I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11405

Upward and Downward Bias When Measuring Inequality of Opportunity

Paolo Brunori
Vito Peragine
Laura Serlenga

MARCH 2018

DISCUSSION PAPER SERIES

IZA DP No. 11405

Upward and Downward Bias When Measuring Inequality of Opportunity

Paolo Brunori
University of Florence and University of Bari

Vito Peragine
University of Bari

Laura Serlenga
University of Bari and IZA

MARCH 2018

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

Upward and Downward Bias When Measuring Inequality of Opportunity

Estimates of the level of inequality of opportunity have traditionally been interpreted as lower bounds due to the downward bias resulting from the partial observability of circumstances that affect individual outcome. We show that such estimates may also suffer from upward bias as a consequence of sampling variance. The magnitude of the latter distortion depends on both the empirical strategy used and the observed sample. We suggest that, although neglected in empirical contributions, the upward bias may be significant and challenge the interpretation of inequality of opportunity estimates as lower bounds. We propose a simple criterion to select the best specification that balances the two sources of bias. Our method is based on cross-validation and can easily be implemented with survey data. To show how this method can improve the reliability of inequality of opportunity measurement, we provide an empirical illustration based on income data from 31 European countries. Our evidence shows that estimates of inequality of opportunity are sensitive to model selection. Alternative specifications lead to significant differences in the absolute level of inequality of opportunity and to the re-ranking of a number of countries. This confirms the need for an objective criterion to select the best econometric model when measuring inequality of opportunity.

JEL Classification: C52, D3, D63

Keywords: inequality of opportunity, model selection, cross-validation, variance-bias trade-off

Corresponding author:
Paolo Brunori
Dipartimento di Scienze per l'Economia e l'Impresa
Via delle Pandette 32
50127 Firenze
Italy
E-mail: paolo.brunori@unifi.it

1 Introduction

The measurement of inequality of opportunity (IOp hereafter) is a growing topic in Economics, and in the past two decades, the number of empirical contributions to this literature has increased substantially: see Ferreira and Peragine (2016), Roemer and Trannoy (2015), and Van de Gaer and Ramos (2016) for a review. The vast majority of these contributions is based on the approach proposed by Roemer (1998) and follows a two-step procedure.[1] First, starting from an outcome distribution (typically income or consumption), a counterfactual distribution is derived, which reproduces only unfair inequalities, i.e., inequalities due to circumstances beyond the individual responsibility, and does not reflect inequality arising from individual choice and effort. Second, a suitable inequality measure is used to quantify inequality in the counterfactual distribution. The empirical literature has used two classes of methods to compute counterfactual distributions of survey data: parametric and non-parametric methods. One of the main drawbacks of both approaches is that, unless all the circumstances beyond an individuals responsibility are observable, they produce biased estimates of IOp. While the magnitude of this bias may be impossible to determine (Bourguignon et al., 2013), under some assumptions it can be shown that the sign of the bias is negative (Roemer, 1998; Ferreira and Gignoux, 2011; Luongo, 2011). This explains why IOp estimates are generally interpreted as lower-bound estimates of the true IOp, whereas the true IOp is interpreted as the estimate one would obtain if all circumstances were observable. The usefulness of those lower-bound measures has been challenged in the recent literature: see Kanbur and Wagstaff (2016), Balcazar (2015), and Wendelspiess (2015). In particular, Balcazar (2015) and Ibarra et al. (2015) suggest that the downward bias may lead to a substantial underestimation of the true IOp in empirical applications. Typically, authors address this problem by using richer data sources and by adopting a variety (or a combination) of empirical strategies: (i) by increasing the number of circumstances, as in Biorklund et al. (2012); (ii) by introducing interaction terms among different circumstances, as in Hufe and Peichl (2015); (iii) by splitting the circumstances into finer partition of categories. These empirical

[1] In this paper we focus on the so called "ex ante" approach. For a comparison of the "ex ante" and "ex post" approaches see Fleurbaey and Peragine (2013).

strategies reduce the downward bias by increasing the explained variability attributable to IOp. In this paper, we emphasize that these procedures are not exempt to risk and might lead to an upper distortion of IOp estimates. Indeed, the reliability of the estimates depends not only on the number of circumstances and the partition of circumstances into categories, but also on the sample distribution. In both parametric and non-parametric approaches, we recognize a trade-off between the downward bias resulting from the observability of circumstances and the upward bias related to the sampling variance of the estimated counterfactual distribution. Although this topic is not new to econometricians and practitioners, the problem of possible upward-biased IOp estimates has been neglected in the literature on IOp measurement. This is surprising because, as we show in the empirical section, such a distortion is likely to be far from negligible. We show that the magnitude of the upward distortion depends upon the strategy used to obtain the counterfactual distribution. This problem is particularly straightforward when applying a parametric approach but can easily be generalized to the non-parametric method. The number of explanatory variables involved and the division into categories, may lead to distortions in both directions: overfitted models result in upward bias, whereas underfitted models reinforce the well-known downward bias because of partial observability. We suggest that, when choosing among alternative specifications, scholars should opt for the best balance between the two sources of bias, and we propose a method to select the best econometric specification that minimizes both types of bias. Our method is based on cross-validation (CV hereafter), which is a methodology commonly adopted by statisticians to evaluate the performance of predictive models and is increasingly used by economists (Varian, 2014). CV directly provides a nearly unbiased measure of the true out-of sample prediction error. The major interest of CV lies in the minimal assumptions required to obtain unbiased measures of model performance (Arlot and Celisse, 2010). The out-of-sample prediction error is estimated by dividing the original sample into training and test sets. The association between circumstances and outcome is first estimated on the training sample under a large number of meaningful model specifications. Next, the derived coefficients are used to predict the outcome on the test sample. The specification selected is the model that, on average, minimizes the prediction error in the test sample. This model selection technique is widely adopted in statistical learning; many routines have been developed and can easily be implemented in commonly used software. To demonstrate the usefulness of our approach, we apply our method to income data from 31 European countries using the EU Survey on Income and Living Conditions (EU-SILC) 2011 database. Our evidence shows that IOp estimates are sensitive to model selection. Alternative specifications lead to significant differences in the absolute level of IOp and, in many cases, to the re-ranking of countries. Our preferred specification is different from the typical model used in the literature; therefore, our estimates differ from those provided by other authors who use the same data to estimate IOp. The rest of this paper is organized as follows: Section 2 introduces the canonical model used to measure IOp and the estimation methods typically used to implement it, and discusses the two possible sources of distortion. Section 3 proposes the Cross-Validation methodology to balance the trade-off between the two types of bias when selecting the specification to estimate IOp. Section 4 presents an empirical implementation, and Section 5 concludes.

2 Downward and upward biased IOp

The canonical equality of opportunity model can be summarized as follows (see Ferreira and Peragine, 2016). Each individual in a society realizes an outcome of interest, y, by means of two sets of characteristics: circumstances beyond individual control, C, belonging to a finite set $\Omega = \{C_1, ..., C_J\}$, and a responsibility variable, e, typically treated as scalar. A function $g : \Omega \times \Re_+ \rightarrow \Re_+$ defines the individual outcome:

$$y = g(C, e).$$

For all $j \in \{1, ..., J\}$, let us denote by K_j the set of possible values taken by circumstance C_j and by $|K_j|$ the cardinality of K_j. For instance, if C_j denotes gender, then $K_j = \{male, female\}$. We can now define a partition of the population into T types, where a type is a set of individuals who share exactly the same circumstances; that is, $T = \Pi_{j=1}^{J} |K_j|$. Let us denote by Y the overall outcome distribution.

The IOp is then defined as the inequality in the counterfactual distribution, $\tilde{Y}$, which reproduces all inequalities due to circumstances and does not reflect any inequality due to effort. A number of methods have been proposed to obtain $\tilde{Y}$, and in general, the selected method affects the resulting IOp measure (Ferreira and Peragine, 2016; Roemer and Trannoy, 2015; Van de gaer and Ramos, 2016). In what follows, we focus on the *ex ante* approach introduced by Bourguignon et al. (2007) and Checchi and Peragine (2010), which is by far the most commonly adopted method in the empirical literature (Brunori et al., 2013)[2]. This approach interprets the type-specific outcome distribution as the opportunity set of individuals belonging to each type. Then, a given value v_t of the opportunity set for any type t, with $t = 1, ..., T$, is selected. Finally, $\tilde{Y}$ is obtained by replacing the outcome of each individual belonging to type t with the value of her type v_t, for all $t = 1, ..., T$.

2.1 Counterfactual estimation

Ex ante IOp can be estimated by either a non-parametric or a parametric approach. Checchi and Peragine (2010) propose a non-parametric estimation of $\tilde{Y}$ following the typical two-stage method: (i) after partitioning the sample into types on the basis of all observable circumstances, they choose the arithmetic mean of type t, denoted by μ_t, as the value v_t of type t; (ii) the counterfactual distribution is constructed by replacing for each individual i belonging to type t, its outcome, y_i with $\tilde{y}_i = \hat{\mu}_t$ — where $\hat{\mu}_t$ is the sample estimate for μ_t — and an inequality measure is applied to the counterfactual distribution $\tilde{Y}$.

Alternatively, Bourguignon et al. (2007) propose parametric measurement of *ex ante* IOp by estimating $\tilde{Y}$, as the prediction of the following reduced form regression:

$$y_i = \sum_{j=1}^{J} \sum_{k=1}^{K_j} \chi_{j_k} c_{ij_k} + u_i, \tag{1}$$

[2] Other well-established approaches can be used to measure IOp. Approaches differ in how they define the principle of equal opportunity and in the way the counterfactual distribution is constructed (Roemer, 1998; Lefranc et al., 2009; Fleurbaey and Shockaert, 2009; Checchi and Peragine, 2010). However, because the construction of these alternative counterfactual distributions generally requires the observation or identification of effort (an extremely difficult variable to measure), they are less frequently adopted in the empirical literature.

where c_{ij_k} identifies each category of the observable characteristics by means of a dichotomous variable, and χ_{j_k} is the corresponding coefficient[3]. In the original specification, the parametric approach consists of ordinary least squares regression where the total outcome variability is explained by a linear combination of regressors with no interaction terms[4]. The parametric approach does not estimate the counterfactual distribution, $\tilde{Y}$, by directly identifying types. It linearly approximates the types' average outcome by estimating the fixed effect of each circumstance on outcome and the regression coefficients and, therefore, obtaining the predicted outcomes. This approach has the main advantage of being much more parsimonious than the non-parametric alternative. In practice, parametric estimations have been proposed as a reasonable choice when few observations are available, see Ferreira and Gignoux (2011) and Ibarra et al. (2015). However, parsimony comes at the cost of imposing the effect of the circumstances on outcome to be fixed and additive. For example, being a women is assumed to have an effect on earning that is independent of all the other circumstances, such as socioeconomic background and race. This assumption constrains the ability of regressors to capture outcome variability.

Recently, Marrero and Rodriguez (2011) and Hufe and Peichl (2015) discuss the importance of considering interaction terms in estimating IOp. Hufe and Peichl (2015) estimate ex ante IOp using the Child & Young Adults Supplement of the National Longitudinal Survey of Youth and alternative model specifications. They implement both a linear model, as in equation (1), and a non-linear model, where circumstances fully interact, and they acknowledge a critical divergence of the IOp estimates among the different specifications.

Indeed, it is important to note that the parametric and the non-parametric methods coincide when all explanatory variables are categorical, and the parametric counterfactual distribution is obtained by the prediction of a regression model where y is regressed on all possible combinations of circumstance values, i.e., all values of all regressors interact with each other, to obtain a model with $T = \Pi_{j=1}^{J}|K_j|$ dummies. In this particular case, each regressor captures the effect of

[3]Note that in principle, one could have non-categorical regressors if cardinal circumstances are observed. However, to the best of our knowledge, this is never the case. Even when a cardinal measure is available, such as parental income or parental years of education, authors tend to construct quantiles and use them as regressors (see, for example, Björklund et al., 2012).

[4]Analogously to the Mincer equation, a log-linear specification is preferred by some authors. (Ferreira and Gignoux, 2011)

belonging to one of all the possible circumstance combinations, which is the effect of belonging to a given type. The estimated model becomes

$$y_i = \sum_{t=1}^{T} \beta_t \pi_{it} + u_i, \tag{2}$$

where π_{it} are T binary variables obtained by interacting all categories of all circumstances. Clearly, the typical (linear) parametric approach, equation (1), explains less inequality than the non-parametric approach, equation (2), simply because model (2) — by construction — allows variability to be explained by the full set of interactions.

Here, a trade-off emerges: while the linear specification might be too restrictive, the inclusion of the full set of combinations among categories might lead to very large sampling variance of the estimated counterfactual distribution, especially when a limited number of observations is available for certain types.

By following the same reasoning, the sampling variance of the estimated counterfactual distribution is also influenced by alternative partitions into categories of observed circumstances: a broadest partition might, again, lead to larger variance in the case of a limited number of observations per type.

Indeed, the reliability of both parametric and non-parametric IOp estimates requires a sufficient number of observations characterizing each circumstance. Specifically, the limitation might be more severe in the case of the non-parametric approach, where a sufficient number of observations for each combination of circumstances is required. In empirical applications, this might represent a serious constraint as in survey data individuals are unlikely to be uniformly distributed across types and across category partitions. For example, a typical argument arises when considering Western countries in which researchers observe both parental education and parental occupation as circumstances. Those variables are usually strongly correlated with each other, i.e., there are very few individuals whose parents are highly educated and employed in elementary occupations or who have no education but work as managers. To overcome this drawback, scholars tend to consider a limited number of circumstances in the definition of types (using either parental education or parental occupation) or to aggregate the different values that a circumstance might take (using blue and white collars rather than more specific occupation categories). These are clearly *ad hoc* solutions, which might greatly affect the shape of the counterfactual distribution and lead to misleading IOp estimates. In what follows, we propose a statistical criterion to properly select among different model specifications or alternative category partitions.

2.2 Variance-bias trade-off in estimating IOp

A number of methodological contributions have shown that if the 'true' set of circumstances is not fully observable, the estimated ex ante IOp will be lower than the 'real' IOp (Roemer, 1998; Ferreira and Gignoux, 2011; Luongo, 2011). This result follows from the assumption of orthogonality between circumstances and effort (see on this Roemer, 1998) and explains why IOp measures are generally interpreted as lower-bound estimates of IOp.

Authors often attempt to solve this problem by using rich datasets that contain the largest possible number of circumstances, including outcome obtained during childhood (Björklund et al., 2012; Hufe et al., 2017). Recently, Niehues and Peichl (2014) endorse an extreme perspective. By exploiting longitudinal datasets, they measure IOp, by including individual fixed effects among circumstances beyond individual control, implying that any unobservable individual characteristic that persists over time is considered to be a source of IOp. This method has been, understandably, proposed as an 'upper-bound' estimate of the true IOp.

However, when using survey data, whenever one attempts to reduce the downward bias by increasing the number of circumstances or the number of categories in which circumstances can be split, she obtains a counterfactual distribution based on a finer partition into types. By construction, this process results in a smaller number of observations in each type[5], which might increase the sampling variance when estimating the counterfactual distribution.

Surprisingly, the empirical literature on IOp estimation has, so far, neglected this second implication. Only recently it has been suggested that this issue may be an important aspect of

[5]Or, if adopting a parametric approach, regression with a larger number of controls and fewer degrees of freedom.

IOp measurement. Brunori et al. (2016) note that the use of very detailed circumstances, such as hundreds of 'villages of birth' in Madagascar or hundreds of 'ethnic groups' in Congo, tends to dramatically increase the IOp estimates.[6]

It is important to note that, when measuring inequality, higher sampling variance of the estimated distribution implies an upward-biased inequality estimate. This result has been proved in Chakravarty and Eichhron (1994) for the case of inequality estimation when the variable of interest is measured with error. The same result can be applied when, instead of the classic measurement error, the partition into types is finer and the type mean is estimated with higher sampling variance. A formal proof based on Chakravarty and Eichhron (1994) is available in Appendix A.[7]

This result has two interesting consequences in the measurement of IOp. First, it states that, if all circumstances are observable and IOp is measured on an appropriate subsample of the original population, such as a typical representative survey, IOp is upward biased. Second, whenever circumstances are not fully observable, two opposite distortions might bias our estimates, and we can no longer claim that the estimated IOp is a lower-bound of the true IOp.

When the sample size is large relative to the number of circumstances included in the model, the downward bias is likely to be large. However, when the sample size is small relative to the number of types/regressors, upward bias might prevail. Appendix B illustrates with a simulation the possible relevance of the upward bias in small samples. However, the absolute and relative sizes of the two biases depend upon a number of factors: the sample size, the joint distribution of outcome and circumstances, and the model specification used to estimate the counterfactual distribution. That is, it is ultimately an empirical issue.

This discussion should clarify that, when estimating IOp, we should consider two different sources of distortion that bias our estimates in opposite directions. The solution to minimize

[6]Note also that the approach proposed by Li Donni et al. (2015), although not explicitly discussed by the authors, represents a possible strategy to address this issue. They define Roemerian types using latent class analysis. That is, they assume that observable circumstances are manifestations of an unobservable membership to a number of latent groups. Their method reduces the number of types and hence avoids large sampling variance in the counterfactual distribution.

[7]Note that Wendelspiess (2015) suggests the opposite direction of bias in a framework in which the outcome is measured with classic measurement error and the sampling variance of the counterfactual distribution is ignored.

the downward distortion cannot consist of *ad hoc* strategies such as simply including a larger possible number of circumstances or considering a broad partition of categories. The choice of the researcher should be based on a statistical criterion. In the following section, we propose a simple method for selecting the best model to measure IOp, a method that exploits the information contained in survey data and minimizes the distortion due to the two biases.

3 Model selection for measuring IOp

We follow Bourgignon et al. (2007) in considering all possible reduced form model but we do not impose a priori restriction on the effect that circumstances might have on outcome. Therefore, in this section, we propose a method to select the most suitable model among all possible specifications.

We consider the following alternatives: (i) a simple linear model with the most parsimonious category partition that provides the lowest extreme IOp estimates (equation 1); (ii) a flexible model that includes the full number of combinations among categories, defined using the finest partition, and leads to the highest value of IOp estimates (equation 2); (iii) all the intermediate specifications that include only subset of category combinations and alternative aggregations of characteristic partitions.

Following a well established approach in the statistical learning literature, we evaluate the variance-bias trade-off in terms of the models' predictive performance. A more flexible model reduces the typical downward bias in IOp measurements and increases the prediction variance causing upward bias. By contrast, a more restricted model reduces the sampling variance and, hence, the upward bias but suffers from omitted variable bias, leading to the typical downward bias, which is well known in the empirical literature. Hence, we exploit the decomposability of the mean squared error (MSE) and select the best model conditioned to available information by means of CV. In a regression setting, the MSE is defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{f}(x_i)\right)^2,$$

where y is the dependent variable, x is the regressors, and $i = 1, ..., n$ are the observations. For given out-of-sample observations y_0 and x_0, the MSE can be decomposed into the variance of $\hat{f}(x_0)$, the square bias of $\hat{f}(x_0)$ and the variance of the error term, ϵ, such as

$$E\left(y_0 - \hat{f}(x_0)\right)^2 = Var(\hat{f}(x_0)) + \left[Bias\left(\hat{f}(x_0)\right)\right]^2 + Var(\epsilon),$$

where $\hat{f}(x_0)$ are the predictions. Since the variance of the prediction error, $Var(\epsilon)$, cannot be reduced, minimizing the MSE implies reducing both bias and variance. Specifically, we aim at minimizing both the bias of the prediction, which accounts for the downward bias of unobserved circumstances, and the variance of the prediction to address the upward bias. The criterion for selecting the best model is comparative and involves two steps: first, we estimate a number of alternatives, such as model (1), model (2) and all the specifications obtained by both interacting only a subset of circumstances and defining categories using different partitions; second, we choose the best specification by means of k-fold CV.[8]

In k-fold CV, the sample is randomly divided into k equal-sized parts. Leaving out part k (test sample), the model is fitted to the other $k - 1$ parts (training sample), and out-of-sample predictions are obtained for the left-out k^{th} part. For each specification, the average of the k MSEs is stored and the best specification is selected by minimizing the average MSE. We select the best specification among various alternatives by means of a simple CV criterion. We claim that estimating $\tilde{Y}$ from the model selected by CV provides the most accurate IOp estimate given the available data. [9]

Note that our strategy might imply the use of a different model for the same country during different time periods and, in general, each time the country's sample differs. As a consequence, when comparing different countries in terms of IOp, we might compare measures obtained with

[8] We are aware that the number of models to test explodes when circumstances are interacted. Moreover, some circumstances can enter into the regression with alternative levels of detail, e.g., country of birth, region of birth, district of birth. When the level of detail to select is not obvious, this further increases the number of models that have to be checked. In these cases, our method should be complemented with an algorithm that can restrict the number of models considered, for example, best subset selection or stepwise selection (Gareth et al., 2013).

[9] Other parsimony criteria could be used to balance variance and bias. However, in contrast to AIC, BIC and adjusted R^2, CV provides a direct estimate of the error based on minimal assumptions. CV is also useful to choose among alternative nonlinear specifications together with non-nested models.

different model specifications. This is in contrast with what generally proposed in the literature. When the same source of data is available for different countries, comparable measures of IOp have usually been computed using the same model specification for all countries, see Marrero and Rodriguez (2012), Brzezinski (2015), Checchi et al. (2016), Suárez and Menéndez (2017). Here, we suggest a different approach: comparable IOp measures should be calculated using the best performing model given the observable circumstances. As a simple example, let us consider the comparison of France and Belgium in terms of IOp. Including 'mother tongue' among circumstances in France would probably make little sense: it would not explain much of the outcome inequality in the country and, therefore, might result in higher sampling variance. However, the same circumstance is likely to be an important source of opportunity inequality in Belgium. Hence, we might infer that 'mother tongue' should be excluded when estimating the French counterfactual distribution but included when estimating Belgian IOp.

We consider our method to be preferable when the intent is to compare the level of IOp in two populations. The derived IOp measures would be the two most reliable estimates of the effect of circumstances on outcome given the information available and the statistical relevance of the characteristics that influence IOp. We believe that the specification used may differ for at least two reasons: first, because the set of available information may not be the same for the two populations; second, and most importantly, because the nature of opportunity inequality, i.e. how circumstances affect individual outcomes, may differ in the analysed populations.

4 An empirical illustration

In this section we provide an empirical illustration based on the EU-SILC 2011 dataset. The EU-SILC is a reference source for comparative statistics on income distribution in the European Union. Because of a special module on the intergenerational transmission of poverty included in a number of EU-SILC waves, the same data have been exploited for other estimates of IOp in the past: Suárez and Menéndez (2017), Marrero and Rodriguez (2012); Brzezinski (2015); Checchi et al. (2016). The year 2011 is the most recent wave, which contains information on family of origin and socioeconomic background. The data refer to 31 countries: Austria (AT), Belgium (BE), Bulgaria (BG), Switzerland (CH), Cyprus (CY), Czech Republic (CZ), Germany (DE), Denmark (DK), Estonia (EE), Greece (EL), Finland (FI), France (FR), Croatia (HR), Hungary (HU), Ireland (IE), Italy (IT), Iceland (IS), Latvia (LV), Lithuania (LT), Luxembourg (LU), Malta (MT), the Netherlands (NL), Norway (NO), Poland (PL), Portugal (PT), Romania (RO), Spain (ES), Slovakia (SK), Slovenia (SI), Sweden (SE), and the United Kingdom (UK). In what follows, we restrict the EU-SILC sample to households whose head is between 26 and 60 years old. The outcome variable is the equivalized disposable income, which is obtained by dividing total household disposable income by the square root of the household size. The circumstances are categorical and identify area of birth and family background (summarized by retrospective questions about parental education and occupation when the respondent was 14 years old). In selecting the best specified model, we consider all possible models ranging from equation (1) to equation (2).

In the first case, we regress outcome on the following four binary variables with no interactions: country of origin (a binary variable that takes the value of one if the respondent was born in the country of residence), father's and mother's occupation (white or blue collar)[10], parental education (low/high).[11]

Those variables are originally coded into a larger number of categories: mother's and father's occupation in 10 categories[12], mother's and father's education in five categories[13] and area of birth in three categories (native, born in Europe, born outside Europe). Interacting all variables coded under the maximum level of detail would result in 7,500 types, a number far greater than the average sample size in EU-SILC. Hence, when estimating equation (2), we

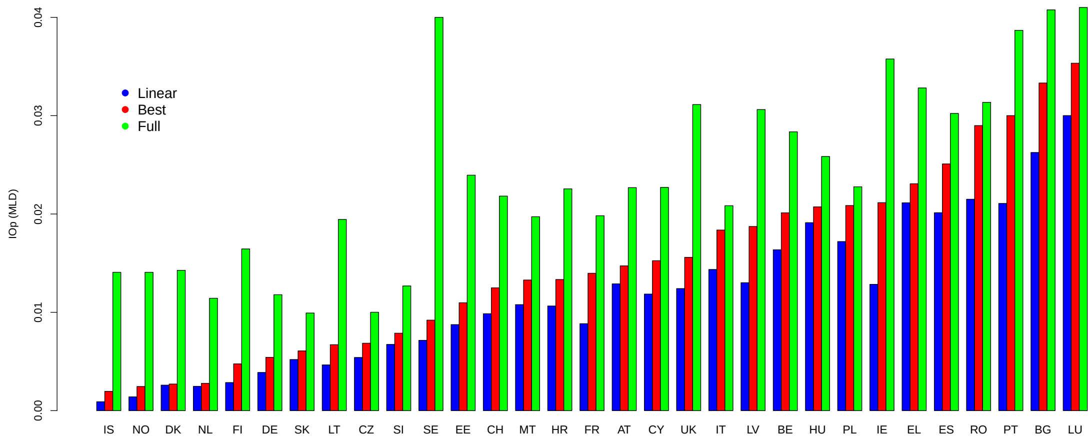
[10]The two categories are based on the International Standard Classification of Occupations, published by the International Labour Office ISCO-08. Blue collar includes parents that who not work or were occupied as: Clerical support workers; Service and sales workers; Skilled agricultural, forestry and fish; Craft and related trades workers; Plant and machine operators; Elementary occupations.

[11]Education categories are based on the International Standard Classification of Education 1997 (ISCED-97). When coded into two categories, low includes ISCED below level 3.

[12]ISCO-08 1-digit: Armed forces occupations; Managers; Professionals; Technicians and associate professionals; Clerical support workers; Service and sales workers; Skilled agricultural, forestry and fish; Craft and related trades workers; Plant and machine operators; Elementary occupations; Did not work/Unknown father/mother

[13]Unknown father/mother, Could neither read nor write; Low level (ISCED 0-2); Medium level (ISCED 3-4); High level (ISCED 5-6).

Figure 1: IOp in 31 European countries under different model specifications



Source: EU-SILC, 2011 . Note: The Figure shows each country's IOp measure obtained with the three alternative methods: (i) the linear, most parsimonious case (*linear*), (ii) the fully interacted model ($full$); (iii) the best model selected ($best$). Countries are ordered according to the IOp level based on the $best$ model specification.

opt for a more compact definition where country of origin is divided in two categories; father's occupation into 10 categories; mother's occupation into two categories; father's education and mother's education into four categories. This model results in 640 possible types. Table 1 shows the descriptive statistics for the data. Intermediate models include subsets of interactions.

Figure 1 shows the level of IOp in the 31 countries. Each bar indicates the mean logarithmic deviation (MLD) of the counterfactual distribution. For each country, the three bars refer to the following cases: (i) the model described in equation (1),(*linear*); (ii) the model described in equation (2), ($full$); (iii) an intermediate measure computed from the best model selected by ten-fold CV ($best$).[14]

Figure 1 shows that the three alternative measures clearly differ among each other, and in some cases (mostly on the left), the best model is very close to the linear model (Denmark and Netherlands, for example). These are mainly Nordic countries characterized by a low level of IOp. Note also that for the same countries, the difference in IOp measured with the *linear* specification and IOp measured with the $full$ model is substantial. This large gap between the

[14]Table 3 in the Appendix contains IOp estimates and relative bootstrapped standard errors based on 500 replications.

two extremes, together with the low level of covariance of circumstances and outcome, is due to the small sample sizes for these countries. When the sample size is limited, such as for Sweden and Iceland, overfitting occurs, even for relatively simple model specifications, and the upward bias discussed above tends to be more pronounced. Interestingly, for Italy, Poland and Hungary, the three countries with the largest sample sizes, the difference between the two models tends to be small. It might be the case that with a sample larger than 12,000, the problem of upward bias becomes less relevant. We further investigate the role of sample size in determining the magnitude of the bias by means of a simulation in Appendix B.

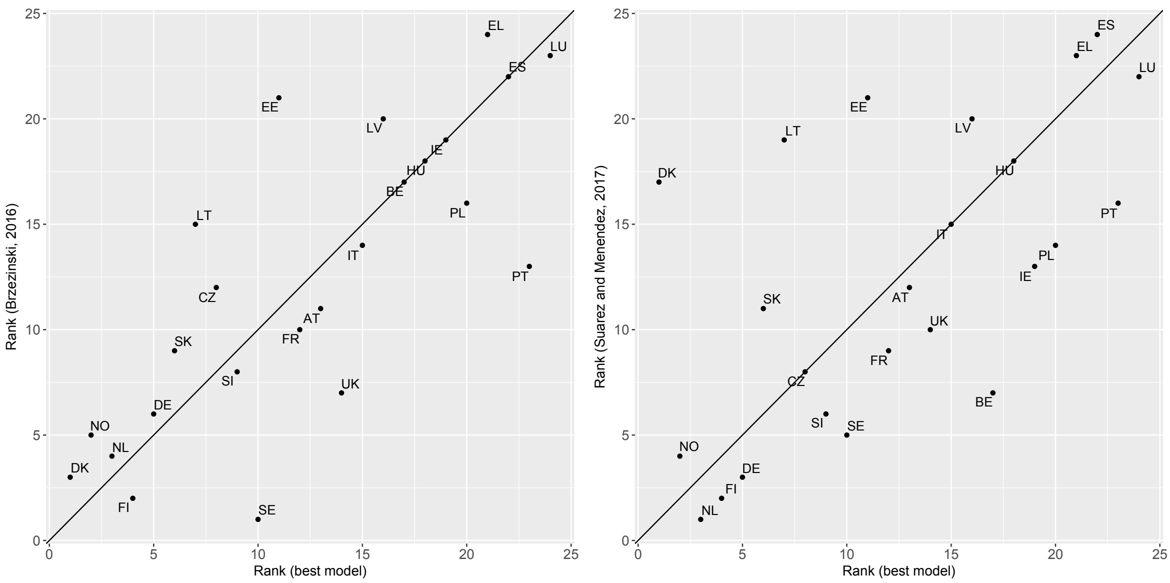
In other cases (concentrated on the right-hand side of the graph), the best model is far from the linear specification and rather close to the most flexible specification. In particular, in Italy, Poland, Romania, Portugal, Bulgaria and Luxembourg, our preferred estimate is closer to the full model than to the linear.

An immediate implication is that the countries' rankings clearly depend on the model specification chosen by the researcher. Consider again Figure 1, where countries are ordered according to the IOp level based on the *best* model specification. The non-monotonicity of the other two series of bins, *linear* and *full*, indicates that the countries' rankings vary with the model specified. For instance, France ranks 19th according to the best model specification but would do much better, ranking 12th, if we consider the most parsimonious specification.

To further investigate the problem of IOp sensitivity to alternative econometric specifications, we consider the measures of IOp proposed in two recent papers (Brzezinski (2015) and Suárez and Menéndez (2017). Both analyses use the same wave of EU-SILC, follow the *ex ante* approach and use equivalized disposable income as outcome variable. The measures provided by those authors differ because they use different model specifications. Suárez and Menéndez (2017) estimate IOp parametrically considering the following circumstances: gender, nationality, urban density, parental education, and parental occupation. Brzezinski (2015) adopt a parametric approach that includes parental education, parental occupation, and nationality. Both models are estimated using log-linear OLS regression with no interactions.

Figure 2 shows the rank correlation of our *best* measure and the two alternative estimates for

Figure 2: IOp estimates in 24 European countries from different studies



Source: EU-SILC, 2011 . Note: the Figure shows the rank correlation of countries in terms of IOp. Our best model specification is compared with Suárez and Menéndez (2017) and Brzezinski (2015)

the 24 countries considered in both studies. We note that the final assessment differs substantially in both cases. Although the rank-correlation is clearly positive and significant, a number of countries lie outside the 45 degree line. Indeed, the re-ranking is substantial in a few cases. For example, in Suárez and Menéndez (2017), Ireland ranks 17th and Belgium ranks 7th, whereas with our *best* measure, they rank first and 17th, respectively. Additionally, in Brzezinski (2015), Portugal ranks 13th, whereas if our *best* specification is adopted, it ranks 23rd.[15]

We believe that this exercise provides convincing evidence that the variance-bias trade-off in IOp measurement is far from negligible in empirical applications. Therefore, it is crucial to introduce a statistical criterion to select the best model among a very large number of possible specifications.

[15]Figure 4 in Appendix C shows a closer but far from perfect rank correlation between the estimates of Brzezinski (2015) and Suárez and Menéndez (2017).

5 Conclusions

The past two decades have seen growing interest from scholars and policy makers in the measurement of inequality of opportunity. A number of methodological contributions have shown that estimates of inequality of opportunity are mostly downward biased. This is a consequence of the partial observability of circumstances beyond individual control that affect individual outcome. This issue has typically been addressed by resorting to rich datasets and adopting broad econometric specifications. However, since IOp is measured as inequality in a counterfactual sample distribution, a second possible source of bias might be related to the sampling variance of the estimated counterfactual distribution. In this paper, we discuss this additional source of bias, which has surprisingly been neglected by the empirical literature on IOp measurement. We show that it implies an upward bias of IOp, which challenges the interpretation of IOp estimates as lower-bound estimates of the real IOp.

We stress that since the empirical specification used to estimate IOp largely influences its magnitude, a reasonable statistical criterion to select among alternative models is required. We suggest that this criterion should minimize the two sources of bias.

We interpret this problem as a typical variance-bias trade-off and propose a simple CV method to find the best-fitting model. Cross-validation methods assess the predictive performance of alternative models to estimate a dependent variable out of sample. Overfitted models tend to be extremely accurate in explaining variability in sample but perform poorly in predicting on a test sample not used to estimate the model. By providing an unbiased assessment of the relative predictive performance of each possible model specification, CV can be used by researchers as a guide to choose the best model to estimate IOp.

The models selected by the algorithm typically differ across countries in terms of the variables considered and the interactions included, suggesting that when attempting to produce comparable IOp estimates, scholars may abandon the idea of specifying the same model for all countries in all time periods. By contrast, comparable estimates may be obtained by using the model specification that best captures the correlation of individual outcome and circumstances beyond individual control separately for each country and time period.

Finally, we show the empirical relevance of our intuition and implement the proposed method to measure IOp in 31 European countries. Our empirical evidence illustrates that the choice of model specification strongly affects the estimated IOp and demonstrates the importance of having a widely accepted criterion to identify the best possible specification.

Table 1: Descriptive statistics

						Mother			Father		
Country	Sample size	Eq. income	Tot. Ineq. (MLD)	Age	Native	White collar	Elementary edu.	Secondary edu.	White collar	Elementary edu.	Secondary edu.
AT	6,630	17,024.37	0.167	43.96	0.788	0.283	0.563	0.350	0.363	0.379	0.437
BE	5,122	14,581.69	0.150	43.14	0.825	0.260	0.524	0.222	0.454	0.463	0.211
BG	6,651	2,114.25	0.246	44.25	0.996	0.421	0.452	0.382	0.236	0.448	0.361
CH	7,212	28,180.78	0.186	44.58	0.683	0.359	0.398	0.413	0.486	0.226	0.490
CY	5,135	11,758.22	0.172	43.48	0.787	0.198	0.654	0.191	0.310	0.638	0.201
CZ	7,068	5,576.24	0.133	44.26	0.963	0.523	0.650	0.277	0.322	0.597	0.219
DE	11,473	15,797.95	0.173	45.27	0.881	0.371	0.274	0.511	0.439	0.120	0.529
DK	2,233	22,189.83	0.229	45.76	0.928	0.570	0.512	0.290	0.461	0.341	0.429
EE	5,464	3,783.51	0.234	43.95	0.869	0.554	0.303	0.409	0.268	0.279	0.366
EL	6,430	7,416.41	0.320	43.65	0.892	0.162	0.585	0.155	0.310	0.579	0.156
ES	16,188	10,231.74	0.303	43.65	0.835	0.137	0.801	0.057	0.342	0.761	0.071
FI	3,439	18,087.89	0.137	43.90	0.949	0.525	0.458	0.266	0.313	0.413	0.210
FR	11,286	15,448.28	0.174	44.15	0.890	0.336	0.721	0.089	0.397	0.700	0.083
HR	6,720	3,557.79	0.234	45.36	0.884	0.234	0.648	0.225	0.283	0.474	0.360
HU	13,533	3,240.24	0.167	44.43	0.989	0.379	0.636	0.262	0.241	0.595	0.263
IE	3,360	15,220.27	0.188	43.08	0.780	0.212	0.524	0.338	0.360	0.562	0.268
IS	1,619	12,815.69	0.104	43.20	0.907	0.494	0.609	0.281	0.449	0.318	0.499
IT	22,696	12,036.84	0.299	43.88	0.875	0.165	0.762	0.125	0.315	0.693	0.147
LT	4,935	3,033.07	0.290	46.36	0.945	0.380	0.493	0.350	0.220	0.550	0.269
LU	7,244	22,727.56	0.170	43.18	0.484	0.233	0.582	0.256	0.374	0.475	0.327
LV	6,748	3,081.47	0.288	44.19	0.880	0.486	0.382	0.429	0.201	0.359	0.324
MT	4,467	7,327.47	0.158	44.23	0.950	0.064	0.662	0.154	0.436	0.571	0.189
NL	5,884	17,126.15	0.125	44.32	0.880	0.259	0.503	0.301	0.517	0.355	0.294
NO	2,568	27,770.11	0.107	43.81	0.915	0.562	0.343	0.440	0.507	0.309	0.391
PL	14,595	3,363.28	0.221	44.14	0.999	0.313	0.480	0.448	0.221	0.426	0.483
PT	6,355	6,235.56	0.222	44.42	0.907	0.177	0.646	0.034	0.268	0.712	0.037
RO	6,549	1,504.92	0.264	43.82	0.998	0.180	0.760	0.141	0.147	0.760	0.103
SE	543	16,165.57	0.127	42.79	0.863	0.593	0.473	0.235	0.436	0.491	0.183
SI	5,234	7,883.03	0.102	43.25	0.872	0.347	0.729	0.166	0.287	0.676	0.175
SK	7,241	4,139.78	0.151	43.36	0.989	0.490	0.418	0.516	0.287	0.334	0.526
UK	6,329	14,177.45	0.251	44.05	0.851	0.443	0.667	0.111	0.462	0.507	0.238

Source: EU-SILC (2011). Descriptive statistics refer only to respondents with no missing information. White collar occupations include: Managers; Professionals; Technicians and associate professionals (ISCO-08).

References

- Arlot, S. and A. Celisse (2010) “A survey of cross-validation procedures for model selection.” Statistics Surveys. 4 (2010), 40–79.

- Balcazar, C. (2015), “Lower bounds on inequality of opportunity and measurement error”, *Economics Letters*, 137: 102-105.

- Björklund, A., Jäntti A. and Roemer J. (2012), “Equality of Opportunity and the Distribution of Long-Run Income in Sweden”, *Social Choice and Welfare*, 39: 675-696.

- Bourguignon, Ferreira F., and Menendez M., (2007), “Inequality of Opportunity in Brazil”, *Review of Income Wealth*, 53: 585-618.

- Bourguignon, Ferreira F., and Menendez M., (2013), “Inequality of Opportunity in Brazil: A Corrigendum”, *Review of Income Wealth*, 59: 551-555.

- Brzenziński M. 2015. ‘Inequality of opportunity in Europe before and after the Great Recession’, Working Paper n. 2/2015 (150). Faculty of Economic Sciences, University of Warsaw.

- Brunori P., Palmisano, F., Peragine, V. (2016), “Inequality of opportunity in Sub Saharan Africa”, *World Bank Policy Research Working Papers* 7782.

- Brunori, P., Ferreira F., Peragine, V. (2013), “Inequality of opportunity, income inequality and mobility: some international comparisons”, in *Getting Development Right: Structural Transformation, Inclusion and Sustainability in the Post-Crisis Era* edited by E. Paus. Palgrave Macmillan.

- Chakravarty S.R., Eichhorn W. (1994) “Measurement of Income Inequality : Observed Versus True Data.” In: Eichhorn W. (eds) *Models and Measurement of Welfare and Inequality.* Springer, Berlin, Heidelberg.

- Checchi, D., and Peragine V. (2010), “Inequality of Opportunity in Italy”, *Journal of Economic Inequality*, 8: 429-450.

- Checchi, D., Peragine V., Serlenga L., (2016), “Inequality of Opportunity in Europe: Is There a Role for Institutions?”, in Cappellari L., Polachek S., Tatsiramos K. (ed.) *Inequality: Causes and Consequences* , Research in Labor Economics, Volume 43, Emerald.

- Ferreira, F., and Gignoux, J. (2011), “The measurement of inequality of opportunity: theory and an application to Latin America”, *Review of Income and Wealth*, 57: 622-657.

- Ferreira F., and V. Peragine. (2016). “Equality of Opportunity: Theory and Evidence”, in *Oxford Handbook of Well-Being and Public Policy* edited by M. Adler and M. Fleurbaey. Oxford University Press.

- Gareth J., Witten D., Hastie T., Tibshirani R. (2013). “An Introduction to Statistical Learning with Applications in R”, Springer, New York.

- Hufe P. and A. Peichl (2015). “Lower Bounds and the Linearity Assumption in Parametric Estimations of Inequality of Opportunity”, *IZA Working Papers*, DP No. 9605.

- Fleurbaey, M., and E. Schokkaert, (2009). “Unfair inequalities in health and health care.” *Journal of Health Economics*, 28, 73-90.

- Fleurbaey M., and V. Peragine, (2013). “Ex Ante Versus Ex Post Equality of Opportunity,” Economica, vol. 80(317), 118-130.

- Kanbur, R. and A. Wagstaff, (2016), “How useful is inequality of opportunity as a policy construct?”, In: Basu K., Stiglitz J.E. (eds) *Inequality and Growth: Patterns and Policy.* International Economic Association Series. Palgrave Macmillan, London

- Ibarra L., Martinez C., Adan L. (2015), “Exploring the sources of downward bias in measuring inequality of opportunity.”, *World Bank Policy Research working paper* no. WPS 7458. Washington.

- Lefranc, A., Pistolesi, N., Trannoy A., (2009) “Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France”, *Journal of Public Economics*, Volume 93, Issues 11-12,1189-1207.

- Li Donni P., Rodriguez J. G., Rosa Dias P., 2015. Empirical definition of social types in the analysis of inequality of opportunity: a latent classes approach. *Social Choice and Welfare* 44, 673-701.

- Luongo P. (2011),“The Implication of Partial Observability of Circumstances on the Measurement of Inequality of Opportunity”, in Rodriguez, J. eds. , *Research on Economic Inequality*, 19: 23-49.

- Marrero, G., and J. Rodrguez. (2011), “Inequality of opportunity in the United States: trends and decomposition.” *Research on Economic Inequality*, 217-246.

- Marrero, G., and J. Rodrguez. (2012). “Inequality of Opportunity in Europe”. *Review of Income and Wealth*, vol. 58: 597-621.

- Niehues, J., and A. Peichl (2014), “Upper bounds of inequality of opportunity: theory and evidence for Germany and the US” , *Social Choice & Welfare*, 43: 63-79.

- Hufe P., Peichl A., Roemer J., Ungerer M. (2017), “Inequality of Income Acquisition: The Role of Childhood Circumstances”, *Social Choice and Welfare*, First Online: 29 March 2017.

- Roemer, J. (1998), “Equality of Opportunity”, Cambridge, MA: Harvard University Press.

- Roemer J., A . and Trannoy , (2015), “Equality of Opportunity”, In *Handbook of Income Distribution* Edited by Anthony B. Atkinson and Francois Bourguignon. vol. 2. Elsevier.

- Suárez A .A. and Menéndez A. J. L. “Income inequality and inequality of opportunity in Europe. Are they on the rise?”, ECINEQ WP 2017 - 436.

- Van de gaer, D. and Ramos, X. (2016). Empirical Approaches to Inequality of Opportunity: Principles, Measures, and Evidence. Journal of Economic Surveys, 30(5): 855-883.

- Varian, H. R. (2014). “Big Data: New Tricks for Econometrics”. *Journal of Economic Perspectives*, 28(2): 3-27.

- Wendelspiess F. C. J. (2015). “Measuring Inequality of Opportunity with Latent Variables”. *Journal of Human Development and Capabilities*, 16:1, 106-121.

A Upward bias when estimating IOp with survey data

Chakravarty and Eichhron (1994) distinguish between the true distribution of income, y, and the observed distribution, $\tilde{y}$, where $\tilde{y} = y + e$ and e is commonly defined as the measurement error such that $e \sim iid(0, \sigma^2)$. By considering a strictly concave von Neumann-Morgenstern utility function, U, they prove by analogy that, if we measure inequality $I(\tilde{y})$ with an inequality index I that satisfies symmetry and the Pigou-Dalton transfer principle, then the inequality of the true y distribution is smaller than observed.

Without loss of generality, we apply their result to the case of non-parametric IOp measurement (eq. 2).

Proposition Let Y be the population income distribution, and $\tilde{Y}$ be the counterfactual distribution estimated with the model in equation (2). Assume that $\tilde{Y}$ is estimated by observing the full set of circumstances and the entire population. Let X be a proper subsample of the entire population and let $\tilde{X}$ be the counterfactual distribution estimated with equation (2) on X. Let $I()$ be any inequality measure that satisfies symmetry and the Pigou-Dalton trnasfer principle. Let $IOp = I(\tilde{Y})$ be a measure of inequality of opportunity in the population and let $E(I\hat{O}p) = E(I(\tilde{X}))$ be the expected value of $I()$ estimated on $\tilde{X}$. Then, $E(I\hat{O}p) > IOp$.

Proof

Let $\mathbf{M} = \mu_1, ..., \mu_T$ be the vector of types' mean outcomes in the population. Let $\hat{\mathbf{M}} = \hat{\mu_1}, ..., \hat{\mu_T}$ be the estimates of types' means based on X. Then, for each $t = 1, ..., n$, $\hat{\mu}_t = \mu_t + \eta$, where $\eta = \frac{\sigma}{\sqrt{N_t}} \sim (0, \chi^2)$ is the standard error of $\hat{\mu}_t$.

Following Chakravarty and Eichhron (1994), we assume that U is a strictly concave function. By Jensen's inequality, we have:

$$E\left(U\left(\hat{\mathbf{M}}|\mathbf{M}\right)\right) < U\left(E\left(\hat{\mathbf{M}}|\mathbf{M}\right)\right). \tag{3}$$

Note that $E\left(\hat{\mathbf{M}}|\mathbf{M}\right) = \mathbf{M}$, therefore:

$$E\left(U\left(\hat{\mathbf{M}}|\mathbf{M}\right)\right) < U\left(\mathbf{M}\right). \tag{4}$$

By taking expectations with respect to **M** on both sides, (4) becomes:

$$E\left(U\left(\hat{\mathbf{M}}\right)\right) < U\left(E(\mathbf{M})\right). \tag{5}$$

Because $E(\eta) = 0$, the two distributions have the same mean. If U is a strictly concave function, then (5) is equivalent to saying that the distribution of **M** Lorenz dominates the distribution of $\hat{\mathbf{M}}$, which implies that $E(I\hat{O}p) > IOp$.

■

Corollary: When one or more of the relevant circumstances is not used to obtain the counterfctual distribution (partial observability of circumstances) and the counterfactual distribution is estimated on a proper subsample of the population, estimated IOp cannot be interpreted as a lower bound of the IOp in the population.

B A simulation to assess the magnitude of the upward bias

The reader may wonder whether the upward bias discussed in this paper actually represents a non-negligible issue in empirical implementations. To provide an idea of the possible magnitude of the bias, we perform a simulation. When estimating inequality of opportunity, the data generating process is typically unknown. We therefore prefer to base the simulation on the entire EU-SILC dataset instead of creating an *ad hoc* dataset.

Assume that the entire EU-SILC dataset is our population of interest. A population composed of 202,843 individuals aged between 26 and 60 years (more than the same age population in Iceland and approximately the same population in Luxembourg). Additionally, assume that a few observable circumstances are the only circumstances that determine inequality of opportunity. Individual outcome is assumed to be the result of the interactions of three circumstances: parental education, parental occupation, and origin. Individuals in the same type share the same highest parental education (five categories), same immigration history (a dummy that takes the value of one if the respondent is a first- or second-generation immigrant), and the same highest parental occupation (ISCO 1 digit).

Under our assumptions, we can observe the real partition of the population into types. The observed between-type inequality is then the real IOp in the population. The residual inequality is assumed to be due to effort. Measured by MLD, IOp in the entire sample is 0.0314, approximately 7% of the total variability.

Our aim is then to understand the circumstances under which an estimate of inequality of opportunity based on a random subsample of this population results in upward bias. To this end, we estimate IOp using samples increasing in size. We start with 500, which is approximately the sample size of the smallest country in EU-SILC (Sweden). We then add 500 observations in each step until we reach a sample of 20,000 observations (not far from Italy's sample size, the largest country in EU-SILC). Each sample is randomly drawn 500 times to obtain 95th percentile bootstrap confidence intervals around the point estimate.

Figure 3 shows the IOp estimates for samples of increasing size. In grey, we provide a histogram showing the frequency of countries' sample size (reported on the right y-axis) in

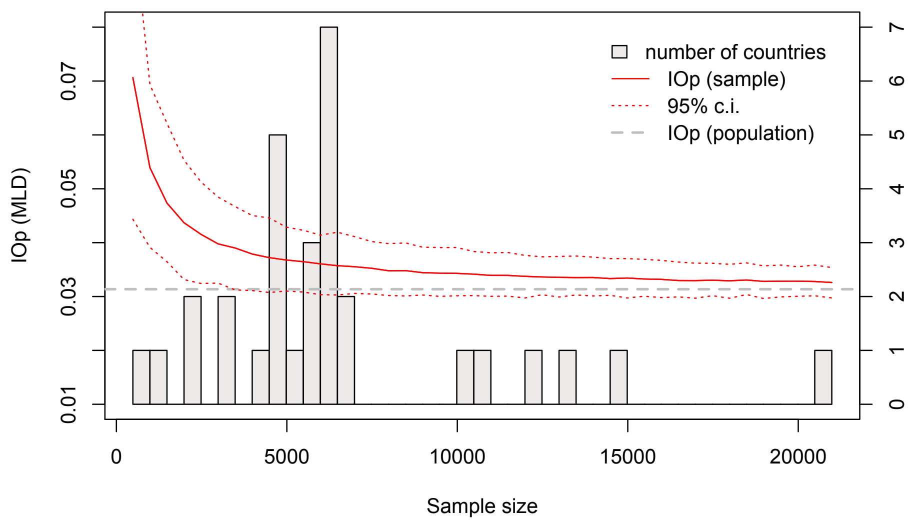
EU-SILC 2011.[16]

The estimates show a marked upward bias for the smallest samples. The average IOp based on the samples is more than 1.2 times higher than the IOp in the population for samples smaller than 4,000. These are not unrealistically small samples: six of the 31 countries have smaller sample sizes. Interestingly, the confidence intervals of the estimates do not contain the population's estimate for all samples smaller than 3,000 (Sweden, Iceland, Denmark, and Norway have smaller sample sizes). Moreover, the upward bias is less than 10% only for sample sizes larger than 9,000. Only France, Germany, Hungary, Poland, Spain, and Italy have larger sample sizes.

Estimates based on the samples approach the IOp in the population rather slowly; at the extreme right of the graph, the bias is approximately 4%. This may be considered a negligible distortion. Interestingly, the reader may recall that in Figure 1 of Section 4, we found a relatively small difference between the IOp estimated with the two extreme specifications for countries with sample sizes larger than 10,000. However, in our simulation, a sample size of 20,000 observations is extremely large as it represents slightly less than 10% of the population.

[16]Note that these are the sample sizes used in the regression; they include only individuals with non-missing information.

Figure 3: IOp estimated on samples of increasing size



Source: EU-SILC, 2011

C Additional tables and figures

Table 2: IOP (MLD) Estimates of 31 countries

Country	Best	Best low	Best high	Linear	Linear low	Linear high	Full	Full low	Full high
AT	0.0147	0.0133	0.0179	0.0129	0.0107	0.0161	0.0227	0.0196	0.0258
BE	0.0201	0.0182	0.0260	0.0164	0.0140	0.0190	0.0284	0.0241	0.0326
BG	0.0333	0.0301	0.0379	0.0263	0.0221	0.0309	0.0408	0.0361	0.0454
CH	0.0125	0.0115	0.0156	0.0099	0.0082	0.0118	0.0218	0.0194	0.0243
CY	0.0152	0.0126	0.0183	0.0119	0.0080	0.0145	0.0227	0.0196	0.0258
CZ	0.0069	0.0058	0.0084	0.0054	0.0042	0.0066	0.0100	0.0083	0.0117
DE	0.0054	0.0050	0.0068	0.0039	0.0031	0.0048	0.0118	0.0101	0.0135
DK	0.0027	0.0018	0.0047	0.0026	0.0016	0.0050	0.0143	0.0104	0.0181
EE	0.0110	0.0088	0.0140	0.0087	0.0074	0.0109	0.0239	0.0200	0.0279
EL	0.0231	0.0189	0.0281	0.0211	0.0168	0.0250	0.0328	0.0273	0.0384
ES	0.0251	0.0239	0.0279	0.0201	0.0181	0.0230	0.0302	0.0277	0.0328
FI	0.0048	0.0044	0.0080	0.0029	0.0020	0.0049	0.0164	0.0122	0.0207
FR	0.0140	0.0121	0.0167	0.0088	0.0076	0.0112	0.0198	0.0172	0.0224
HR	0.0133	0.0113	0.0165	0.0106	0.0082	0.0130	0.0226	0.0188	0.0263
HU	0.0207	0.0189	0.0231	0.0191	0.0170	0.0209	0.0258	0.0242	0.0275
IE	0.0212	0.0195	0.0290	0.0128	0.0098	0.0170	0.0358	0.0309	0.0406
IS	0.0020	0.0015	0.0050	0.0009	0.0003	0.0023	0.0141	0.0104	0.0177
IT	0.0184	0.0163	0.0201	0.0144	0.0129	0.0167	0.0208	0.0192	0.0225
LT	0.0067	0.0053	0.0089	0.0047	0.0027	0.0068	0.0194	0.0159	0.0229
LU	0.0353	0.0320	0.0405	0.0300	0.0273	0.0358	0.0410	0.0376	0.0444
LV	0.0187	0.0165	0.0238	0.0130	0.0110	0.0170	0.0306	0.0267	0.0345
MT	0.0133	0.0118	0.0165	0.0108	0.0086	0.0135	0.0197	0.0160	0.0234
NL	0.0028	0.0021	0.0039	0.0025	0.0017	0.0037	0.0114	0.0095	0.0133
NO	0.0025	0.0022	0.0063	0.0014	0.0009	0.0029	0.0141	0.0110	0.0172
PL	0.0209	0.0190	0.0239	0.0172	0.0148	0.0187	0.0228	0.0201	0.0254
PT	0.0300	0.0267	0.0362	0.0211	0.0167	0.0248	0.0387	0.0335	0.0439
RO	0.0290	0.0255	0.0333	0.0215	0.0176	0.0243	0.0314	0.0274	0.0353
SE	0.0092	0.0066	0.0283	0.0072	0.0028	0.0135	0.0453	0.0320	0.0587
SI	0.0079	0.0061	0.0100	0.0067	0.0055	0.0088	0.0127	0.0108	0.0146
SK	0.0061	0.0053	0.0077	0.0052	0.0044	0.0066	0.0099	0.0083	0.0115
UK	0.0156	0.0132	0.0195	0.0124	0.0095	0.0155	0.0311	0.0280	0.0342

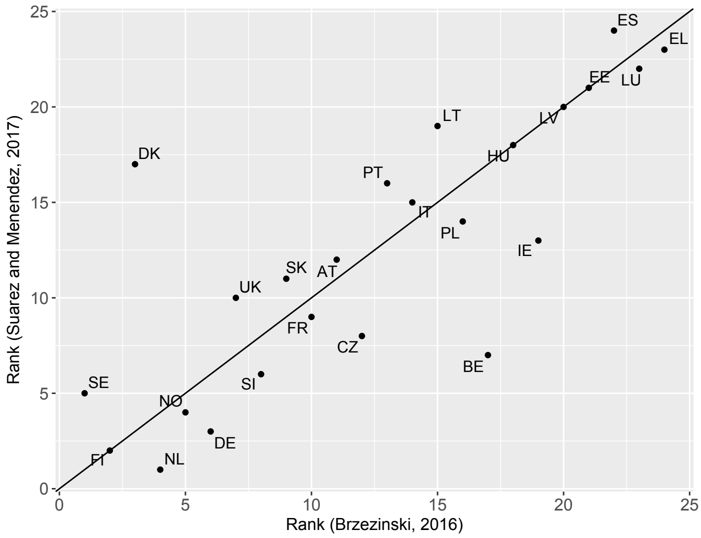
Notes: Source: EU-SILC 2011. Notes: IOp (MLD) estimates derived from (i) the linear most parsimonious case (linear); (ii) the fully interacted model (full); (iii) the best model selected (best). Bootstrapped confidence intervals.

Table 3: Model specifications

Country	Regressors included in the 'best' model specification							
AT	father occ. (10)	mother occ. (10)	mother edu. (5)	(birth area) x (father white)	(father white) x (highest par. edu.)	(mother white) x (highest par. edu.)		
BE	father occ. (10)	mother occ. (10)	mother edu. (5)	(birth area) x (highest par. edu)	(father white) x (mother white)	(mother white) x (highest par. edu.)		
BG	birth area (3)	father occ. (10)	father edu. (5)	mother edu. (5)	(father white) x (mother white)	(father white) x (highest par. edu.)		
CH	father white	father edu. (5)	mother edu. (5)	(birth area) x (mother white)	(birth area) x (highest par. edu)	(mother white) x (highest par. edu.)		
CY	father occ. (10)	father edu. (5)	mother edu. (5)	(birth area) x (mother white)				
CZ	father occ. (10)	mother occ. (10)	father edu. (5)	(birth area) x (father white)	(birth area) x (mother white)			
DE	father edu. (5)	mother edu. (5)	(birth area) x (mother white)	(father white) x (highest par. edu.)				
DK	birth area (3)	father white	mother white	mother edu. (5)				
EE	mother white	father edu. (5)	mother edu. (5)	(birth area) x (father white)				
EL	mother occ. (10)	(birth area) x (mother white)	(birth area) x (highest par. edu)	(father white) x (highest par. edu.)				
ES	father occ. (10)	mother occ. (10)	father edu. (5)	mother edu. (5)	(birth area) x (mother white)	(father white) x (highest par. edu.)		
FI	mother occ. (10)	father white	mother edu. (5)	(birth area) x (mother white)	(birth area) x (highest par. edu)			
FR	father occ. (10)	mother white	father edu. (5)	mother edu. (5)	(birth area) x (father white)			
HR	birth area (3)	father occ. (10)	mother white	father edu. (5)	mother edu. (5)			
HU	birth area (3)	mother white	father edu. (5)	mother edu. (5)	(father white) x (highest par. edu.)			
IE	birth area (3)	father occ. (10)	mother occ. (10)	father edu. (5)	mother edu. (5)	(father white) x (mother white)	(father white) x (highest par. edu.)	(mother white) x (highest par. edu.)
IS	mother occ. (10)	(birth area) x (father white)	(birth area) x (mother white)	(father white) x (mother white)				
IT	father occ. (10)	father edu. (5)	mother edu. (5)	(birth area) x (mother white)	(birth area) x (highest par. edu)	(father white) x (mother white)	(mother white) x (highest par. edu.)	
LT	father white	father edu. (5)	mother edu. (5)	(birth area) x (mother white)				
LU	father white	father edu. (5)	mother edu. (5)	(birth area) x (highest par. edu)	(mother white) x (highest par. edu.)			
LV	father occ. (10)	mother occ. (10)	father edu. (5)	mother edu. (5)	(birth area) x (highest par. edu)			
MT	father occ. (10)	father edu. (5)	mother edu. (5)	(birth area) x (father white)	(father white) x (mother white)			
NL	mother edu. (5)	(birth area) x (father white)	(birth area) x (mother white)	(father white) x (mother white)				
NO	father occ. (10)	mother edu. (5)	(birth area) x (father white)	(birth area) x (highest par. edu)	(father white) x (highest par. edu.)			
PL	mother occ. (10)	father edu. (5)	(birth area) x (mother white)	(birth area) x (highest par. edu)	(father white) x (mother white)			
PT	mother occ. (10)	father edu. (5)	mother edu. (5)	(birth area) x (father white)	(birth area) x (highest par. edu)	(father white) x (highest par. edu.)	(mother white) x (highest par. edu.)	
RO	mother occ. (10)	mother edu. (5)	(birth area) x (mother white)	(father white) x (mother white)	(father white) x (highest par. edu.)	(mother white) x (highest par. edu.)		
SE	birth area (3)	mother edu. (5)	(father white) x (mother white)	(father white) x (highest par. edu.)				
SI	birth area (3)	mother occ. (10)	father edu. (5)	mother edu. (5)	(father white) x (mother white)			
SK	father occ. (10)	mother occ. (10)	highest par. edu. (5)	(birth area) x (father white)	(birth area) x (mother white)	(father white) x (mother white)		
UK	mother edu. (5)	(birth area) x (father white)	(father white) x (mother white)	(father white) x (highest par. edu.)	(mother white) x (highest par. edu.)			

Notes: Source: EU-SILC 2011. Numbers in parentheses refer to the number of categories. Complete regression tables are available upon request.

Figure 4: IOp estimates for 24 European countries from different studies



Source: EU-SILC, 2011 . Note: the figure shows the rank correlation of countries in terms of IOp. The ranking proposed by Suárez and Menéndez (2017) is compared with the ranking proposed by Brzezinski (2015)