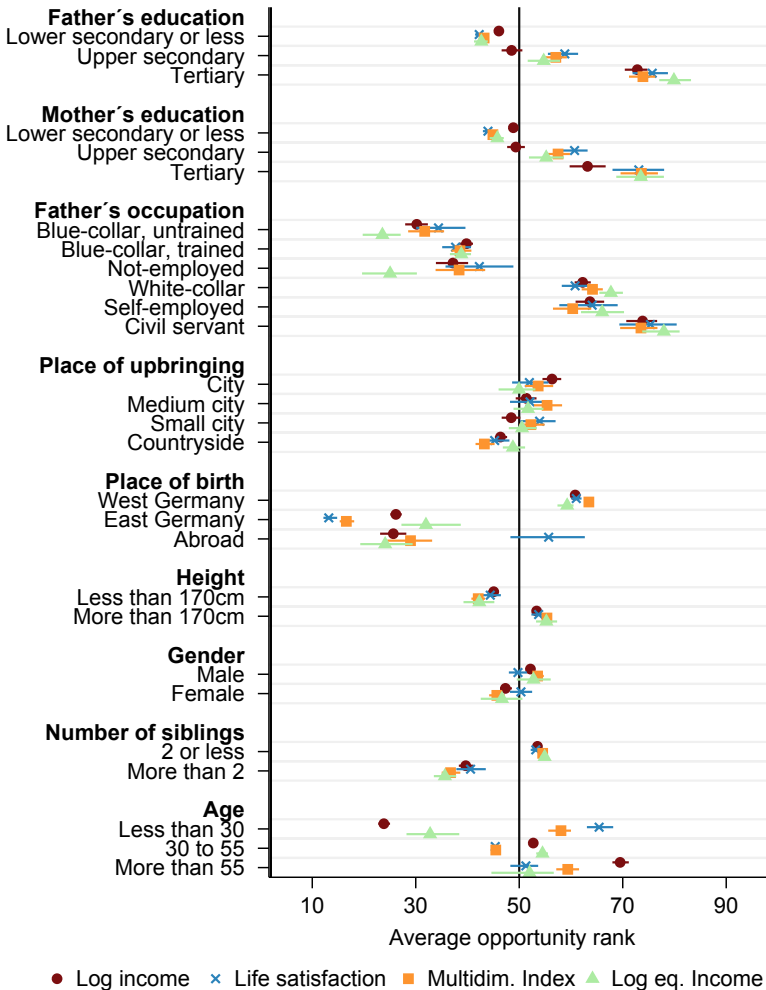


Figure 2. Who Are the Opportunity-Deprived?

Notes: The figure shows the average opportunity rank ( $r_{it}^{opportunity} = H_1(\beta^C + \gamma\beta^E)a_{it}^C$ ) 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 95th percentile confidence bands. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

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.

Meaningful 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 understand the life satisfaction scale more optimistically than Germans. Alternatively, it may be because people born abroad tend to have other circumstances, which are particularly good for life satisfaction. We can indirectly check which effect is driving the result by excluding place of birth from the regression, calculating new opportunity ranks, and recomputing the average opportunity rank of individuals born abroad. With this approach, the direct channel from being born abroad on life satisfaction is omitted. Now the average opportunity rank of people born abroad falls below 50 (not shown in the figure). This suggests that the previous higher rank was solely driven by the direct positive effect of being born abroad on life satisfaction, and that individuals born abroad fare worse than people born in Germany with respect to the remaining circumstances. Hence, it may be that the high-opportunity rank with respect to life satisfaction for people born abroad is solely 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 healthier. Since young people have a lower preference for health (as shown in Table A.2 in the Appendix), young people are not faring better with equivalent incomes, despite the fact that they are more healthy. With respect to life satisfaction, the middle-aged are opportunity deprived while the young and people above 55 are doing better. This mimics the well-known U-curved relationship between subjective well-being and age (Blanchflower and Oswald, 2008; Clark *et al.*, 1996).

Unlike the other categories shown in the figure, individuals will in most cases experience all age categories during their life. It is questionable whether resources should be allocated such that individuals have equal opportunities in every part of their life. For this reason, later on we will 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 life cycle rather than in every point of their life (for a similar approach, see Almås *et al.*, 2011).

In sum, there is relatively large agreement about who the opportunity-deprived are across the four measures. Hence, if a policymaker strives to target individuals with low opportunities, it matters relatively little how welfare is measured. Although this suggests that equalizing opportunities does not depend on how welfare is measured, it may very well be the case that the public policies needed to prevent inequality of opportunity from arising in the first place depend on which measure of welfare is used. To study this properly, a causal framework is needed, which goes beyond the scope of this paper.

#### 4.2. *Equality of Opportunity over Time*

Before analyzing how equality of opportunity has evolved over time, and whether this depends on the well-being measure used, we start by analyzing how the level of well-being has evolved over time and how inequality in well-being has developed over time. Panel (a) of Figure 3 shows the level of well-being in Germany from 1992 to 2016. In the figure, and in the time-series figures to follow, we use 3-year moving averages to smoothen out erratic trends. The level of well-being is normalized to 100 in 1992 to foster comparisons between the different measures.

The figure shows that from around 2005, well-being has increased with all four measures, but prior to that, life satisfaction and the multidimensional index followed different patterns than the remaining two measures. Panel (b) of Figure 3 shows the development in inequality in the four welfare measures using the Theil index. Inequality in log incomes has increased substantially over most of the period and inequality in life satisfaction has decreased over the past 10 years, while inequality in the other two measures has followed intermediate trends.<sup>12</sup>

Thus, the measure of welfare matters both when we look at the development of welfare and the inequality of welfare over the past 25 years in Germany. How do things look for the development of inequality of opportunity over time? This is displayed in Figure 4.

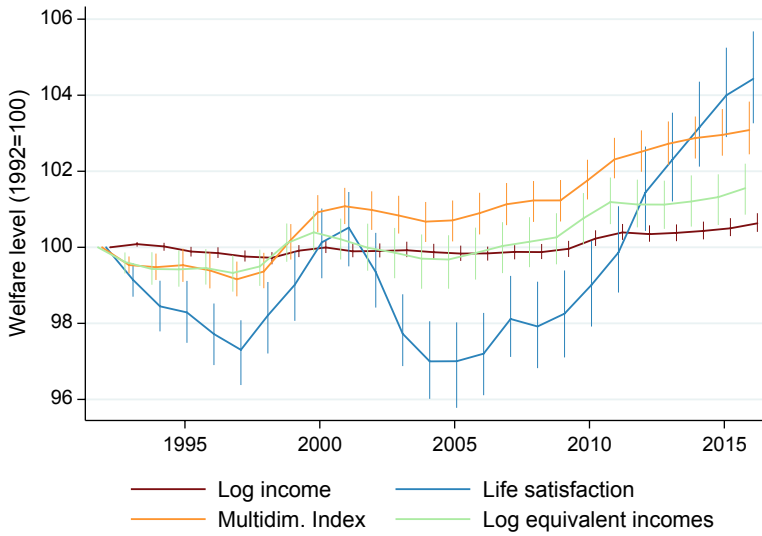
Although the confidence bands are quite wide, the measures follow broadly the same pattern. From 1992 to 2005, inequality of opportunity increased, and from then and until 2016, it gradually fell. Some of the welfare measures, particularly equivalent incomes, are noisier, which can be explained by the added uncertainty from the fairly complicated process of constructing this welfare measure. Our setup does not allow us to speak to the causes of the changes in inequality of opportunity, but we note that the fall from around 2005 coincides with the implementation of the Hartz plan, which could explain the pattern.

#### 4.3. *Altering the Responsibility Cut*

The analysis above was based on important normative assumptions regarding what individuals were held responsible for. We assumed that individuals were responsible for four variables (*ivar*): their education, work hours, and whether they are (i) self-employed or (ii) 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 circumstances and effort 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

<sup>12</sup>For a more detailed analysis of inequality in life satisfaction over time, see Clark *et al.* (2016).

(a) Level of Well-Being



(b) Inequality of Well-Being

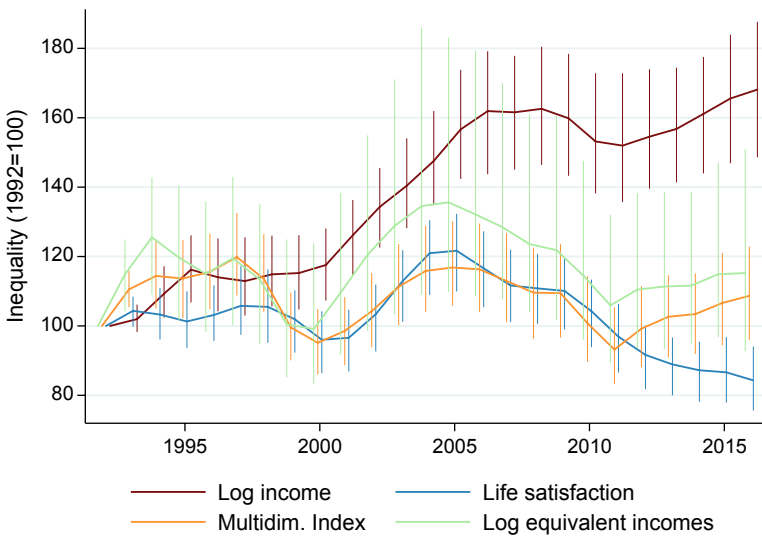


Figure 3. Development in the Level and Inequality of Well-Being, 1992–2016

Notes: Development in the level of well-being and inequality in well-being from 1992 to 2016. All measures are normalized to equal 100 in 1992 to foster comparisons. Inequality is measured using the Theil index. Bars indicate bootstrapped 95th percentile confidence bands. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

assumed that individuals were not responsible for well-being differences due to preference heterogeneity arising from circumstances (*pref*).



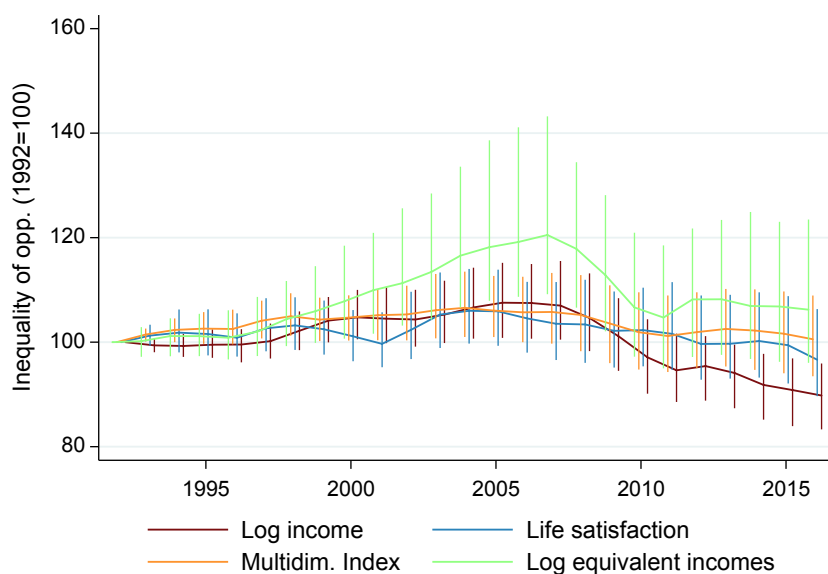


Figure 4. Inequality of Opportunity over Time

Notes: Development in inequality of opportunity in each well-being variable from 1992 to 2016 using the Magdalou–Nock divergence measure with  $s = 1$  and the Egalitarian Equivalent allocation rule. Bars indicate bootstrapped 95th percentile confidence bands. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Our baseline effort set contained *4var* and *residual*. In this section, we try to shift the responsibility cut by altering 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 5 and 6. Figure 5 uses our most narrow effort set, only *4var*, while Figure 6 uses the widest possible effort set, *4var*, *residual*, *cor*, *age*, *pref*.

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 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 40 with the smallest effort set and 15 with the largest effort set. In other words, East Germans are not doing too poorly if we only hold individuals responsible for four variables (which almost mimics just looking at outcomes), but the more we hold individuals responsible for, the more disadvantaged they are. This suggests that East Germans have relatively high effort levels.

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. The results are displayed in Figure 7. 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.

Our baseline results are displayed in panel (c). If we no longer hold individuals responsible for the four variables we had deemed effort (panel (b)), then our

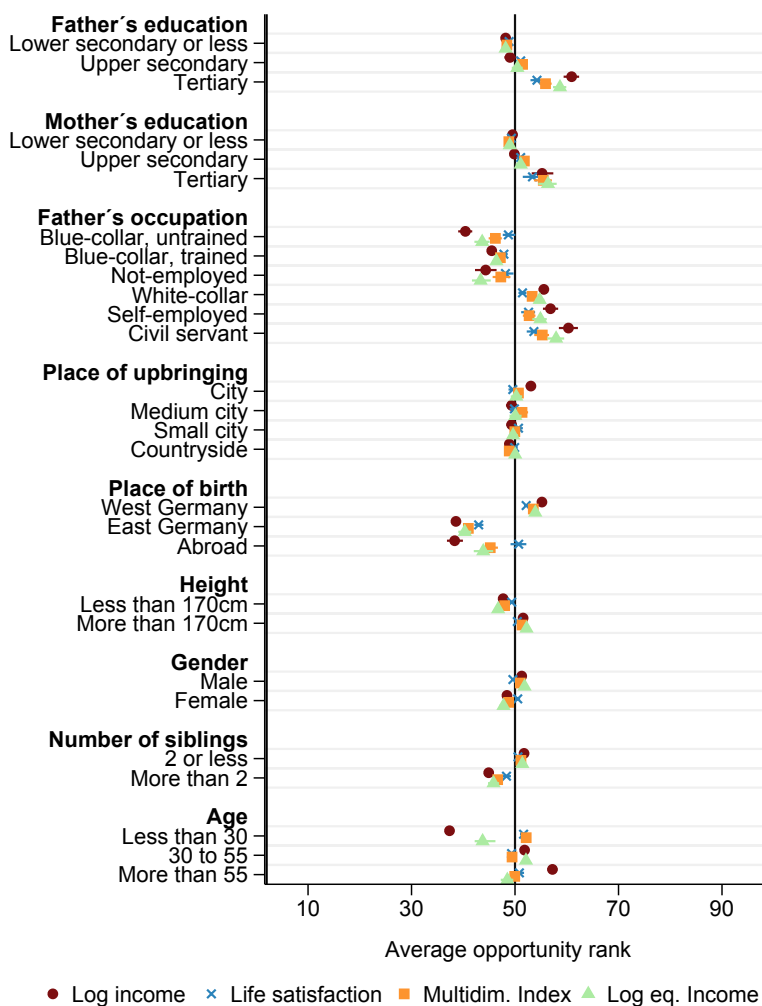


Figure 5. Changing the Responsibility Cut:  $\text{Effort} = \{4\text{var}\}$

Notes: The figure shows the average opportunity rank for people sharing a particular circumstance using a narrow effort set. Bars indicate bootstrapped 95th percentile confidence bands. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

main results become more clear; inequality of opportunity increased until about 2005 and then decreased for all four measures. If we only hold individuals responsible for the four variables and shift the residual to the other side of the responsibility cut (panel (a)), then the picture looks very different, particularly for log income. Since the four effort variables are able to explain very little of the variance, the measure almost boils down to the development in inequality over time, which, as we have already established, follows different trends. Other studies that have switched the residual to the other side of the responsibility cut likewise found this to have a great impact on the results (see, e.g., Devooght, 2008; Almås *et al.*, 2011).

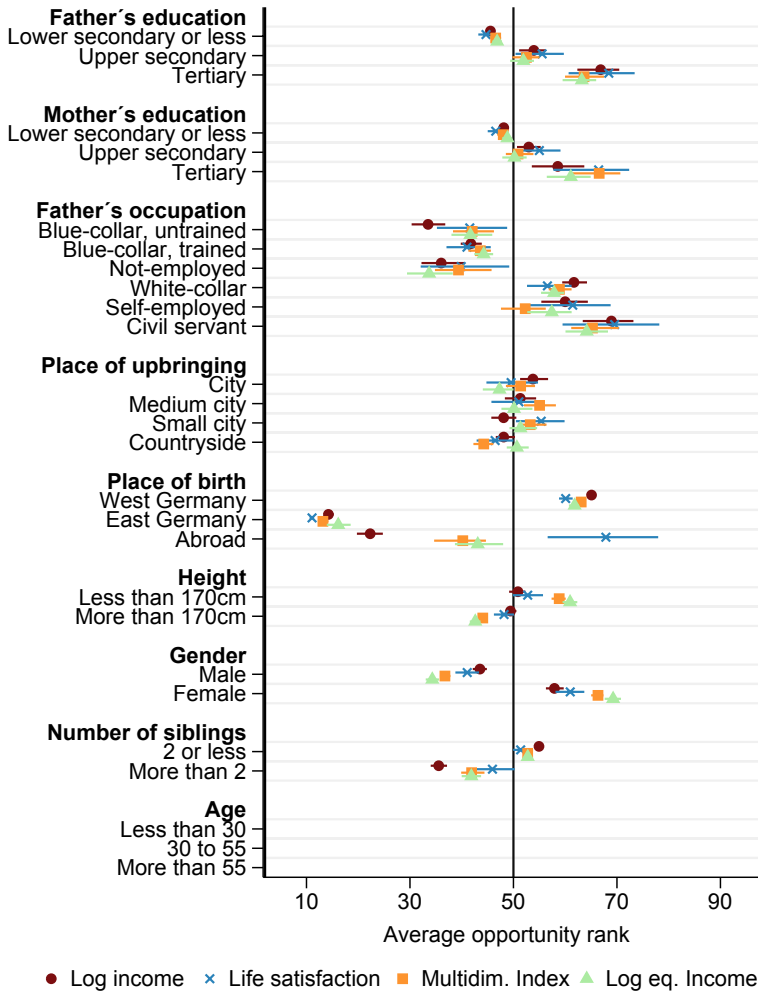


Figure 6. Changing the Responsibility Cut: Effort = {4var, residual, cor, age, pref}  
 Notes: The figure shows the average opportunity rank for people sharing a particular circumstance using a wide effort set. We no longer report differences in the average opportunity rank by age groups, as age is not considered a circumstance in this specification. Bars indicate bootstrapped 95th percentile confidence bands. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

If we also hold individuals responsible for the correlation between effort and circumstances (panel (d)), then we still see an increase in inequality of opportunity in all measures from 1992 to 2005, but from then on the multidimensional index diverges from the rest. In the other cases where we add more to the responsibility cut (panels (e) and (f)), the overall trend changes, but stays broadly similar across the four well-being measures. In these cases, inequality of opportunity has remained rather flat or decreased a little from 1992 to 2016.

In sum, we find that when characterizing the opportunity-deprived, neither the measure of welfare nor the precise location of the responsibility cut is of great

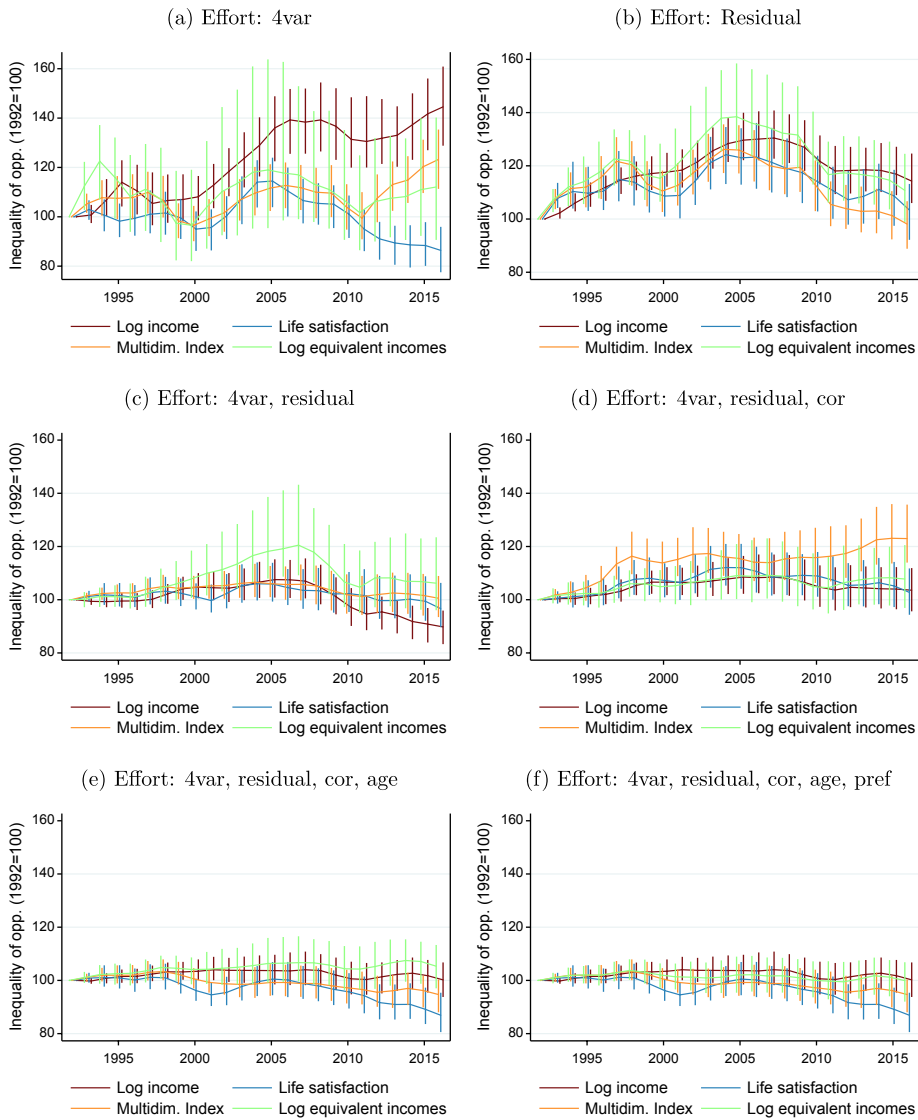


Figure 7. Altering the Responsibility Cut

Notes: Development in inequality of opportunity in each well-being variable from 1992 to 2016 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 that we are equalizing lifetime opportunities). *Cor*: The correlation between effort and circumstance variables is not considered a circumstance. *Pref*: Individuals are held to be fully responsible for their preferences (this only applies to the equivalent income measure). Our baseline specification used *4var* and *residual* as effort. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

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. Although the trend may change depending on what we hold

individuals responsible for, for a given responsibility cut, the well-being measures follow broadly the same trend.

#### 4.4. *Robustness Checks*

In this section, we test whether our findings are sensitive to using other divergence measures and fair allocation rules. Our primary interest is not whether the trends change as we change the method but, rather, whether the four well-being measures follow similar trends regardless of what method we use. First, we try to use the Generalized Proportional allocation rule rather than the Egalitarian Equivalent allocation rule. Since the results now will depend on which norm vector we use, we show the results using the mode/mean circumstance (for categorical and continuous variables, respectively), as well as the worst and best circumstances, as defined by their bivariate relationship with welfare. The results are shown in Figure A.2 in the Appendix. Although large confidence bands prevent us from nailing down the trends with precision, particularly when the worst circumstances are used as the norm vector, in the other two cases, point estimates suggest that the measures follow broadly the same trend.

Next, we try to use the Magdalou–Nock divergence measures with  $s = 0$  and  $s = 2$  rather than  $s = 1$ . The results are shown in Figure A.3 in the Appendix. Again, we face wide confidence intervals. When  $s = 0$ , point estimates suggest that inequality of opportunity has decreased over the past 15 years with all measures. Things look a bit different with  $s = 2$ , where no trend is visible over the past 10 years with life satisfaction and the multidimensional index, while inequality of opportunity for income decreased. Since  $s = 2$  gives less weight to the bottom of the distribution, this suggests that the improvements in equality of opportunity found in our baseline scenario have mostly arisen since the most opportunity-deprived were given greater chances. At the top of the distribution, the relationship between opportunities and outcomes remains equally strong.

We also try using different divergence measures; the Fairness Gini and the correlation between opportunity ranks and welfare ranks. The latter comes with a number of distinct advantages: (1) the measure is constrained to be between 0 and 1, so there is little need to index the numbers to 1992 = 100; (2) the confidence bands are narrower; (3) it minimizes our reliance on cardinality, which is problematic in the case of life satisfaction; (4) it ensures that we are comparing well-being measures that follow the same distributions; (5) it is more intuitive to understand; (6) it allows that zero and negative well-being levels can easily be included in the analysis; and (7) it generalizes a frequently used measure of intergenerational mobility, and as such the results can be compared with certain intergenerational mobility studies. The results are displayed in Figure A.4 in the Appendix. Both with the Fairness Gini and the Spearman correlation, the results suggest that inequality of opportunity has improved for all four measures over the past 10–15 years.

## 5. CONCLUSION

We have investigated whether equality-of-opportunity estimates depend on what, precisely, we seek to equalize opportunities for. On the basis of the

philosophical literature on well-being, we have constructed four measures of welfare that are candidates for what we ought to equalize opportunities for. Upon constructing these, we have analyzed if the way in which welfare is measured matters for (1) characterizing the opportunity-deprived and (2) tracking inequality of opportunity over time. We have 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 policymakers 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 to equalize opportunities. 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 M. Fleurbaey, *The Oxford Handbook of Well-Being and Public Policy*, Oxford University Press, New York, 2016.
- Akay, A., O. Bargain, and H. X. Jara, "Back To Bentham: Should We? Large-Scale Comparison of Decision versus Experienced Utility for Income-Leisure Preferences," *IZA Discussion Paper Series 10907*, 2017.
- Almås, I., A. W. Cappelen, J. T. Lind, E. Ø. Sørensen, and B. Tungodden, "Measuring Unfair (In) equality," *Journal of Public Economics*, 95(7), 488–99, 2011.
- Arneson, R. J., "Equality and Equal Opportunity for Welfare," *Philosophical Studies*, 56(1), 77–93, 1989.
- Barry, B., *Why Social Justice Matters*, Polity Press, Malden, MA, 2005.
- Benjamin, D. J., O. Heffetz, M. S. Kimball, and A. Rees-Jones, "Can Marginal Rates of Substitution be Inferred from Happiness Data? Evidence from Residency Choices," *The American Economic Review*, 104(11), 3498–528, 2014.
- Blanchflower, D. G. and A. J. Oswald, "Is Well-Being U-Shaped over the Life Cycle?" *Social Science & Medicine*, 66(8), 1733–49, 2008.
- , "Well-Being over Time in Britain and the USA," *Journal of Public Economics*, 88(7), 1359–86, 2004.
- Bossert, W., "Redistribution Mechanisms Based on Individual Characteristics," *Mathematical Social Sciences*, 29(1), 1–17, 1995.
- Bossert, W. and M. Fleurbaey, "Redistribution and Compensation," *Social Choice and Welfare*, 13(3), 343–55, 1996.
- Brunori, P., P. Hufe, and D. G. Mahler, "The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees," *World Bank Policy Research Working Paper 8349*, 2018.
- Brunori, P., V. Peragine, and L. Serlenga, "Upward and Downward Bias when Measuring Inequality of Opportunity," *Social Choice and Welfare*, <https://link.springer.com/article/10.1007/s00355-018-1165-x>, forthcoming.
- Calsamiglia, C., "Decentralizing Equality of Opportunity," *International Economic Review*, 50(1), 273–90, 2009.
- Cappelen, A. W. and B. Tungodden, "Fairness and the Proportionality Principle," *Social Choice and Welfare*, 49(3–4), 709–19, 2017.
- Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner, "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," *The American Economic Review*, 104(5), 141–7, 2014.
- Clark, A. E., S. Flèche, and C. Senik, "Economic Growth Evens Out Happiness: Evidence from Six Surveys," *Review of Income and Wealth*, 62(3), 405–19, 2016.
- Clark, A. E., A. Oswald, and P. Warr, "Is Job Satisfaction U-Shaped in Age?" *Journal of Occupational and Organizational Psychology*, 69(1), 57–81, 1996.



- Cohen, G. A., "Equality of What? On Welfare, Goods and Capabilities," *Recherches Économiques de Louvain/Louvain Economic Review*, 56(3–4), 357–82, 1990.
- Cowell, F. A., "Measures of Distributional Change: An Axiomatic Approach," *The Review of Economic Studies*, 52(1), 135–51, 1985.
- Crisp, R., "Well-Being," in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Stanford University, <https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=well-being>, 2016.
- Decancq, K., M. Fleurbaey, and E. Schokkaert, "Happiness, Equivalent Incomes and Respect for Individual Preferences," *Economica*, 82(s1), 1082–106, 2015.
- Decancq, K. and D. Neumann, "Does the Choice of Well-Being Measure Matter Empirically?" in M. D. Adler and M. Fleurbaey (eds), *The Oxford Handbook of Well-Being and Public Policy*, Oxford University Press, New York, 553–87, 2016.
- Devooght, K., "To Each the Same and to Each His Own: A Proposal to Measure Responsibility-Sensitive Income Inequality," *Economica*, 75, 280–95, 2008.
- Dworkin, R., "What is Equality? Part 2: Equality of Resources," *Philosophy & Public Affairs*, 10(4), 283–345, 1981.
- Ferreira, F. H. G. and J. Gignoux, "The Measurement of Inequality of Opportunity: Theory and an Application to Latin America," *Review of Income and Wealth*, 57(4), 622–57, 2011.
- Ferreira, F. H. G. and V. Peragine, "Individual Responsibility and Equality of Opportunity," in M. D. Adler and M. Fleurbaey (eds), *The Oxford Handbook of Well-Being and Public Policy*, Oxford University Press, New York, 746–84, 2016.
- Fitoussi, J.-P., A. Sen, and J. E. Stiglitz, *Mismeasuring Our Lives: Why GDP Doesn't Add Up*, The New Press, New York, 2010.
- Fleurbaey, M., "On Fair Compensation," *Theory and Decision*, 36(3), 277–307, 1994.
- \_\_\_\_\_, "The requisites of equal opportunity", in W. Barnett, H. Moulin, M. Salles et N. Schofield, eds, *Social Choice, Welfare and Ethics*, Cambridge University Press, New York, 1995.
- Fleurbaey, M. and F. Maniquet, *A Theory of Fairness and Social Welfare*, Cambridge University Press, New York, 2011.
- Fleurbaey, M. and E. Schokkaert, "Unfair Inequalities in Health and Health Care," *Journal of Health Economics*, 28(1), 73–90, 2009.
- Griffin, J., *Well-Being: Its Meaning, Measurement, and Moral Importance*, Clarendon Press, Oxford, 1986.
- Jusot, F., S. Tubeuf, and A. Trannoy, "Circumstances and Efforts: How Important is Their Correlation for the Measurement of Inequality of Opportunity in Health?" *Health Economics*, 22(12), 1470–95, 2013.
- Magdalou, B. and R. Nock, "Income Distributions and Decomposable Divergence Measures," *Journal of Economic Theory*, 146(6), 2440–54, 2011.
- Parfit, D., *Reasons and Persons*, Clarendon Press, Oxford, 1984.
- Ramos, X. and D. Van de gaer, "Empirical Approaches to Inequality of Opportunity: Principles, Measures, and Evidence," *Journal of Economic Surveys*, 30(5), 855–83, 2016.
- \_\_\_\_\_, "Is Inequality of Opportunity Robust to the Measurement Approach?" *IZA Discussion Paper Series 11157*, 2017.
- Ravallion, M., "Inequality and Poverty when Effort Matters," *Econometrics*, 5(4), 50, 2017.
- Rawls, J., *A Theory of Justice*, Harvard University Press, Cambridge, MA, 1971.
- Roemer, J. E., "A Pragmatic Theory of Responsibility for the Egalitarian Planner," *Philosophy & Public Affairs*, 22(2), 146–66, 1993.
- \_\_\_\_\_, *Equality of Opportunity*, Harvard University Press, Cambridge, MA, 1998.
- \_\_\_\_\_, "On Several Approaches to Equality of Opportunity," *Economics and Philosophy*, 28(2), 165–200, 2012.
- Roemer, J. E. and A. Trannoy, "Equality of Opportunity," in A. B. Atkinson and F. Bourguignon (eds), *Handbook of Income Distribution*, Volume 2A–2B, North Holland, Amsterdam, 217–300, 2015.
- \_\_\_\_\_, "Equality of Opportunity: Theory and Measurement," *Journal of Economic Literature*, 54, 1288–332, 2016.
- Sen, A., *Commodities and Capabilities*, North-Holland, Amsterdam, 1985.
- \_\_\_\_\_, "Equality of What?" In S. M. McMurrin (ed.), *The Tanner Lecture on Human Values*, Vol. 1, University of Utah Press, Salt Lake City/Cambridge University Press, Cambridge, UK, 197–220, 1980.
- Stevenson, B. and J. Wolfers, "Happiness Inequality in the United States," *Journal of Legal Studies*, 37(S2), S33–79, 2008.
- Swift, A., *Justice, Luck and the Family*, Princeton University Press, Princeton, NJ, 2005.
- Van de gaer, D., *Equality of Opportunity and Investment in Human Capital*, Katholieke Universiteit Leuven, Leuven, 1993.

## Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site:

**Appendix A.1:** Summary Statistics

**Table A.1:** Summary Statistics A.2: Equivalent Income Regression Output

**Table A.2:** Equivalent Income Regression A.3: Decomposition of Equivalent Incomes

**Figure A.1:** Equivalent Income Illustration A.4: Regression Output

**Table A.3:** Regressing Welfare on Circumstances and Effort A.5: Robustness Checks

**Figure A.2:** Changing the Allocation Rule: Generalized Proportionality Principle

**Figure A.3:** Changing the Parameter of the Magdalou–Nock Divergence Measure

**Figure A.4:** Changing the Divergence Measure