Households or Individuals? Gender and the Level of Analysis when Measuring Opportunity Inequality

Hugo del Valle-Inclán*

This version: February 2021

When estimating inequality of opportunity, the level of analysis implies a normative choice regarding gender. This, however, has been largely overlooked in the literature. We first analyze theoretically what are the implications of the income aggregates used in empirical studies as the advantage variable, some aggregates at the household level and others focusing on individuals. Next, using data from 31 European countries we show that inequality of opportunity estimates are highly sensitive to the level of analysis, and that this variation responds almost entirely to the contribution of the circumstance gender to overall inequality of opportunity. We conclude that if gender is to be considered a source of unfair inequality, then an aggregate at the individual level must be employed. (120 words)

JEL Codes: D63, D70, I24

Keywords: inequality of opportunity, income aggregate, gender

*University of Vigo. Email address: hvalle-inclan@uvigo.es. The research leading to these results has received funding from the *Xunta de Galicia*, programs ECOBAS (ED431E 2018/02) and RGEAF (ED431B 2019/35). The responsibility for all conclusions drawn from the data lies entirely with the author.

1. Introduction

> If women have a comparative advantage over men in the household sector when they make the same investments in human capital, an efficient household with both sexes would allocate the time of women mainly to the household sector and the time of men mainly to the market sector.
>
> Gary Becker, *A Treatise on the Family*, p. 38 (1991 [1981])

Standard analysis of economic inequality studies the distribution of income, wealth or consumption at the household level. This level of analysis is often justified on the basis that the object of study is the access of individuals to economic resources (Jenkins and Van Kerm 2009), which may be better captured by focusing on households rather than on individuals themselves[1].

However, when measuring inequality of opportunity (IOP henceforth) the appropriate level of analysis remains unclear. If income is the outcome of interest, both household and individual levels are commonly considered in the literature, but the choice is rarely justified and seldom are robustness issues discussed.

In this article we first show that IOP estimates are highly sensitive to this choice, and secondly argue that if one believes gender to be a potential source of IOP, then income at the individual level must be considered. Specifically, it is so because assuming within-household redistribution virtually nullifies the contribution of the circumstance gender to overall IOP.

Nonetheless, why are different levels of analysis common in the literature? One explanation is that both levels have their own advantages and shortcomings, so different researchers might be inclined to follow different approaches. Yet, the lack of a common and widely accepted methodology compromises comparability, making IOP measurement inconsistent. In this article we explore the implications of the level of analysis, and argue in favor of focusing on individuals.

On the one hand, the IOP approach is not utilitarian, i.e., the interest is not in the access to economic resources. Instead, the object of study is how certain personal traits or decisions condition certain outcomes, so a priori it could be argued that the focus ought to be on the individual level—furthermore, it makes little sense to assume within-household redistribution of opportunities. Indeed the original proposal of the IOP approach referred to individual outcomes (Roemer 1993, 1998), and many empirical studies consider individual income as the outcome

[1]As long as we assume (perfect) within-household redistribution, even though it rarely holds. The problems associated with this assumption are well known in inequality research—see for instance Haddad and Kanbur (1990) for an early reference. Recent contributions include Lechene et al. (2019), who show that poor people may live in non-poor households, Fremeaux and Leturcq (2020), who find stark differences between household and personal wealth inequality, and Sauer et al. (2020), who offer a general perspective.

of interest, examples of which include Andreoli and Fusco (2017), Björklund et al. (2012), and Fleurbaey, Peragine, and Ramos (2017).

On the other hand, focusing on the individual level neglects household bargaining processes that affect labor market participation. This refers to the fact that, for instance, an individual with a comparative advantage over her spouse in the household sector might rationally decide to stop working. In such situation it is obvious that not all income differences between this inactive person and a full-time worker could be attributed to IOP. In fact, partly because of this problem it is also common to estimate IOP focusing on the household level.[2] Examples of studies considering household income are Brunori, Palmisano, et al. (2019), Ferreira, Lakner, et al. (2018), and Singh (2012).

Focusing on the household level effectively abstracts from the problem of labor market participation, but that is not fully satisfactory either. To begin with, part of the effect of circumstances works precisely through participation in the labor market—specially that of gender—, so ideally we would account for this issue rather than abstracting from it. In addition, relating personal circumstances to income generated by groups of potentially heterogeneous individuals (namely households) may bias the analysis. Hence, both the household and individual levels can be problematic.

In response to these shortcomings other approaches have been proposed. A third solution could be seen as something in between the first two. It consists of referring to household income, but keeping only household heads in the sample. An argument in favor of this choice is that it may alleviate the problems the two previous approaches have, while keeping their benefits. Articles following this methodology include Ferreira and Gignoux (2011), Marrero and Rodríguez (2013), and Palomino et al. (2019).[3]

A fourth solution builds onto the first one, and consists of explicitly modeling selection into employment. This is done to account for the household bargaining processes affecting labor market participation that were referred above.[4] To the best of our knowledge, only Bourguignon et al. (2003)[5] and Checchi, Peragine, and Serlenga (2016) have followed this approach.

The third and fourth solutions are also at the household and individual level, respectively, but their particular characteristics may alleviate the issues suffered by the first two. Throughout this text we will refer to these four possible definitions of the outcome of interest as income aggregates, and investigate to what extent choosing one or the other matters for the measurement of IOP. This is, we will check if IOP estimates are sensitive to considering either a) individual income, b) household income, c) household income but keeping only heads in the

[2]Other reasons may include that individual income is specially prone to suffer short-term shocks, inertia from the economic inequality analysis, or taking an (inadequate) utilitarian perspective of IOP.

[3]In actuality Ferreira and Gignoux (2011) keep household heads and also spouses, and justify their choice on the basis that in some countries of their sample family background information was only collected for these individuals. Nonetheless we include their study in this list because its strategy resembles the one we are featuring.

[4]Selection into the labor market and the Heckman correction are discussed in the appendix.

[5]The working paper version of Bourguignon et al. (2007). In the peer-reviewed article results using a selection model are mentioned, but not reported.

sample, or d) individual income using a selection model into employment.

We will find that estimates are indeed sensitive to this methodological choice, although the relevant feature is not the aggregate itself, but the level of analysis they imply. This is, only little differences arise when comparing results obtained with aggregates at the household level (the second and third options), or with concepts focusing on the individual (the first and fourth). Nevertheless, IOP estimates do vary substantially if we make comparisons across the two levels of analysis.[6] As a matter of fact, we observe that this feature can be a more meaningful methodological choice than the selection of circumstances or the measurement approach, although the last two have received much more attention. We also find that this variation responds almost entirely to the contribution of the circumstance gender. In particular, considering an aggregate at the household level artificially nullifies the role of this characteristic. We conclude that if gender is to be considered a source of IOP, income at the individual level must be employed.

It is perhaps surprising that despite the multiple stances regarding the definition of income in the IOP literature, little attention has been payed to this feature. As mentioned above, studies generally state which income concept will consider and offer scant or no justification of it, while normative or robustness issues are rarely discussed. In fact some articles review exhaustively the many methods available to measure IOP (Ramos and Van de gaer 2016; Roemer and Trannoy 2016), discuss the robustness of IOP estimates to the measurement technique (Ramos and Van de gaer 2020), or study how accounting for welfare in ways other than income (subjective life satisfaction, access to leisure time and the sort) affects IOP results (Mahler and Ramos 2019), that nonetheless largely overlook the role of the level of analysis.

To our knowledge, this issue has only been analyzed before by Suárez Álvarez and López Menéndez (2020), who, using data of developing economies, conclude that the economic indicator chosen is unimportant. Previously, Bourguignon et al. (2003) noted differences between IOP estimates for individual earnings and household per capita incomes, which they attributed to the effect of assortative mating, fertility decisions, and non-labor income sources. We add to these contributions by analyzing the underlying mechanisms, and arrive at different conclusions.

The rest of the article is organized as follows. Section 2 presents the basics of IOP measurement, section 3 explores the implications of four income aggregates from a conceptual viewpoint, and in section 4 we discuss the methodology and results of an empirical demonstration. Section 5 concludes.

2. The measurement of opportunity inequality

In the "canonical" model of IOP, as described by Ferreira and Peragine (2016, p. 755), an individual outcome y is determined by a vector of personal circumstances $C = (c^1, \ldots, c^K)$ and a

[6] Yet, we prefer to report results with the four aggregates—two for each level of analysis—instead of just one per level for two reasons: first, comprehensiveness, and second, because all four aggregates have been used in the IOP literature.

scalar of effort $e \in [0, 1]$. The individual outcome is an economic good, i.e., it is universally desired with no satiation. Circumstances are characteristics that cannot be chosen, and therefore individuals should not be held responsible for them. These include gender, race, geographical origin, family background and the sort. Effort refers to the intensity with which individuals devote themselves to work or how responsibly they behave, and can, conversely, be decided. We have described it as a scalar, what is common in the literature, but it can be thought of as a vector.

Circumstances and effort belong to the finite sets Ω and θ, respectively. Then, y is a function $\Phi : \Omega \times \theta \to \mathbb{R}$, such that:

$$y = \Phi(C, e). \tag{1}$$

This can be seen as a reduced-form model in which a given outcome depends on circumstances and effort only, according to which all individuals sharing the same circumstances and exerting the same degree of effort would enjoy the same amount of outcome. Note that in eq. (1) the effect of circumstances on effort is not explicitly addressed.[7]

In recent years there has been an explosion in the number of methods available to empirically assess the extent of unfair inequality.[8] In this article we will use a popular procedure, the parametric approach proposed by Bourguignon et al. (2007) and Ferreira and Gignoux (2011).

This method is *ex-ante*, and no *ex-post* approaches will be considered in this text. Ex-ante techniques measure IOP considering circumstances only, while ex-post methods account for effort. Despite arguably being more normatively appealing, ex-post approaches are applied less frequently because they require more data. For more on the differences (and clash) between the ex-ante and ex-post perspectives see Fleurbaey and Peragine (2013).

In addition, this technique return IOP estimates that are generally interpreted as *lower-bounds*. This means that estimates represent the minimum value we can expect from them. This is so because the vector of circumstances C we are able to observe is a subset of the "true" vector C^*, such that $|C| \leqslant |C^*|$.[9] Nevertheless, recent research has questioned the interpretation as lower-bounds, pointing out that, on the one hand, estimates may suffer from an upward bias due to sampling variance (Brunori, Peragine, et al. 2019), and that, on the other hand, the parametric approach proposed by Bourguignon et al. (2007) and Ferreira and Gignoux (2011)[10] may produce IOP measurements transgressing the principle of transfers (Teyssier 2017), in which the proof by Ferreira and Gignoux (2011) relies. We will return to this below.

We proceed now to briefly describe this measurement approach. Suppose we have a population of individuals denoted by $i \in \{1, \ldots, I\}$, each of whom is fully characterized by the elements (y, C, e). A way to partition this population is proposed. We group individuals into

[7] For the constrain of circumstances on effort, see Roemer (1998).

[8] For surveys of the existing approaches to measure IOP see Ramos and Van de gaer (2016), Roemer and Trannoy (2016) or Ferreira and Peragine (2016).

[9] For the proof see Ferreira and Gignoux (2011), for a definition of *upper-bound* estimates go to Niehues and Peichl (2014), and for further discussion see Hufe et al. (2017).

[10] Among other measurement techniques (Van de gaer and Ramos 2020). We name this one only because it is the one employed in this text.

types $t_n \in T_n$, within which all members share the same combination of circumstances C^n. This partition is such that $t_1 \cup \ldots \cup t_N = \{1, \ldots, I\}$, $t_n \cap t_{n'} = \varnothing$, and $C_i = C_{i'}\ \forall i | i \in t_n, i' \in t_n$, $\forall n$.[11] Based on the realizations r_k of each circumstance c^k, the number of types is given by $N = \prod_{k=1}^{K} r_k$.[12]

The parametric procedure proposed by Bourguignon et al. (2007) and Ferreira and Gignoux (2011) consists of constructing a smoothed counterfactual of y by means of a (possibly log-linearized[13]) OLS regression of y against a vector of circumstances C:

$$y_i = \beta C_i + \varepsilon_i. \tag{2}$$

If estimates $\hat{\beta}$ were reliable, a parametric estimation of $\tilde{\mu} = \{\mu_1, \ldots, \mu_N\}$, the smoothed counterfactual of y, could be obtained by

$$\mu_n = \hat{\beta} C_i, \forall i \in n. \tag{3}$$

The elements of $\tilde{\mu}$, $\mu_n = \{\mu_n, \ldots, \mu_n\}$, are subvectors of length equal to the number of individuals belonging to type $t_n, \forall n$, and are given by (3) for each n. Of course, all individuals sharing the same circumstances will have the same predicted outcome.

Notice that in $\tilde{\mu}$ there is no inequality within types, remaining only inequality between types, i.e., inequality due to circumstances.[14] Then, an inequality index $I(\cdot)$ is applied to $\tilde{\mu}$ in order to obtain an absolute

$$\text{IOP}_{abs} = I(\tilde{\mu}), \tag{4}$$

and a relative measure of IOP

$$\text{IOP}_{rel} = \frac{I(\tilde{\mu})}{I(y)}. \tag{5}$$

Therefore, the criterion to identify the existence of IOP are differences between mean outcome levels across types. This is, $\exists t_n, t_{n'} \in T_n$ such that $\mu_n \neq \mu_{n'}$.

Notice that the decomposition of overall inequality into a "between" and "within" group inequality components precedes the measurement of IOP (see for example Foster and Shneyerov 2000; Shorrocks 1980). Since Checchi and Peragine (2010) introduced it in this field, the procedure has become a common practice in the IOP literature.

Despite having many alternatives to choose from, we use this method because it allows to break apart total IOP into the contribution of each circumstance by means of a Shapley-value decomposition (Shorrocks 2013; see also Ferreira, Gignoux, and Aran 2011), which is an useful

[11]Superscripts of C denote specific combinations of circumstances, while subscripts refer to the circumstances of particular individuals.

[12]In empirical applications of IOP it is common practice to consider only discrete variables as circumstances c^k, because continuous variables would dramatically increase the number of types, leading to very few observations, if any, in each type.

[13]The log-linearized version is advised against, see Teyssier (2017) and Van de gaer and Ramos (2020).

[14]Assuming all inequality between types is IOP implies regarding as normatively irrelevant possible differences in the amount of *absolute* effort exerted across types. See Roemer (1998).

feature we will put to use.

We will now discuss which properties could be required from the inequality index $I(\cdot)$ to be employed. We would like our measure to be Lorenz-consistent, for which it must satisfy the principles of symmetry, Pigou-Dalton transfer, and population and scale[15] invariance (e.g. Foster and Lustig 2019). This limits the choice to the known as summary indexes, of which the most commonly used are the Atkinson, Gini and generalized entropy measures. In addition we need our index to be additively decomposable, given that we intend to split total inequality into its fair and unfair shares. And, moreover, we would like it to be path-independent decomposable (Foster and Shneyerov 2000). This last requirement conveniently narrows the possible choices to just one, the Theil 0, also known as mean log deviation.[16] This index is defined as follows:

$$\text{MLD}(X) = \frac{1}{N}\sum_{i=1}^{N} \ln \frac{\bar{x}}{x_i}, \tag{6}$$

where N is the size of the sample, $x_i \in X$ is the outcome of observation i, and $\bar{x}$ is the mean of X. Notice that MLD is defined for positive x_i only, and also that $\text{MLD} = 0$ iff $\forall i, i' \in I, x_i = x_{i'}$, and $\text{MLD} > 0$ otherwise.

Still, in spite of the foregoing, the Gini index is sometimes employed in the IOP literature, and since we acknowledge that each inequality index implies a normative choice (Atkinson 1970) we will also employ the Gini index at some points, when testing for robustness.

3. Conceptual framework

In this section we study the implications from a conceptual perspective of choosing each one of the four income aggregates referred above. This choice will affect estimates through three channels: bias, sample composition, and its interplay with the inequality index employed. Bias alludes to assuming within-household redistribution, sample composition to the fact that each aggregate requires a specific data cleaning process, and the interplay with the inequality index arises from transformations performed to the data in order to build the aggregates, such as averaging income across household members to obtain per capita values.

Throughout this section we will study each of the channels separately, remaining agnostic about all factors but the one at hand. Subsequently, all effects combined will be analyzed empirically in section 4.

[15]It is standard to analyze IOP in relative terms, which is why we opt for scale invariance. However, the assessment could be also conducted in absolute terms, in which case this property shall be substituted with translation invariance.

[16]For further discussion on the axiomatic properties one might wish to require in the IOP setting see e.g. Ferreira and Gignoux (2011).

3.1. Bias

Let us start by analyzing the source of bias, abstracting by now from the other two channels. To do so we revisit the "canonical" model of IOP discussed in section 2:

$$y_i = \Phi(C_i, e_i), \tag{7}$$

where y_i is an individual outcome, C_i is a vector of personal circumstances and e_i is an effort scalar of individual $i \in I$. However, to simplify notation we will exclude effort from this discussion. Note that this omission does not compromise our analysis, since we will use ex-ante measures of IOP only, which do not take effort into account.[17] Hence, we reformulate (7):

$$y_i = \Psi(C_i). \tag{8}$$

We will begin with the case of per capita household income. This aggregate consists of adding the income received by all members of the household and then find the mean value,[18] which is imputed to each household member. To wit, independently of the income sources each particular member has, the analysis assumes they all receive the same amount. Then, crucially, these equal amounts are put in relation with the potentially different personal circumstances of every particular household member. In fact, according to eq. (8) estimates of IOP may be biased as long as $C_{ij} \neq C_{i'j}$ for some $i \in I, j \in J$, where J denotes households—this is, as long as at least two members of one household are heterogeneous with respect to some circumstance.

How large can the bias be, and which sign will it have? To see it begin by noticing that if $C_{ij} \neq C_{i'j}$, then $i \in t_n$ and $i' \in t_{n'}$, so averaging income across household members with heterogeneous circumstances equates to a transfer between types. To the extent these transfers are mostly progressive or regressive, and how large they are, will determine the sign and magnitude of the bias.

Let us back up a moment and state some definitions. A progressive transfer is defined as redistributing from the relatively rich to the relatively poor, without bringing the rich to a poorer situation than the poor. A regressive transfer is defined inversely. Additionally, an inequality index $I(\cdot)$ is said to satisfy the Pigou-Dalton transfer principle if its value diminishes after a progressive transfer has taken place, and inversely with a regressive transfer. This is, for all $Y, Y^* \in \mathbb{R}^n_+$ such that there exists $\delta > 0$ and $i, k \in I$ so that $y^*_i = y_i + \delta \leqslant y^*_k = y_k - \delta$, while for all $l \notin \{i, k\}$ it holds that $y^*_l = y_l$, then $I(Y^*) < I(Y)$.

[17] In addition, to assume homogeneity with respect to circumstances *and* effort is stronger than assuming it only for the former. We will get to this below.

[18] Apart from per capita, equivalent household income is commonly employed in the literature as well. This is obtained by applying equivalences of scale, which are used to account for non-linearities in the growth of households' needs with each additional member (e.g. Cowell 2000). Although the conceptual analysis we conduct in this section does not apply to equivalent income, in practice it entails a negligible difference with respect to per capita income. The same results of our empirical demonstration but employing equivalized household income are shown in section 4.3.

Now let us rewrite (8) as

$$y_j^H = \Upsilon(C_{1j}, \ldots, C_{ij}, \ldots, C_{I^j j}), \tag{9}$$

where I^j denotes individuals living in household j, and also the set $\{1j, 2j, \ldots, I^j j\}$. Outcome y_j^H, which refers to per capita household income, is a function of not only individual i's circumstances, but of all other members of her household j too. This is, of course, a consequence of assuming redistribution within the household.

If all household members are homogeneous with respect to C, or if household j has only one member (this is, $I^j = \{1\}$), then eq. (9) reduces to (8). However, in households with more than one member $C_{ij} \neq C_{i'j}$ will generally hold for two reasons: first, positive assortative mating[19] is not fully prevalent, and second, most couples are heterosexual. Positive assortative mating refers to homogeneity with respect to every circumstance except gender, while heterosexuality, naturally, concerns gender only. Table 1 confirms that the vast majority of European households with more than one member are conformed by individuals who are heterogeneous with respect to at least one circumstance, being gender the main driver behind it (only adults are considered—go to section 4.1 for a full description of the data). For instance, about 97% of these households are conformed by individuals who are heterogeneous with respect to gender (cross-country average). This includes couples, family and any other kind of arrangement; regarding relationships only, circa 99% of the declared couples in 2010 were heterosexual, cross-country average (not reported in table 1). With respect to other circumstances, consider the case of parental education: around 35% of households in our sample are constituted by adults whose parents attained a different educational level. Moreover, if we consider several circumstances at once $C_{ij} \neq C_{i'j}$ holds in nearly all cases, as the column "Homogeneity in 4 out of 4 [circumstances]" illustrates (again, go to section 4.1 for a full description of the circumstances and the rest of the data).

Returning to our previous question, the magnitude and sign of the bias will depend on the amount and progressiveness/regressiveness of the transfers between types produced by averaging income in the household. This in turn depends on the size of households on one side, and on the degree and pattern of heterogeneity of their members on the other. Note that taking $y_j^H = \Psi(C_i)$, i.e., assuming that per capita household income is a function of the circumstances of each member separately, may give rise to a bias in each and every household, and hence each family unit can slant estimates in its own way—be it positively or negatively.

The potentially different biases arising from each household will happen at once, and therefore might offset each other partially or completely in aggregate terms. If these effects cancel each other out or on the contrary they align and add up will depend on the pattern of heterogeneity of households' members. For instance, provided there is a consistent heterogeneity pattern of household members across the population, such as mostly heterosexual couples of

[19] Also known as homogamy, this is a mating pattern according to which people of similar economic, cultural and other societal characteristics tend to partner with one another.

Table 1: Households homogeneity

	Gender	Immigrant status	Parental education	Parental occupation	Homo. in 4 out of 4	Homo. in 3 out of 4	Homo. in 2 out of 4
Austria	1.37	90.45	51.10	67.67	0.46	34.58	78.36
Belgium	2.21	87.09	59.00	57.74	1.17	35.51	72.20
Bulgaria	4.71	99.16	64.69	66.82	2.63	47.94	81.80
Croatia	6.25	88.88	62.53	52.71	3.09	35.37	72.78
Cyprus	2.31	83.66	68.43	59.64	1.21	41.51	75.57
Czech R.	2.76	94.86	57.94	59.79	1.74	40.22	73.87
Denmark	-	-	-	-	-	-	-
Estonia	3.08	87.20	50.24	54.89	1.85	27.38	68.56
Finland	-	-	-	-	-	-	-
France	2.08	89.44	70.97	52.16	1.24	36.41	79.23
Germany	1.54	93.61	57.52	54.71	0.58	35.27	73.43
Greece	6.01	93.99	73.54	70.95	3.95	57.44	84.35
Hungary	3.05	98.91	67.73	62.85	1.53	47.96	82.31
Iceland	-	-	-	-	-	-	-
Ireland	5.23	83.42	55.14	54.95	2.60	30.47	69.98
Italy	4.58	92.79	74.26	58.40	3.09	44.98	82.85
Latvia	3.86	83.60	52.33	50.55	2.17	27.02	64.12
Lithuania	2.72	89.39	62.14	53.45	1.53	35.33	71.79
Luxembourg	1.58	82.22	64.12	63.38	0.41	38.33	76.53
Malta	3.72	89.81	76.13	52.99	2.50	40.79	81.79
Netherlands	-	-	-	-	-	-	-
Norway	-	-	-	-	-	-	-
Poland	2.91	99.72	64.79	70.90	1.67	49.79	85.95
Portugal	4.48	91.70	92.09	69.21	3.43	62.98	92.51
Romania	3.29	99.84	83.33	77.72	2.50	69.16	91.73
Slovakia	3.62	98.19	61.35	53.67	2.07	36.14	76.74
Slovenia	-	-	-	-	-	-	-
Spain	4.15	94.57	79.81	59.30	2.73	49.80	86.27
Sweden	-	-	-	-	-	-	-
Switzerland	1.80	79.83	54.09	55.54	0.52	26.66	69.51
United K.	2.18	91.33	52.90	50.65	0.65	27.01	71.60
Average	3.31	90.99	64.84	59.61	1.89	40.75	77.66

Note: This table shows the percentage of households of which members have homogeneous circumstances. Only households with 2 or more members are considered, which is why some countries have missing information (in these countries there is information on parental features about one member per household only—see section 4.1). "Homo. in x out of 4" refers to homogeneity with respect to x circumstances out of the 4 considered. Cross-sectional files of the EU-SILC database.

women who tend to earn less than their male partners, the overall bias produced by taking $y_j^H = \Psi(C_i)$ instead of $y_i = \Psi(C_i)$ will be negative.

Let us analyze this issue more formally. Remember that if $C_{ij} \neq C_{i'j}$ then $i \in t_n$ and $i' \in t_{n'}$, and denote personal income of individual i living in household j by y_{ij}. Then, all $i \in I, j \in J$ such that $y_{ij} < y_{i'j}$ and $\mu_n < \mu_{n'}$ will produce a negative bias (where μ_n denotes mean individual income of type n—see section 2). This is, each person living in a household with another individual who enjoys higher personal income and belongs to a richer type than herself will bias IOP estimates downwards. This is so because in this particular case taking $y_j^H = \Psi(C_i)$ equates to a progressive transfer between types.

On the contrary, a positive bias will be produced for all $i \in I, j \in J$ such that $y_{ij} > y_{i'j}$ and $\mu_n < \mu_{n'}$ (or equivalently, $y_{ij} < y_{i'j}$ and $\mu_n > \mu_{n'}$); namely, for each individual who lives in a household with another individual who has lower (higher) personal income and, nonetheless, belongs to a richer (poorer) type. In this case taking $y_j^H = \Psi(C_i)$ equates to a regressive transfer between types.

In addition, no bias will arise if $y_{ij} = y_{i'j}$, even if $\mu_n \neq \mu_{n'}$. In this case taking $y_j^H = \Psi(C_i)$ will leave every type's mean income μ unaltered.

Remark 1 Let y^H be per capita household income, and $\tilde{\mu}^H$ its smoothed counterfactual distribution. All $i \in I, j \in J$ such that $y_{ij} < y_{i'j}$ and $\mu_n < \mu_{n'}$ will negatively bias $I(\tilde{\mu}^H)$.

Proof By imputing $y_{ij} = y_{i'j} = y_j^H$ it follows that $\mu_n < \mu_n^H$ and $\mu_{n'}^H < \mu_{n'}$. Provided $I(\cdot)$ satisfies the Pigou-Dalton transfer principle and as long as $\mu_n^H \leqslant \mu_{n'}^H$, it follows that the imputation reduces $I(\tilde{\mu}^H)$. ■

Corollary 1.1 All $i \in I, j \in J$ such that $y_{ij} > y_{i'j}$ and $\mu_n < \mu_{n'}$, or $y_{ij} < y_{i'j}$ and $\mu_n > \mu_{n'}$, will positively bias $I(\tilde{\mu}^H)$.

Corollary 1.2 All $i \in I, j \in J$ such that $y_{ij} = y_{i'j}$, even if $\mu_n \neq \mu_{n'}$, will leave $I(\tilde{\mu}^H)$ unbiased.

In conclusion, considering per capita household income may bias IOP estimates with respect to the "canonical" model (7) in any direction. The bigger the size of households, and the more heterogeneous their members are with respect to C and y, the larger the bias can be. However, since effects take place within households, every household can bias results, all at once. It is hence possible that, provided there is no consistent pattern of heterogeneity, the negative and positive biases offset each other. Nonetheless, we will see that in practical applications considering y^H will generally bias IOP estimates downwards, specifically by virtually nullifying the contribution of gender to overall IOP. This is a consequence of two facts: a) women tend to earn less than their male partners[20] (and hence mean incomes of women's types are generally lower than those of males'), and b) in most households live adults of both genders (see table 1).

[20] Our data is from European economies, but this pattern extends well beyond. For instance, Bertrand et al. (2015) find that in the US most women have lower salaries than their husbands, and that when this is not the case, marital satisfaction is lower and divorce more likely.

In other words, it is consequence of a consistent pattern of heterogeneity with respect to gender and personal incomes. Therefore, when empirically assessing the extent of IOP we will generally find that $I(\tilde{\mu}^H) < I(\tilde{\mu})$, being $\tilde{\mu}$ the smoothed counterfactual of individual income y_{ij}. We will explore this further in section 4.2.

The next aggregate we are going to discuss is household income but retaining only household heads in the sample, which we denote by y^{HE}. This aggregate is, obviously, also at the household level. Furthermore, it also suffers from the biases described above because, again, the amount of income we are imputing to each head is determined not only by her circumstances, but by every other household members' as well.

However, the magnitude of the bias *per household* will be smaller with y^{HE} than with y^H, due to a sample effect—i.e., only the head from each household will be considered in the analysis, instead of every member. Emphasis on per household is due because in aggregated terms the total bias of taking $y_j^{HE} = \Phi(C_i)$ can be larger or smaller than with $y_j^H = \Phi(C_i)$. It will depend on the extent to which the potentially different biases arising from each household offset each other, or if on the contrary they add up.

Notice that by construction the sample size when considering y^{HE} cannot be larger than with y^H. This is, let N^H be the sample size of y^H and N^{HE} be that of y^{HE}, then $N^{HE} \leq N^H$. In fact, N is the only potential difference between y^H and y^{HE}, and consequently y^{HE} equates to a reweighting of y^H. When considering y^{HE} the income of each household, independently of their number of members, is related to one set of personal circumstances only—those of the head. This is, the number of times we take $y_j^{HE} = \Phi(C_i)$ is once per household only; however, with y^H we take $y_j^H = \Phi(C_i)$ for each household member. Thus, just as with the sample size, the magnitude of the bias per household when considering y^{HE} cannot be larger than with y^H too. Recall that we stated that the magnitude of the bias with y^H will in part depend on the size of households—indeed, keeping only heads in the sample works through this mechanism.

Remark 2 Let y^H be household income, y^{HE} household income but keeping only heads in the sample, and $\bar{B}^H$ and $\bar{B}^{HE}$ the magnitude of their biases per household, respectively. Since $N^{HE} \leq N^H$, then $\bar{B}^{HE} \leq \bar{B}^H$.

Nevertheless, as we will see in section 4.2, in practice results change only slightly when we consider y^{HE} instead of y^H, suggesting there is no meaningful difference between the two. However, total bias with y^{HE} may be larger or smaller than with y^H, and will depend on the sample composition—we will return to this in sections 3.2 and 4.2. In any case, it will be shown that the contribution of the circumstance gender to total IOP is also nearly nullified with y^{HE}.

Let us now consider personal income. Even though the "canonical" model of IOP assumes y to be an individual outcome, individual measures of income can be problematic.[21] As we mentioned at the beginning of this article, focusing on the individual level neglects household bargaining processes that affect labor market participation, which is one of the reasons why

[21]The situation is of course different when the outcome of choice is, say, health status or educational attainment.

considering income at the household level is standard in the analysis of income inequality (e.g. Jenkins and Van Kerm 2009). These processes refer to situations in which, for instance, an individual with a comparative advantage over her spouse in the household sector might rationally decide to stop working. Failure to account for this issue will lead to an overestimation of IOP, because observed differences between type's mean income μ will not only respond to IOP. For example, the educational level of each household member might influence the household bargaining process, and personal effort plays a role in educational choices.

However, abstracting completely from the household bargaining process may produce underestimates of IOP, since part of the effect of circumstances on personal achievement works precisely through labor supply decisions. This is specially true for gender. For instance, Kleven, Landais, Posch, et al. (2019) and Kleven, Landais, and Søgaard (2019), using data from European countries, find that having children strongly impacts labor market participation and hours worked of women, but not of men. They do not find that public policies explain this difference, but rather that it is rooted in gender identity norms. Furthermore, Goldin (2014) and Goldin and Katz (2016) highlight that "family friendliness" of occupations is the main factor behind the gender pay gap. In fact there is a large strand of literature suggesting that gender identity norms condition females joining the labor force and their willingness to work long hours.[22] Simultaneously, other personal circumstances can influence participation in the labor market too. Family background may shape labor supply in many ways, including social connections and genetic transmission of characteristics valued in the labor market (Corak 2013; Papageorge and Thom 2019; Roemer 2004). Additionally, certain social groups such as immigrants may be specially prone to suffer from unemployment and underemployment (Bisin et al. 2011; De La Rica et al. 2015), what can impact their incentives to take part in the labor market.

In sum, when considering income at the personal level we are not estimating eq. (8), but

$$y_{ij}^W = \Gamma(C_{ij}, p(C_{1j}, \ldots, C_{ij}, \ldots, C_{I^j j})), \tag{10}$$

where y_{ij}^W denotes individual income (in our application we will consider labor earnings, so the W is for wage), and $p(\cdot)$ denotes the bargaining process in household j that conditions labor market participation of individual i. Note that $p(\cdot)$ depends on the circumstances of all members of the household, $\{1, \ldots, i, \ldots, I^j\}$, not only on C_{ij}. Hence, as in eq. (9), household composition affects results with y_{ij}^W, although in (10) it does so indirectly.

The bargaining process $p(\cdot)$ affects IOP estimates in two ways: on the one hand, by pooling non-comparable individuals (such as full-time workers and people willingly out of the labor market), and on the other hand, through the effect of C on labor market participation. Let us call the former effect u, which cannot be identified with IOP, and the latter effect v, which is behind some part of overall IOP. Hence, the bargaining process in a given household is:

$$p(\cdot) = \tau(u, v(C_{1j}, \ldots, C_{ij}, \ldots, C_{I^j j})), \tag{11}$$

[22] See e.g. Bursztyn et al. (2017) and Teso (2019) for experimental evidence.

where $v(\cdot)$ depends on the circumstances of all members of the household.

If $p(\cdot)$ is not taken into account when estimating IOP, all differences between the null income of individuals out of the labor force and the positive income of people at work are assumed by construction to be unfair, which is obviously problematic. Yet, removing $p(\cdot)$ completely will "leave out" part of IOP. Therefore, we would ideally account for $v(\cdot)$, because it is related to IOP, while abstracting from u, which is not. Then, instead of eq. (10) we would take

$$y_{ij}^{W} = \Theta(C_{ij}, v(C_{1j}, \ldots, C_{ij}, \ldots, C_{I^j j})). \tag{12}$$

A possible way to get rid of u is to use a sample composed of individuals with a similar level of labor market attachment. For example, Andreoli and Fusco (2017) study individuals who spent most of the income reference period as full-time workers, while Bourguignon et al. (2007) use a sample of males active in the labor market who report positive earnings.[23]

Nonetheless this approach is not free from problems. Specifically, when constructing a sample of individuals with a similar level of labor market attachment the researcher faces a trade-off: the more similar they are, the smaller the influence of u will be, but a larger part of $v(\cdot)$ will be removed as well, and vice versa. Consider a few examples. In a sample of employed individuals some will work full-time while others may willingly do it part-time, so u would not be removed completely. This is so because these individuals would not be perfectly comparable, i.e., their labor market attachment differs.[24] Inversely, a sample composed of full-time workers weakens the assumption that u has been removed, but parallelly it implies removing some part of $v(\cdot)$ too—the reason being that part-time working is not randomly assigned and in some degree depends on C. Moreover, by considering only people at work we are excluding the unemployed (who by definition would prefer to be working), again removing part of the effect of $v(\cdot)$, since unemployment is also not randomly distributed.

In sum, using y^W we fail to properly account for the household bargaining process $p(\cdot)$. We are forced to either pool non-comparable individuals—letting u affecting our estimates—or miss out some IOP produced by the influence of C on labor market participation—removing part of the effect of $v(\cdot)$. In light of this shortcoming we believe it is preferable to explicitly model participation in the labor market, what we discuss in the rest of this section.

Now we are going to discuss our last income aggregate: personal income with a selection model into employment, which we denote by y^{SW}.[25] The implications for IOP estimation are that if the assumptions of the selection model are satisfied we could remove u completely. However, contrary to what occurs with the previous three aggregates, $v(\cdot)$ would not be removed

[23] However, other studies considering individual income simply drop all observations with zero income (Fleurbaey, Peragine, and Ramos 2017), or keep a sample of males with similar age (Björklund et al. 2012).

[24] This would not be the case if we could control for the willingness of part-time workers to work more hours, and include in the sample only those who would prefer to be working full-time; that is, if we could account for underemployment. Regrettably though, in the data source we will employ—the EU-SILC database—this is not possible.

[25] A description of the selection bias, the Heckman correction method and the actual selection model we employ can be found in the appendix.

alongside it. The reason for this is that the selection model imputes *expected* incomes to individuals who do not match a certain level of labor market attachment, based on the income of similar individuals who do match it. Hence, provided that the model works as intended, a sample of individuals with comparable income can be constructed while the effect of C on labor market participation is retained. This is, u can be removed while preserving the effect of $v(\cdot)$.

Assumption 1 The data with which we employ our selection model is fit for it and hence the model works as intended.

Remark 3 If Assumption 1 holds, with a sample such that $u^W = u^{SW} = 0$ we would have $v^W(\cdot) \leqslant v^{SW}(\cdot)$.

In the appendix we describe the Heckman selection model we use and discuss that no evidence against Assumption 1 is found in our analysis, so y^{SW} is our preferred income aggregate. Nonetheless, as we will see in section 4.2, results with y^{SW} and y^W do not differ substantially and hence the choice between them is not a major methodological feature.

3.2. Sample composition

The second channel through which the choice of an income aggregate will affect IOP estimates are the differences between the samples employed. Depending on the aggregate a specific data cleaning process is in order, and a fair comparison of IOP using different aggregates must factor this in. For instance, if an estimation of IOP is being conducted with "labor earnings" as the outcome of interest, observations with missing or non-positive values on the outcome variable "labor earnings" might be dropped, while that would not be done if the outcome of interest is "household income". Likewise, we would not drop observations with missing or non-positive labor earnings if we are going to employ a selection model into employment. For this reason we keep four different samples throughout the analysis, one for each aggregate. These samples are nonetheless identical in everything other than the requirements of the idiosyncratic data processing of each aggregate. However, we will find that the sample composition plays only a secondary role in shaping the differences across estimates.

Table 2 shows some descriptive statistics of each sample, by country. These consist of the percentage of individuals with the referenced characteristics (namely gender, immigrant status, parental education and parental occupation, which constitute the set of circumstances we are going to employ), alongside sample size, the average value of the income aggregate, and its standard deviation.[26] This table spans 6 pages, the first 3 of which relate to samples of aggregates at the household level (y^H and y^{HE}), relating the last 3 to samples of individual outcomes (y^W and y^{SW}). Also, each page contains statistics of two different samples for 10 or 11 countries.

Let us now review the differences across samples. Comparing samples with aggregates at the household level (first three pages of table 2) we find that the sample size is always notably

[26]Note that the data are described extensively in section 4.1.

larger with y^H than with y^{HE}'s, even up to more than two times larger. The only exception is Sweden, where y^{HE} sample is just a little smaller. In fact, samples with y^{HE} can only be as large or smaller than with y^H, since we are using the same income aggregate but retaining only individuals who are the head of their household (see Remark 2 in section 3.1). Regarding the mean outcome level, there seems to be only small differences. If anything, y^{HE} appears to be slightly higher (with the exception of Sweden), but this differences are not meaningful. Likewise, the standard deviation of the mean outcome is generally higher in the sample of y^{HE}, but just slightly (the only exception is Denmark). Moving on to circumstances, we see that in y^H's samples there tends to be gender-parity, with each sex representing about half of them. However, women are less frequently household heads, in some cases by a substantial difference. Consider for instance Cyprus, Greece or Romania, where only 17%, 19% and 18% of the heads are female, respectively. On the other end we find Bulgaria, Latvia and Sweden, where the proportion of women is actually somewhat higher in y^{HE}'s samples. In addition, differences with respect to gender are solely small in Denmark, Estonia, Finland, Lithuania and Slovenia. Nonetheless, there appears to be a general trend and it is likely to respond to gender identity norms, which incentivize women to allocate a larger portion of their time, compared to men, to the household sector (see the discussion of personal income in section 3.1). Apropos of immigration, no major differences seem to arise between the samples of the two aggregates at the household level. This may be due to the fact that in most households where there is one immigrant, all household members have the same status too (see table 1), so this social group is well represented in y^{HE}'s samples. Finally, regarding family background, we observe similar distributions of parental education and occupation levels across the samples of these two aggregates. At most, in y^{HE}'s samples we find distributions tending to some extent towards the higher levels, but in an almost negligible fashion. If anything, we could have expected household heads to enjoy a more privileged background than the general population.

We will compare now samples with aggregates at the individual level, y^W and y^{SW} (last three pages of table 2). When using a selection model into employment, y^{SW}, the sample size is of course larger, with no exception. The average income is about the same with both aggregates; depending on the country it is larger in one or the other sample, but in no case the difference is substantial. Nevertheless, we do find disparities when it comes to standard errors. SEs are always larger with y^W than with y^{SW}, and the disparities are in some cases sizable, such as in Spain, Ireland or the United Kingdom. With respect to gender, the proportion of women is generally similar, although always higher in the y^{SW}'s samples (there are three exceptions: Estonia, Lithuania and Latvia). These discrepancies, although small, go in line with what has been discussed above about the allocation of female's time between the household and market sectors. A stark difference appears in Malta, where women move from representing 32% of the y^W sample to 49% with y^{SW}. Regarding immigration we find a parallel situation to that of gender: the relative presence of immigrants is broadly similar across the samples of the two aggregates, but it tends to be higher in the y^{SW}'s ones (with the exception of Bulgaria, Finland,

Table 2: Summary statistics

	AT	BE	BG	CH	CY	CZ	DE	DK	EE	ES
y^H										
Sample size	5,607	4,521	5,994	5,913	4,310	6,023	9,187	2,430	4,404	13,318
Outcome										
Average	18,145.97	16,461.79	2,400.21	31,005.40	13,764.40	5,981.33	17,298.12	23,998.32	4,749.96	11,564.07
Standard deviation	10,303.90	8,344.25	1,568.46	19,183.52	8,534.25	3,140.15	9,913.20	12,166.18	3,091.80	7,505.50
Demography										
Female	52.26	50.90	50.12	53.59	54.48	57.84	52.45	52.30	52.52	51.44
Immigrant status	15.62	17.14	0.50	22.73	19.28	3.19	5.68	6.50	12.15	9.36
Parental education										
Primary or less	36.74	49.72	50.52	21.77	72.65	58.79	8.73	30.91	32.61	83.92
Secondary	47.62	25.64	36.70	59.77	18.38	30.98	59.64	41.93	44.73	6.92
Tertiary or more	15.64	24.64	12.78	18.47	8.98	10.23	31.63	27.16	22.66	9.17
Parental occupation										
Elementary	7.62	5.24	10.66	3.57	13.18	4.35	2.55	0.45	5.86	16.11
Skilled workers	74.25	56.56	66.93	52.48	69.88	62.96	55.32	62.63	56.38	62.94
Professionals	18.14	38.20	22.41	43.95	16.94	32.69	42.14	36.91	37.76	20.95
y^{HE}										
Sample size	3,323	2,615	2,589	3,463	2,065	3,278	5,416	2,221	2,242	6,547
Outcome										
Average	18,594.90	16,895.71	2,491.03	32,029.23	14,203.48	6,072.56	17,571.30	24,136.32	4,941.92	11,974.32
Standard deviation	10,719.31	8,768.28	1,671.40	19,961.37	8,982.61	3,316.70	10,321.32	12,147.13	3,287.09	7,936.56
Demography										
Female	39.15	28.83	57.24	39.39	17.48	33.59	36.34	50.20	47.59	38.08
Immigrant status	14.32	16.21	0.73	21.77	13.90	3.33	4.67	6.12	11.78	9.35
Parental education										
Primary or less	36.08	48.53	52.84	21.48	72.93	57.72	8.03	29.67	32.78	84.36
Secondary	47.88	25.85	34.61	59.89	18.74	31.24	59.56	42.28	43.89	6.58
Tertiary or more	16.04	25.62	12.55	18.63	8.33	11.04	32.40	28.05	23.33	9.06
Parental occupation										
Elementary	7.40	5.01	11.47	3.26	13.17	4.00	2.49	0.41	6.02	16.22
Skilled workers	73.07	54.80	66.59	51.92	70.27	62.45	54.56	62.13	55.53	62.18
Professionals	19.53	40.19	21.94	44.82	16.56	33.56	42.95	37.46	38.45	21.60

Note: This table shows summary statistics for each sample and country. Samples are referenced by the notation employed throughout this section: y^H refers to household income, y^{HE} to household income but keeping only heads in the sample, y^W to labor earnings, and y^{SW} to labor earnings using a selection model into employment. Aside from sample size, the average value of the outcome and its standard deviation, numbers refer to the percentage of individuals with the referenced characteristics. Cross-sectional files of the EU-SILC database.

Continuation of table 2: Summary statistics

	FI	FR	GR	HR	HU	IE	IS	IT	LT	LU
y^H										
Sample size	2,893	9,182	4,271	4,920	11,970	2,725	1,460	18,246	4,465	6,114
Outcome										
Average	19,940.98	17,035.51	9,118.64	4,326.60	3,608.06	16,800.68	14,436.53	12,915.08	3,237.43	26,344.59
Standard deviation	10,142.73	10,879.76	6,239.23	2,541.15	2,091.37	10,064.13	7,280.12	8,318.67	2,208.47	14,929.95
Demography										
Female	46.83	52.53	51.65	51.50	53.78	57.14	50.41	51.67	54.42	51.80
Immigrant status	4.48	10.55	10.04	11.44	1.04	21.43	8.77	8.37	6.47	52.03
Parental education										
Primary or less	49.31	76.38	77.17	56.42	60.94	44.00	27.88	76.43	58.45	52.01
Secondary	28.17	10.51	14.26	35.75	28.34	37.94	54.66	18.07	29.05	33.69
Tertiary or more	22.52	13.11	8.57	7.83	10.73	18.06	17.47	5.50	12.50	14.30
Parental occupation										
Elementary	5.72	19.46	4.68	26.12	12.35	16.22	3.42	12.82	22.75	4.24
Skilled workers	58.66	48.75	77.83	53.13	65.91	50.64	55.89	62.77	52.63	61.63
Professionals	35.62	31.79	17.49	20.75	21.75	33.14	40.68	24.41	24.61	34.13
y^{HE}										
Sample size	2,591	5,249	1,936	2,067	6,350	1,614	913	9,026	2,272	3,366
Outcome										
Average	20,086.99	17,472.37	9,526.28	4,450.45	3,701.29	16,847.63	14,657.73	13,625.75	3,291.88	27,234.81
Standard deviation	10,226.46	11,221.78	6,645.23	2,678.22	2,239.09	10,356.75	7,660.03	9,003.14	2,362.76	15,953.38
Demography										
Female	45.69	36.27	19.27	25.21	32.24	47.34	38.23	33.45	54.36	33.69
Immigrant status	4.12	9.87	10.80	11.71	0.93	22.74	9.20	7.82	6.07	51.07
Parental education										
Primary or less	49.46	75.10	76.76	59.46	61.31	44.55	29.46	75.53	61.27	51.66
Secondary	27.71	11.03	14.00	32.41	26.91	37.86	53.23	18.26	26.76	34.22
Tertiary or more	22.83	13.87	9.25	8.13	11.78	17.60	17.31	6.22	11.97	14.11
Parental occupation										
Elementary	5.74	18.92	4.70	27.53	12.36	16.67	3.72	12.39	23.50	4.01
Skilled workers	58.51	48.20	76.91	51.72	65.21	49.75	55.42	62.23	52.42	61.47
Professionals	35.76	32.88	18.39	20.75	22.43	33.58	40.85	25.38	24.08	34.52

Note: This table shows summary statistics for each sample and country. Samples are referenced by the notation employed throughout this section: y^H refers to household income, y^{HE} to household income but keeping only heads in the sample, y^W to labor earnings, and y^{SW} to labor earnings using a selection model into employment. Aside from sample size, the average value of the outcome and its standard deviation, numbers refer to the percentage of individuals with the referenced characteristics. Cross-sectional files of the EU-SILC database.

Continuation of table 2: Summary statistics

	LV	MT	NL	NO	PL	PT	RO	SE	SI	SK	UK
y^H											
Sample size	5,565	3,643	4,593	2,390	13,156	5,248	5,384	2,390	4,129	6,139	5,295
Outcome											
Average	3,564.05	8,574.19	17,675.26	30,437.24	3,983.38	7,183.69	1,642.07	18,771.40	9,020.86	4,713.98	16,079.58
Standard deviation	2,495.79	4,708.19	9,400.43	14,069.41	2,685.52	5,246.52	1,095.10	8,926.64	4,420.93	2,383.17	10,689.58
Demography											
Female	55.45	52.05	53.43	45.94	53.43	52.76	50.50	51.34	53.02	53.51	54.90
Immigrant status	13.23	5.57	5.55	8.28	0.15	7.34	0.09	6.32	11.65	1.17	12.58
Parental education											
Primary or less	41.67	73.57	35.97	23.35	49.29	93.14	85.20	34.06	66.14	35.84	52.45
Secondary	41.80	19.30	38.65	42.51	43.36	3.07	11.65	40.59	21.77	55.76	25.10
Tertiary or more	16.53	7.14	25.39	34.14	7.35	3.79	3.16	25.36	12.09	8.41	22.46
Parental occupation											
Elementary	11.61	11.94	3.27	1.92	6.97	9.32	11.92	-	17.51	11.86	6.44
Skilled workers	55.45	62.48	50.08	47.82	74.48	75.82	76.75	-	55.58	59.06	49.84
Professionals	32.94	25.58	46.66	50.25	18.55	14.86	11.33	-	26.91	29.08	43.72
y^{HE}											
Sample size	2,997	1,748	2,881	1,733	6,517	2,432	2,443	2,375	2,787	2,932	3,055
Outcome											
Average	3,657.06	8,746.77	18,036.60	30,675.10	4,156.07	7,360.27	1,676.80	18,769.83	9,281.21	4,737.43	16,290.33
Standard deviation	2,633.11	5,016.24	9,585.14	14,238.85	2,904.31	5,526.34	1,148.35	8,945.88	4,614.41	2,451.23	10,953.10
Demography											
Female	59.46	24.66	37.90	35.55	39.99	29.65	18.38	51.49	49.55	35.37	43.47
Immigrant status	13.18	5.38	5.87	6.87	0.08	7.40	0.08	6.32	11.48	1.36	12.18
Parental education											
Primary or less	43.91	74.37	36.58	23.54	50.22	92.43	85.80	33.98	65.37	37.93	52.08
Secondary	39.24	19.22	37.70	42.35	41.68	3.33	11.01	40.72	21.10	53.41	25.14
Tertiary or more	16.85	6.41	25.72	34.10	8.10	4.24	3.19	25.31	13.53	8.66	22.78
Parental occupation											
Elementary	11.68	11.44	3.30	1.79	6.78	9.58	12.44	-	16.86	12.11	6.28
Skilled workers	55.16	62.30	50.16	47.55	74.18	74.10	76.59	-	55.18	58.77	49.20
Professionals	33.17	26.26	46.55	50.66	19.04	16.32	10.97	-	27.95	29.13	44.52

Note: This table shows summary statistics for each sample and country. Samples are referenced by the notation employed throughout this section: y^H refers to household income, y^{HE} to household income but keeping only heads in the sample, y^W to labor earnings, and y^{SW} to labor earnings using a selection model into employment. Aside from sample size, the average value of the outcome and its standard deviation, numbers refer to the percentage of individuals with the referenced characteristics. In the case of Sweden we are forced to exclude the circumstance parental occupation due to the very small number of respondents. Cross-sectional files of the EU-SILC database.

Continuation of table 2: Summary statistics

	AT	BE	BG	CH	CY	CZ	DE	DK	EE	ES
y^W										
Sample size	4,420	3,474	4,437	4,940	3,367	4,677	7,471	2,099	3,161	8,879
Outcome										
Average	36,003.02	37,235.43	4,041.75	60,867.89	25,720.09	11,125.30	33,149.83	50,718.54	9,759.85	23,228.05
Standard deviation	24,817.35	19,698.89	2,533.61	40,649.19	16,428.79	6,564.32	21,182.96	24,248.22	6,928.02	16,092.22
Demography										
Female	47.24	47.21	49.70	47.49	48.23	52.21	48.51	51.21	53.18	44.62
Immigrant status	13.91	13.44	0.52	21.76	18.06	3.08	5.38	5.53	11.23	8.44
Parental education										
Primary or less	35.34	45.54	44.58	20.45	71.28	56.68	8.22	30.63	29.83	81.41
Secondary	48.10	27.40	40.66	60.67	19.13	32.56	59.39	41.83	45.84	7.69
Tertiary or more	16.56	27.06	14.76	18.89	9.59	10.75	32.39	27.54	24.33	10.90
Parental occupation										
Elementary	6.81	4.84	8.29	3.28	12.80	3.68	2.44	0.48	5.25	13.45
Skilled workers	74.34	54.61	66.49	51.84	69.62	62.26	54.62	61.89	54.57	63.18
Professionals	18.85	40.56	25.22	44.88	17.58	34.06	42.94	37.64	40.18	23.37
y^{SW}										
Sample size	5,420	4,174	5,964	5,763	4,111	5,968	9,132	2,328	4,338	12,135
Outcome										
Average	35,353.84	37,365.16	3,911.13	59,228.11	24,918.21	10,867.96	32,399.33	50,326.14	9,584.20	23,487.60
Standard deviation	22,217.58	18,029.12	2,137.09	37,649.82	15,194.29	5,764.52	19,333.50	21,820.21	6,060.12	13,978.17
Demography										
Female	51.51	50.57	50.13	52.63	53.66	57.89	52.24	52.41	52.77	49.65
Immigrant status	14.94	15.36	0.50	22.38	19.27	3.18	5.55	5.67	12.19	9.09
Parental education										
Primary or less	36.49	49.31	50.29	21.34	72.54	58.66	8.65	30.63	32.57	83.18
Secondary	47.75	26.09	36.85	60.18	18.24	31.13	59.70	42.05	44.86	7.14
Tertiary or more	15.76	24.60	12.86	18.48	9.22	10.20	31.65	27.32	22.57	9.67
Parental occupation										
Elementary	7.55	5.32	10.43	3.49	13.26	4.29	2.56	0.47	5.83	15.35
Skilled workers	74.30	56.61	67.04	52.56	69.55	62.92	55.34	62.46	56.39	63.02
Professionals	18.15	38.07	22.54	43.95	17.20	32.79	42.09	37.07	37.78	21.62

Note: This table shows summary statistics for each sample and country. Samples are referenced by the notation employed throughout this section: y^H refers to household income, y^{HE} to household income but keeping only heads in the sample, y^W to labor earnings, and y^{SW} to labor earnings using a selection model into employment. Aside from sample size, the average value of the outcome and its standard deviation, numbers refer to the percentage of individuals with the referenced characteristics. Cross-sectional files of the EU-SILC database.

Continuation of table 2: Summary statistics

	FI	FR	GR	HR	HU	IE	IS	IT	LT	LU
y^{W}										
Sample size	2,215	7,298	2,871	2,841	8,286	1,689	1,238	13,201	3,158	4,565
Outcome										
Average	38,077.41	30,203.79	21,401.82	11,097.48	7,212.12	39,772.35	30,406.04	28,915.87	6,566.77	52,536.85
Standard deviation	19,173.14	20,803.70	16,103.89	6,229.17	4,599.30	27,886.91	18,442.64	19,044.67	4,683.75	34,242.51
Demography										
Female	46.24	49.48	42.39	46.88	49.96	52.46	47.82	43.41	54.88	43.94
Immigrant status	3.52	9.17	9.33	9.82	1.10	20.13	8.80	7.70	6.46	51.72
Parental education										
Primary or less	47.90	74.62	74.61	50.19	56.11	39.08	27.06	73.71	55.22	49.29
Secondary	29.25	11.26	15.57	39.88	31.51	40.08	55.33	20.01	30.78	34.72
Tertiary or more	22.85	14.11	9.82	9.93	12.38	20.84	17.61	6.27	14.00	15.99
Parental occupation										
Elementary	5.14	17.66	4.39	22.03	9.86	12.37	3.31	11.15	20.65	3.92
Skilled workers	58.13	48.62	76.21	53.26	65.41	50.74	55.41	62.11	52.79	60.13
Professionals	36.73	33.72	19.40	24.71	24.73	36.89	41.28	26.74	26.57	35.95
y^{SW}										
Sample size	2,860	8,728	3,823	4,487	11,720	2,652	1,453	16,629	4,429	5,978
Outcome										
Average	38,601.31	29,550.33	21,295.08	11,113.56	7,038.80	42,207.14	30,171.25	28,515.34	6,427.83	51,667.65
Standard deviation	16,655.86	18,523.89	14,252.18	5,270.85	3,977.16	23,798.22	16,773.32	16,918.38	4,073.13	31,395.40
Demography										
Female	46.88	52.31	48.21	50.08	53.81	57.20	50.31	48.88	54.59	51.17
Immigrant status	3.31	10.15	10.31	10.88	1.05	20.14	8.74	8.20	6.57	51.69
Parental education										
Primary or less	49.46	76.13	76.38	54.76	60.91	44.00	27.87	75.53	58.66	51.52
Secondary	28.01	10.67	14.67	37.04	28.39	37.93	54.78	18.73	29.01	33.99
Tertiary or more	22.53	13.20	8.95	8.20	10.70	18.06	17.34	5.74	12.33	14.49
Parental occupation										
Elementary	5.76	19.16	4.81	25.01	12.15	15.91	3.44	12.02	22.98	4.22
Skilled workers	58.56	48.87	77.24	53.35	66.08	50.79	55.95	62.87	52.56	61.58
Professionals	35.68	31.98	17.94	21.64	21.77	33.30	40.61	25.11	24.45	34.21

Note: This table shows summary statistics for each sample and country. Samples are referenced by the notation employed throughout this section: y^{H} refers to household income, y^{HE} to household income but keeping only heads in the sample, y^{W} to labor earnings, and y^{SW} to labor earnings using a selection model into employment. Aside from sample size, the average value of the outcome and its standard deviation, numbers refer to the percentage of individuals with the referenced characteristics. Cross-sectional files of the EU-SILC database.

Continuation of table 2: Summary statistics

	LV	MT	NL	NO	PL	PT	RO	SE	SI	SK	UK
y^W											
Sample size	3,783	2,277	3,979	2,100	8,525	3,728	3,937	2,105	3,375	4,884	4,163
Outcome											
Average	7,763.48	19,234.02	39,719.51	55,832.43	9,216.10	15,812.74	3,412.16	34,396.53	18,666.37	8,711.12	32,891.68
Standard deviation	6,127.80	10,450.83	24,596.60	30,407.09	6,147.90	13,724.68	2,325.19	17,488.69	11,268.33	4,518.40	28,443.86
Demography											
Female	56.38	32.15	50.29	44.33	45.89	48.36	43.33	49.31	50.61	50.57	51.72
Immigrant status	13.06	5.45	4.98	7.43	0.14	8.13	0.10	5.70	10.99	1.19	11.48
Parental education											
Primary or less	38.22	69.61	34.68	22.43	42.80	92.36	82.60	33.63	63.38	32.08	50.78
Secondary	43.80	21.78	39.31	42.52	47.92	3.35	13.54	40.24	23.29	58.68	25.85
Tertiary or more	17.98	8.61	26.01	35.05	9.28	4.29	3.86	26.13	13.33	9.23	23.37
Parental occupation											
Elementary	10.26	9.75	3.04	1.67	5.98	9.01	10.64	-	16.27	10.05	5.89
Skilled workers	54.11	60.78	49.31	46.90	72.13	74.28	76.50	-	54.61	58.37	48.69
Professionals	35.63	29.47	47.65	51.43	21.89	16.71	12.85	-	29.13	31.57	45.42
y^{SW}											
Sample size	5,535	3,366	4,310	2,237	12,846	4,895	4,830	2,353	3,881	6,021	5,026
Outcome											
Average	7,654.87	18,854.79	39,099.99	54,869.17	9,194.14	15,424.33	3,360.98	33,779.61	18,445.27	8,550.42	31,918.69
Standard deviation	5,156.52	8,820.54	22,657.05	28,056.89	5,141.67	12,019.38	2,115.82	16,015.28	10,304.87	3,834.45	24,514.06
Demography											
Female	55.47	48.51	52.95	45.64	53.25	52.36	47.72	51.17	53.54	53.68	54.93
Immigrant status	13.39	5.59	5.27	7.47	0.16	7.56	0.10	6.33	11.62	1.16	10.60
Parental education											
Primary or less	41.77	72.73	35.73	22.98	49.34	92.87	84.39	34.25	65.89	35.66	51.73
Secondary	41.82	19.88	38.77	42.47	43.29	3.15	12.17	40.46	21.98	55.90	25.63
Tertiary or more	16.40	7.40	25.50	34.56	7.37	3.98	3.44	25.29	12.14	8.44	22.64
Parental occupation											
Elementary	11.53	11.53	3.36	1.97	6.90	9.36	11.28	-	17.39	11.73	6.33
Skilled workers	55.57	62.24	49.77	47.70	74.47	75.10	76.81	-	55.48	59.08	49.50
Professionals	32.90	26.23	46.87	50.34	18.63	15.55	11.90	-	27.13	29.20	44.17

Note: This table shows summary statistics for each sample and country. Samples are referenced by the notation employed throughout this section: y^H refers to household income, y^{HE} to household income but keeping only heads in the sample, y^W to labor earnings, and y^{SW} to labor earnings using a selection model into employment. Aside from sample size, the average value of the outcome and its standard deviation, numbers refer to the percentage of individuals with the referenced characteristics. In the case of Sweden we are forced to exclude the circumstance parental occupation due to the very small number of respondents. Cross-sectional files of the EU-SILC database.

Hungary, Island, Luxembourg, Portugal, Romania, Slovakia and the United Kingdom). And that is also the situation with family background, since in general we find a larger proportion of individuals with the least privileged background (as measured by parental education and occupation) in the sample of y^{SW} compared to y^{W}. Of course, it is those who face more difficulties to enter the labor maker—such as women, immigrants, and people coming from less privileged backgrounds—that we would expect to appear in larger numbers in the samples with y^{SW}, compared to y^{W}'s.

To end this comparison of samples we are going to contrast the two levels of analysis discussed in the two paragraphs above. About the sample size, with y^{H} it is in every case the largest. On the other end, with y^{HE} the sample size is almost always the smallest (with the exceptions of Denmark, Finland and Sweden, in which the sample with y^{W} is[27]). Hence, broadly speaking, $N^{H} > N^{SW} > N^{W} > N^{HE}$, which represent respectively the sample size of y^{H}, y^{SW}, y^{W} and y^{HE}. Regarding mean outcome, with aggregates focusing on the individual (y^{W} and y^{SW}) we find levels about twice as high than with aggregates that look at households (y^{H} and y^{HE}). Standard errors with y^{H} are around half of what we find with y^{W}, and in some cases less than half. Therefore, recalling the discussion above, generally we have that $SE^{W} > SE^{SW} > SE^{HE} \approx SE^{H}$, which refers to the standard errors of y^{W}, y^{SW}, y^{HE} and y^{H}, respectively. About gender, the proportion of females is largest in the y^{H} sample (with the exception of Estonia, Lithuania and Latvia, where is highest in y^{W}), but not far from what we find with y^{SW}. The lowest proportion by far appears with y^{HE}. With respect to immigration, the samples of aggregates at the household level have in general a larger proportion of immigrants, but the difference is small and even sometimes indiscernible, such as in Bulgaria, Czech Republic, Hungary, Poland, Romania and Slovakia (countries of which immigrant population is very small). Finally, regarding parental education and occupation we find no major differences across the samples of the four aggregates. If anything, the y^{H}'s samples have in most cases a bigger proportion of individuals coming from less privileged backgrounds (as measured by parental education and occupation) with respect to y^{W}'s samples, which is in line with what could have been expected, provided we only consider individuals with a minimum level of labor market attachment y^{W}'s samples (see section 4.1 for a full description of the data).

Summing up, we find that the main differences across the samples of our four aggregates are their size, the mean income and its standard error, and the presence of females. Divergences

[27]It might come as a surprise that there are more household heads (sample size with y^{HE}) than people are in the labor market (sample of y^{W}). There are two reasons for this: on the one hand, the data source we employ, the EU-SILC database, collects information on family background from only one person in each household, who is usually the head, in a small group of countries (namely Denmark, Finland, Iceland, the Netherlands, Norway and Sweden—however, in all other countries this information is collected from various members of each household). Hence, if researchers want to account for parental education and occupation in these few countries they are forced to keep only one individual per household, who is usually the head, artificially lowering the sample size with y^{W} and evening it to the one with y^{HE}. On the other hand, the kind of employment we consider with y^{W} is employees or self-employed who worked full- or part-time during at least 7 months in the reference period, in order to study individuals with a minimum level of attachment to the labor market. This is, many individuals who are in the labor market, but just intermittently, are not considered in the sample of y^{W}. This and all other details about the data we employ are discussed in extenso in section 4.1.

relating to immigration and parental background do not seem substantial. Moreover, the discrepancies in mean income respond principally to the level of the aggregate—household or individual—rather than to the aggregate itself. In fact, as we will later verify, the sample composition does not play a major role in shaping the differences between the IOP estimates obtained with our four income aggregates, even though its influence is not negligible. We will return to this in section 4.3, where we employ one artificially identical sample for all aggregates.

3.3. The inequality index

The third source of discrepancy between IOP estimates obtained with different income aggregates is their interplay with the inequality index employed. Recall that the most commonly employed measure of inequality in the IOP literature is the mean log deviation (see section 2), which is defined as in eq. (6), or alternatively:

$$\mathrm{MLD}(X) = \ln \bar{x} - \overline{\ln x_i}. \tag{13}$$

Suppose $X \in \mathbb{R}^n_+$ is a distribution of individual income, and $x_i \in X$ with $i \in I$ representing individuals. We can transform X to per capita household income by taking the average in each residential unit, and represent it by X^H, with $x_j \in X^H$, where $j \in J$ are households. Then, to calculate the MLD of X^H we would substitute x_i by x_j in eq. (13). Note, however, that $\bar{x}$ takes the same value in both cases. Hence, $\mathrm{MLD}(X^H) = \ln \bar{x} - \overline{\ln x_j}$. Now, informed by the strict concavity of the logarithmic function we know that $\overline{\ln x_i} < \overline{\ln x_j}$, and consequently $\mathrm{MLD}(X^H)$ $< \mathrm{MLD}(X)$.

Returning to our previous notation, the fact just described entails that, abstracting from the other effects discussed, $I(\tilde{\mu}^H) < I(\tilde{\mu})$, where $\tilde{\mu}^H$ is the smoothed counterfactual of household income, y^H_j, and $\tilde{\mu}$ is the smoothed counterfactual of individual income y_i.

4. Empirical demonstration

With the aim of illustrating the importance of the income aggregate, in this section we estimate IOP using four different income concepts: per capita disposable household income, which will be denoted by y^H, per capita disposable household income but keeping only household heads in the sample (y^{HE}), gross labor earnings (y^W), and gross labor earnings with a Heckman selection model into employment (y^{SW}). We will first, in section 4.1, describe the data and method employed, to subsequently discuss the results in section 4.2 and the robustness tests in section 4.3.

4.1. Data and methodology

We make use of the cross-sectional files of the European Survey of Income and Living Conditions, 2019 revision. This is a well-known and researched database for the study of inequality, poverty, and social exclusion that offers harmonized data on income and circumstances at the individual and household level. It has been conducted yearly since 2003 for up to 31 European countries in its most recent waves. However, the EU-SILC has information about parental features in two waves only—2004 and 2010—and since some data to account for family background is generally required to estimate IOP, we are forced to use one these two waves.[28] We decide to use the wave referring to 2010, because it is more recent and covers more countries.[29]

We will employ four different samples, one for each income aggregate. The rationale is that each aggregate requires its own data processing, and a valid comparison would take this into account. For instance, we would drop all observations with missing values of "labor earnings" if we are estimating IOP using precisely "labor earnings" as the outcome of interest, but would not do so if we are considering "household income". Likewise, we would not drop observations with missing or non-positive income if we are going to employ a selection model into employment. For that reason we keep four different samples throughout the analysis, that nonetheless are identical in everything other than the idiosyncratic data processing requirements of each aggregate.

Common to the four samples is the following. We restrict them to people aged 30 to 59 to account for life-cycle effects, which is a common procedure in the literature (e.g. Marrero and Rodríguez 2012). We choose 30 years as the lower limit because income at this age is a good predictor of long-term earning potential (Chetty et al. 2014), and 59 as the upper limit because the EU-SILC does not collect information on family background for older individuals. We also remove from our sample all observations with missing values in any circumstance, and in addition, we cap very high values in each income distribution, in particular by replacing all values above the 99th percentile with the value at that percentile.

The construction of the four income aggregates will be detailed now. The per capita disposable household income, y^H, is calculated as "total disposable household income" (variable HY020) divided by "household size" (variable HX040).[30] In the specific sample of this aggregate we drop non-positive or missing observations in variables HY020 or HX040.

Per capita disposable household income but keeping only household heads in the sample, y^{HE}, is defined in exactly the same way as y^H, the only difference being that the sample employed is limited to household heads only. The head of the household is identified by the

[28]The importance of socioeconomic origin in determining economic and social outcomes has been explored extensively by the literature on intergenerational mobility (e.g. Corak 2013). In the context of IOP see Ferreira and Peragine (2016), table 25.8, for a list of the circumstances employed by eight studies, all of which include parental background.

[29]We include in our analysis all countries available, which in 2010 are Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland and the United Kingdom.

[30]Total disposable household income is defined as the sum of all household members' gross personal income components, plus gross income components at household level, minus regular taxes on wealth, regular inter-household transfers paid and tax on income and social insurance contributions. Gross means that neither taxes nor social contributions have been deducted at source. Household size corresponds to the number of members in the household.

variable "Person responsible for the accommodation" (HB080).[31] As with y^H, we also drop non-positive or missing observations in variables HY020 or HX040.

Gross labor earnings, y^W, is defined as "Employee gross cash or near cash income" (variable PY010G), plus "Gross non-cash employee income" (PY020G), plus "Cash benefits or losses from self-employment (Gross)" (PY050G). We exclude "Value of goods produced by own-consumption (Gross)" (PY070G) because it is not available in all countries. Gross means that neither taxes nor social contributions have been deducted at source, which we opt to consider because it allows to have an account of how the market rewards each type, excluding state intervention. The sample with this aggregate is limited to employees or self-employed who worked full- or part-time during at least 7 months during the reference period,[32] and observations with non-positive or missing values in all of the variables PY010G, PY020G or PY050G are dropped. We would like to account for underemployment, by considering the willingness of part-time workers to work more hours. Regrettably though, the variable of the EU-SILC that could allow us to do so, "Reason for working less than 30 hours" (PL120), has only a small number of observations that would remarkably reduce our sample size.

Finally, the aggregate gross labor earnings with a Heckman selection model into employment, y^{SW}, is defined in exactly the same way as y^W, but we impute expected incomes to people out of employment. The kind of employment modeled is also employees or self-employed who worked full- or part-time during at least 7 months during the reference period. The Heckman correction method and the actual selection model employed are described in the appendix. From the sample of this aggregate we only drop observations with missing values in the variables included in the selection model, while we naturally keep observations with non-positive or missing values in the earnings variables (PY010G, PY020G or PY050G).

Regarding circumstances, we consider gender, immigrant status, parental education and parental occupation, which are commonly used in the literature.[33] As gender we consider binary gender, since it is the information available in the database.

With respect to immigrant status, we differentiate between individuals born in the country of residence and those born outside. In spite of being frequently used, immigrant status is not considered a circumstance by some researchers. For that reason we would like to justify this choice. Although migration clearly falls within the control of individuals (except in extreme situations such as famines, wars, political prosecution or natural disasters), we believe it can be considered a circumstance on the basis that the country where we live largely determines our income (Milanovic 2015), and unless we emigrate, the country where we live is the country where we were born, what is outside our control. Moving to a country with a more favorable

[31]Defined as the person owning or renting the accommodation. In case two persons share responsibility, the oldest one is registered as the head.

[32]We use the variables "Number of months spent at full-time work as employee" (PL073), "Number of months spent at part-time work as employee" (PL074), "Number of months spent at full-time work as self-employed (including family worker)" (PL075), and "Number of months spent at part-time work as self-employed (including family worker)" (PL076).

[33]See Ferreira and Peragine (2016), table 25.8, for a list of circumstances employed in eight studies.

income distribution may return a gain, but at a cost in terms of effort that those already born in such country do not have to assume. Furthermore, natives do not face possible discrimination due to their national origin.

Regarding paternal education, we group individuals according to the highest educational level attained by any of their parents: pre-primary, primary or lower secondary education (levels 0, 1, and 2 of ISCED-97), upper secondary and post-secondary non-tertiary education (levels 3 and 4 of ISCED-97), and first and second stage of tertiary education (levels 5 and 6 of ISCED-97). This is, we distinguish three levels of parental education.

Finally, regarding parental occupation, we also create three groups of individuals according to the highest job category of their parents, which correspond to elementary occupations (group 9 of ISCO-88), plant and machine operators and assemblers, craft and related trades workers, skilled agricultural and fishery workers, service workers and shop and market sales workers, and clerks (groups 8, 7, 6, 5 and 4 of ISCO-88), and technicians and associate professionals, professionals, and legislators, senior officials and managers (groups 3, 2 and 1 of ISCO-88).

However, in the case of Sweden we are forced to exclude the circumstance parental occupation due to the very small number of respondents, what would drive the number of observations per type unacceptably low. In addition, in Denmark, Finland, Iceland, Netherlands, Norway, Slovenia, and Sweden information on parental features is collected for one individual per household only, either the head or the spouse. This means that to account for these parental features we can only analyze one individual per household in these countries, in all our four samples. This is the reason why in table 1 these countries have missing information, as heterogeneity within households cannot be tested if we have information of one member only. Notice that, in the mentioned countries, this fact prevents us from truly estimating IOP using y^H with the EU-SILC database.

Therefore, we have up to 36 types, product of 2 genders $\times$ 2 geographical origins $\times$ 3 levels of parental education $\times$ 3 levels of parental occupation. However, the number of types falls shorter than 36 in some datasets, because some combinations of circumstances are infrequent and do not appear in the data. To prevent the possible bias produced by types with very few observations that may contain extreme values (see Brunori, Peragine, et al. 2019), we retain only those types with at least a minimum number of observations, which we set to 10.

For our analyses it is advised to make use of the bootstrap (see e.g. Efron and Tibshirani 1993). When appropriate we will perform estimations with 1,000 replications, and following Andreoli and Fusco (2017) we stratify by region[34] (see also Goedemé 2013).

Finally, regarding the empirical methods to estimate IOP, we apply the parametric approach proposed by Bourguignon et al. (2007) and Ferreira and Gignoux (2011). We employ this method because, as detailed in section 2, we are interested in decomposing the contribution of individual circumstances to total IOP by means of the Shapley-value. The inequality measure

[34] Variable DB040, which refers to the region of the residence of the household at the date of interview, classified according to NUTS-08.

employed, as was also detailed in section 2, is the mean log deviation.

4.2. Results

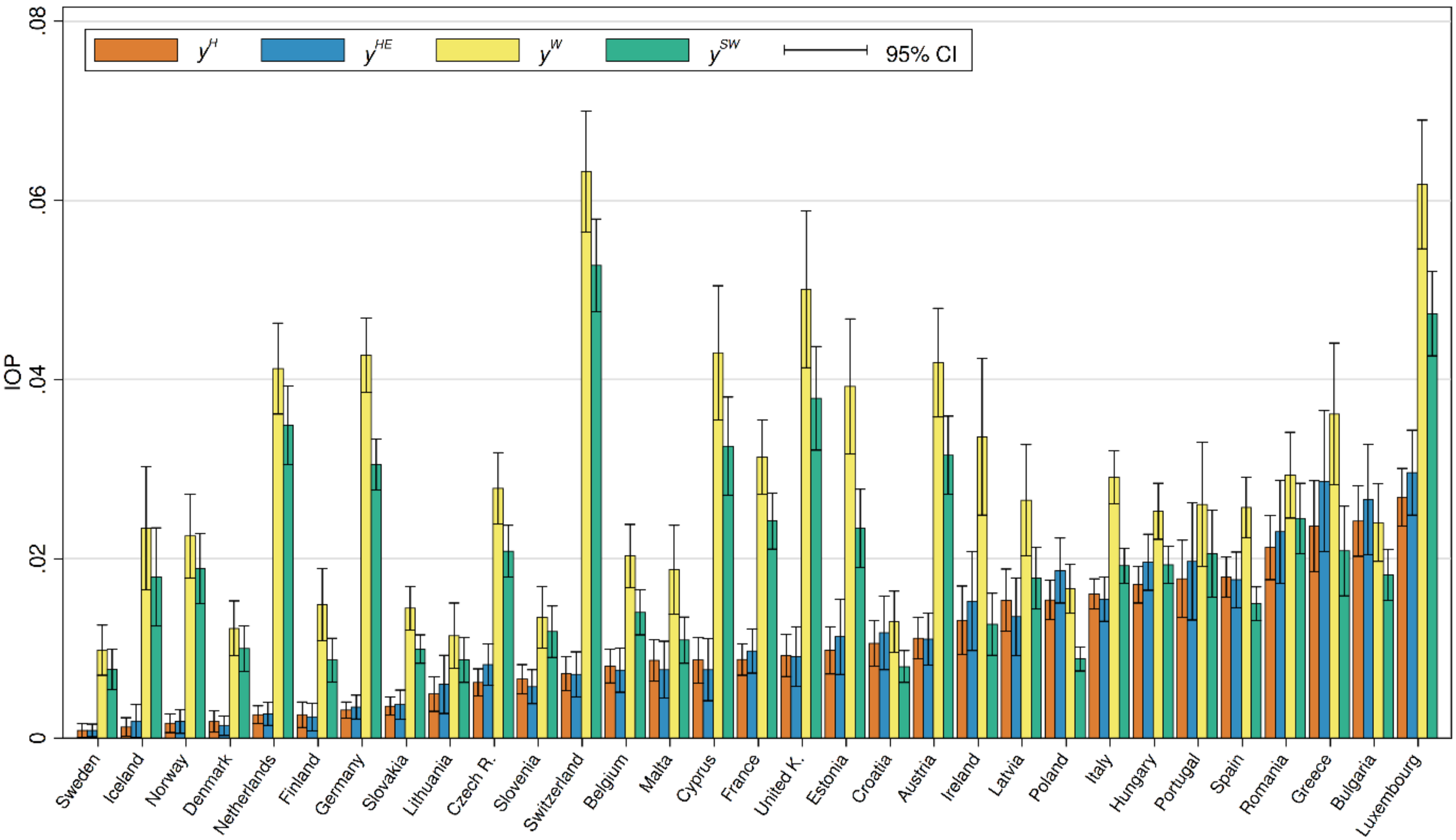
We will now test empirically the effects described in section 3 by comparing IOP estimates obtained with our four income aggregates. First we study the levels of IOP, then look at the resultant rankings, and lastly decompose total IOP into the contribution of each particular circumstance. Grosso modo, we find that IOP tends to be lower when we focus on the household level, almost entirely due to the fact that this level of analysis virtually nullifies the contribution of the circumstances gender, in line with what was discussed in section 3.

Figure 1 displays the level of absolute IOP by country, with bootstrapped confidence intervals in brackets. As we readily see, IOP differs remarkably depending on the income concept considered. Estimates with y^H and y^{HE} tend to be lower than those with y^{SW}, and specially with y^W, even though there are exceptions: in Poland and Bulgaria (countries ordered as in fig. 1) estimates with y^W are as low or lower than with y^H or y^{HE}; and in Croatia, Ireland, Poland, Hungary, Spain, Greece and Bulgaria (which are about a fifth of the countries in the analysis) estimates with y^{SW} are also as low or lower than y^H or y^{HE}. This highlights that while we can observe some regularity, it should not be assumed that using aggregates at the individual level will necessarily return higher levels of IOP. Other tendencies to notice is that estimates with y^H and y^{HE} are generally similar, and that with y^{SW} IOP is always lower than with y^W. In conclusion, fig. 1 clearly shows that IOP estimates are sensitive to the income aggregate, and that the channels described in section 3 can indeed move the results in any direction.

Aside from comparisons relating to the income aggregate considered, a notable feature to appreciate in fig. 1 is how low the level of IOP is in some countries, and how much higher is in others—compare for instance Norway and Greece with aggregate y^H, or Switzerland and Croatia with y^{SW}. As we discussed in section 2, the interpretation of estimates as lower-bounds leads to values that might appear to be "too low". Regrettably this is presently a shortcoming of the literature, common to all applications of the measurement methods we employ.

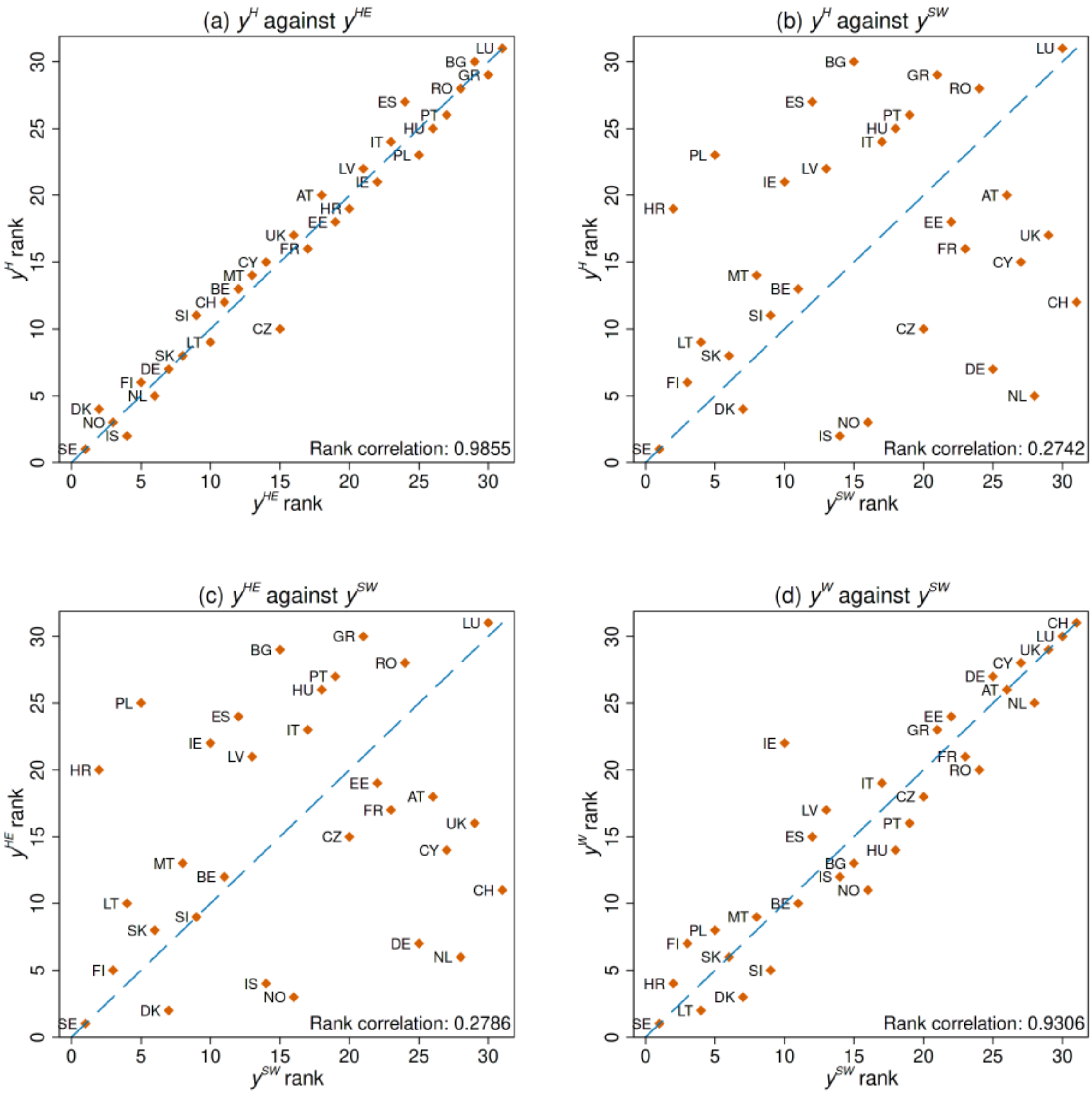
Another way of assessing the extent to which the choice of the aggregate affects IOP estimates is by looking at ranks. Figure 2 compares country rankings according to absolute IOP calculated with our four income aggregates. It displays four scatter plots, of which axes represent the rank of the 31 countries in the analysis with respect to their level of absolute IOP, from smaller to bigger. Panel (a) contains the first scatter plot, which compares the ranking according to IOP with y^H against that with y^{HE}. Panel (b) contrasts the ranking of y^H against the one with y^{SW}, (c) y^{HE} against y^{SW}, and (d) y^W against y^{SW}. Hence, panels (a) and (d) compares rankings with aggregates at the same level of analysis, either the household or the individual. On the contrary, panels (b) and (c) compare ranks with aggregates at both levels. Unsurprisingly, panels (a) and (d) show similar rankings, with Spearman's rank correlations around 0.95 in both cases. However, important differences arise in panels (b) and (c). The

Figure 1: Absolute IOP comparison



Note: Absolute IOP by country. There are four bars per country, representing the level of IOP with each one of our four income aggregates. Countries are ordered with respect to y^H. Bootstrapped confidence intervals stratified by region shown in brackets (1,000 replications). Estimates refer to the year 2010. Cross-sectional files of the EU-SILC database.

Figure 2: Comparison of country rankings with respect to absolute IOP



Note: This graph contrasts country rankings according to absolute IOP estimates obtained using our four income aggregates. Each panel contrasts two particular country ranks, and the Spearman's rank correlation is reported at the bottom of each panel. Estimates refer to the year 2010. Cross-sectional files of the EU-SILC database.

rank correlation of IOP estimates obtained with an aggregate at the individual level, y^{SW}, and estimates using aggregates at the household level, both y^H (panel (b)) and y^{HE} (panel (c)), is below 0.3, with a significance level (not reported) indicating that we cannot reject the hypothesis of independence.[35] Indeed, it appears that when focusing on the household level

[35] Considering y^W instead of y^{SW} here does not alter the conclusions, nor it does comparing relative levels of IOP instead of absolute ones.

Table 3: Averaged Shapley-value decomposition of IOP

	y^H	y^{HE}	y^W	y^{SW}
Gender	0.0002	0.0003	0.0152	0.0098
Immigrant status	0.0017	0.0017	0.0019	0.0011
Parental education	0.0048	0.0050	0.0066	0.0057
Parental occupation	0.0040	0.0042	0.0054	0.0043
Total	0.0107	0.0113	0.0290	0.0209
Gender (%)	1.9	3.0	52.3	47.0
Immigrant status (%)	15.9	14.9	6.4	5.4
Parental education (%)	45.1	44.6	22.8	27.1
Parental occupation (%)	37.1	37.4	18.5	20.5
Total	100.0	100.0	100.0	100.0

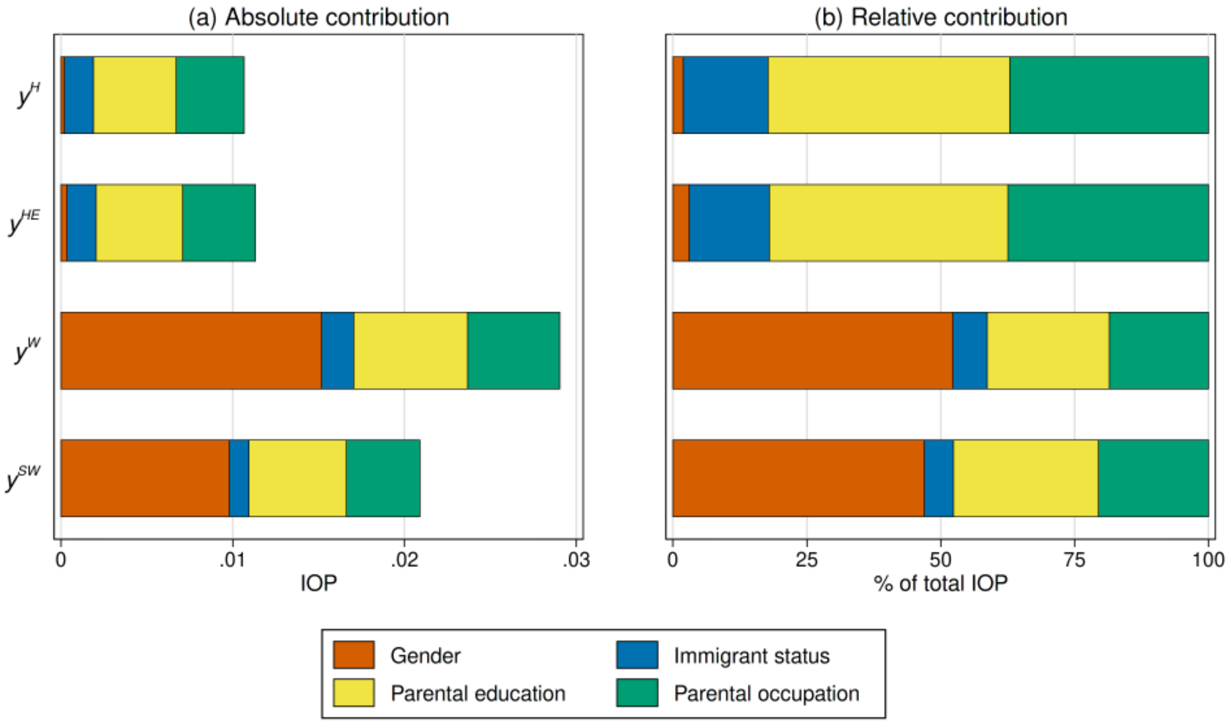
Note: Contribution of each circumstance to total IOP, averaged across 30 countries (Sweden is excluded because it lacks information on parental occupation), using our four income aggregates. The first five rows show mean absolute contributions, the following five display mean relative contributions to total IOP. The same information of this table can be found in fig. 3. Estimates refer to the year 2010. Cross-sectional files of the EU-SILC database.

we are measuring a different phenomena than when looking at the individual, underlining the importance of the income aggregate as a methodological feature.

We have seen that IOP estimates are sensitive to the level of analysis, but not so much to the particular aggregate employed. Now, to understand what is driving the differences it is useful to study how the role of each circumstance changes with each aggregate. To this end we will perform Shapley-value decompositions, which allow to decompose total IOP into the contribution of each particular circumstance (Shorrocks 2013; see also Ferreira, Gignoux, and Aran 2011). A convenient property of this technique is that the individual contributions of all circumstances add up to exactly the total amount of IOP. However, it shall be noted that these decompositions must be taken as approximations only. Table 3 shows the contribution to total IOP of each circumstance, averaged across 30 countries (Sweden is excluded because it lacks information on one circumstance, namely parental occupation, so its decomposition is not comparable), both in absolute and relative terms. Additionally, fig. 3 shows the same information in a more graphical way.

The Shapley-value decomposition of IOP suggests that the difference between the household and individual levels of analysis is largely explained by the contribution of the circumstance gender. Its contribution is almost negligible when choosing an income aggregate at the household level, but it becomes by far the most important circumstance when we focus on individuals. Interestingly enough, the absolute contributions of all other circumstances—immigrant status, parental education, and parental occupation—are broadly similar across all income aggregates (see the first five rows of table 3), stressing that the sensitivity of results to different levels of

Figure 3: Averaged Shapley-value decomposition of IOP



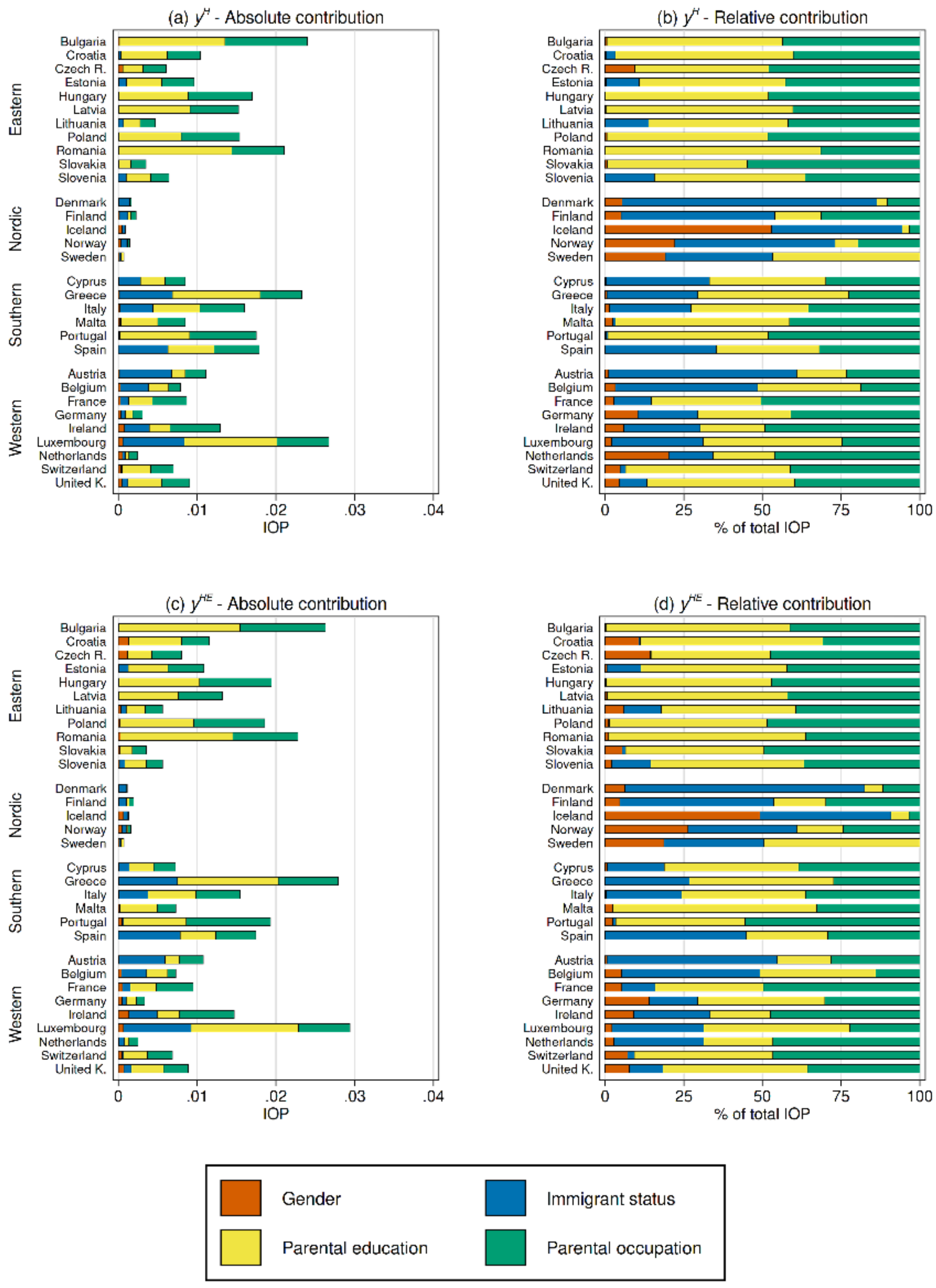
Note: Contribution of each circumstance to total IOP, averaged across 30 countries (Sweden is excluded because it lacks information on parental occupation), using our four income aggregates. Panel (a) refers to mean absolute contributions, panel (b) shows mean relative ones. The same information of this graph can be found in table 3. Estimates refer to the year 2010. Cross-sectional files of the EU-SILC database.

analysis responds to the role of gender. Indeed, the decomposition procedure suggests that the effect of gender is virtually nullified when choosing an income aggregate at the household level, product of the bias described in section 3. The decomposition of estimates with y^H and y^{HE} returns remarkably similar results, and while the contributions of the circumstances do differ between estimates with y^W and y^{SW}, they do so to a small extent only (recall fig. 2, which shows very similar country rankings with y^W and y^{SW}).

However, averaging across countries as in table 3 and fig. 3 may conceal variability worth considering. Analyzing the role of circumstances in a case by case fashion leads to figs. 4 and 5, where countries have been grouped by geographical location. The first thing to notice is that grouping by region appears to be convenient because there is some consistency of the role of circumstances within these "natural" groups. A second takeaway is that, again, there are no major differences between aggregates at the same level of analysis, i.e., between y^H and y^{HE} or between y^W and y^{SW}. This seems to be a persistent finding in both the averaged and the case by case perspectives, and of course goes in line with what the study of rankings suggested (see fig. 2).

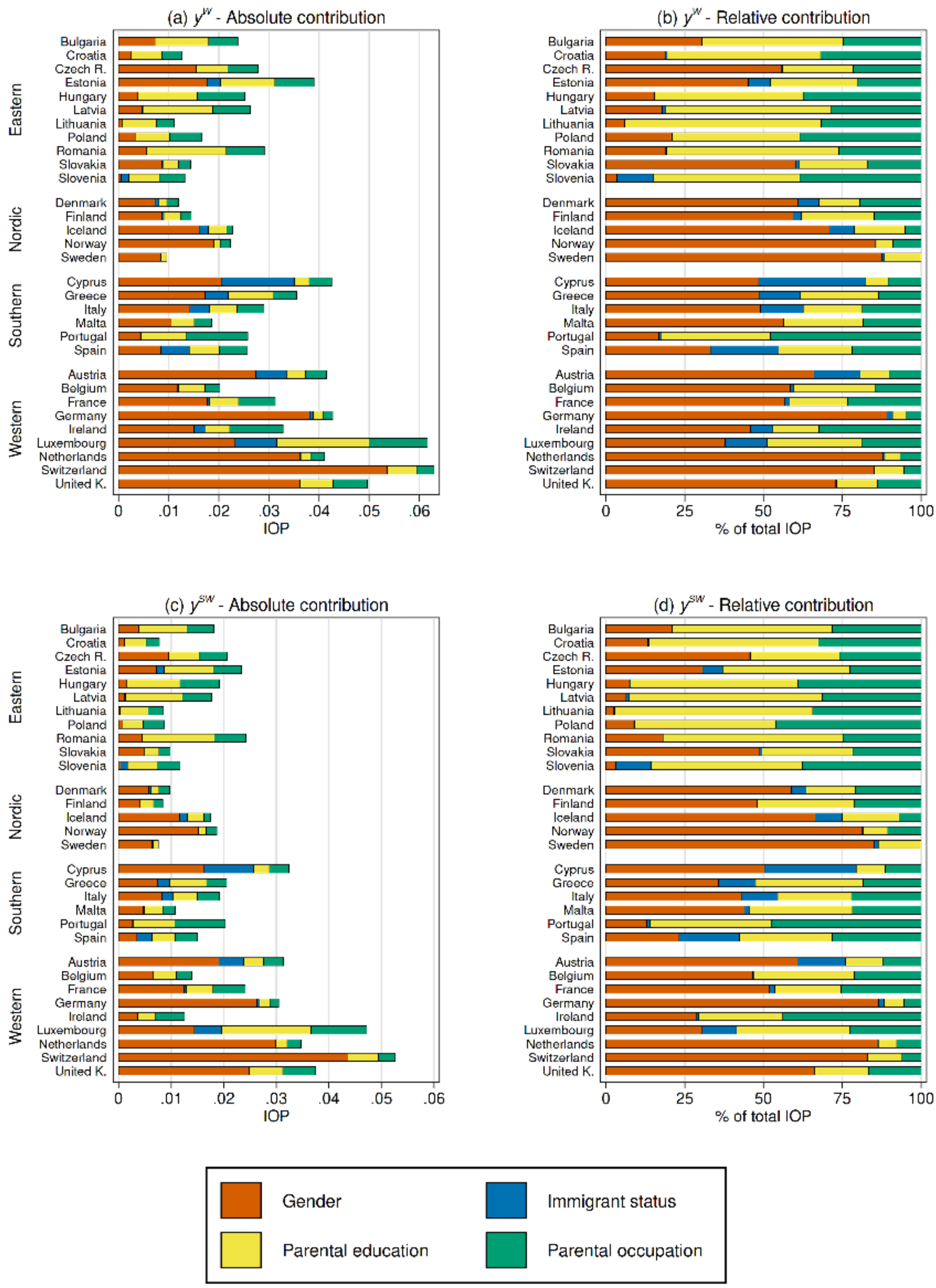
Let us consider fig. 4 now, where IOP estimated with aggregates at the household level is decomposed. This graph displays the contribution of each circumstance to total IOP by

Figure 4: Shapley-value decomposition of IOP at the household level (y^H and y^{HE})



Note: Contribution of each circumstance to total IOP by country, using our two income aggregates at the household level. Panels (a) and (b) refer to the aggregate y^H, while panels (c) and (d) relate to y^{HE}. Additionally, panels (a) and (c) show absolute contributions, while panels (b) and (d) display relative ones. In Sweden the circumstance parental occupation is not available. Estimates refer to the year 2010. Cross-sectional files of the EU-SILC database.

Figure 5: Shapley-value decomposition of IOP at the individual level (y^W and y^{SW})



Note: Contribution of each circumstance to total IOP by country, using our two income aggregates at the individual level. Panels (a) and (b) refer to the aggregate y^W, while panels (c) and (d) relate to y^{SW}. Additionally, panels (a) and (c) show absolute contributions, while panels (b) and (d) display relative ones. In Sweden the circumstance parental occupation is not available. Estimates refer to the year 2010. Cross-sectional files of the EU-SILC database.

country. Panels (a) and (b) refer to the aggregate y^H, while panels (c) and (d) relate to y^{HE}. Additionally, panels (a) and (c) show absolute contributions, while panels (b) and (d) display relative ones. In Eastern countries IOP is high, and family background appears to be the most important feature, while immigrant status and gender are nearly irrelevant in most cases. Nordic countries have by large the lowest level of IOP, and family background relinquishes its relevance to immigrant status and gender—see panels (b) and (d). In this group of countries gender appears to play a much bigger role than in the rest, although it does so only in relative terms. Considered as the amount of inequality due to gender, i.e. absolute contribution—panels (a) and (c)—, in the Nordic countries IOP due to gender is not particularly large. The importance of this circumstance in relative terms—panels (b) and (d)—is due to the small amount of total absolute IOP in these countries.[36] In Southern economies family background accounts for most of IOP too, although immigration is important as well (with the exception of Malta and Portugal). Again, gender plays little to no role. Finally, in Western European economies both family background and immigrant status are important and add up to most of IOP, although gender plays some role too[37].

If we look at fig. 5 we find a contrasting picture—one in which gender is the most important circumstance. In Eastern countries gender now contributes a relevant share of total IOP (with the remarkable exception of Lithuania and Slovenia), but family background remains an important source of inequality too. Nordic economies have again the lowest level of IOP, but immigrant status is not as important anymore, and instead gender (specially in Norway and Sweden) accounts for most of total IOP. In the south of Europe gender becomes a very important circumstance, even though immigrant status retains its relevance (with the exception of Malta and Portugal). Finally, it is in Western countries where we find the highest levels of IOP—panels (a) and (c)—. In all of these societies gender is the single most important circumstance, accounting in all but Ireland and Luxembourg for more than 50% of total IOP.

Summing up, the main conclusion we can extract from studying the role of circumstances in a case by case fashion concurs with what we had concluded from the study of averages in table 3 and fig. 3. This is, that using an aggregate at the household level almost nullifies the role of gender, and that there are no major differences between estimates with aggregates at the same level of analysis, i.e., between y^H and y^{HE} nor between y^W and y^{SW}.

4.3. Robustness

In this section we test if our previous results are robust to some methodological variations. These alternative approaches are three, to wit a) using the same sample across income aggre-

[36] Nevertheless there is an important caveat to remember here: with y^H as the outcome of interest—panels (a) and (b)—we must be careful interpreting the data from the Nordic countries and the Netherlands. For these countries the EU-SILC database contains information on family background from one member per household only, and hence it is not possible to truly estimate IOP using y^H, provided we want to account for socioeconomic background (we discussed this issue in section 4.1 and table 1). However, this caveat does not apply with y^{HE}—panels (c) and (d).

[37] Again, we must interpret results with y^H in the Netherlands with caution; see the previous footnote.

Table 4: Spearman's rank correlations with alternative methodologies

	y^H with y^{HE}	y^H with y^{SW}	y^{HE} with y^{SW}	y^W with y^{SW}
Equal samples	-	0.2613	-	-
Gini	0.9762	0.1919	0.2165	0.9323
Equivalent income	0.9722	0.2520	0.2754	0.9351

Note: This table shows Spearman's rank correlations between IOP estimates obtained using our four income aggregates. Each column corresponds to the correlation between two particular aggregates, and each row shows the statistic with a particular methodological variation. Estimates refer to the year 2010. Cross-sectional files of the EU-SILC database.

gates, b) using the Gini index instead of the mean log deviation, and c) considering equivalized household income instead of per capita. We will conclude that our previous findings hold, but we also obtain further evidence suggesting that the bias related to the circumstance gender described in section 3.1 is the main source of discrepancy across levels of analysis.

Consider table 4, which shows the rank correlations of IOP estimates obtained with our four income aggregates. This table is similar to fig. 2, but showing only Spearman correlations. Each column corresponds to correlations between results with two particular aggregates, and each row displays the results with one of the three methodological variations included in the robustness test. The first row, referring to using the same sample across all income aggregates, responds to the fact that each aggregate requires its own data processing, what naturally gives rise to different samples (see section 3.2). Therefore this test is aimed at checking to what extent this factor explains the differences in estimates. This row displays information in one column only—that referring to the correlation of y^H with y^{SW}—, because the only difference between y^H and y^{HE}, and also between y^W and y^{SW}, is precisely the sample composition, and hence artificially using a unique sample returns equal results with y^{HE} and y^H, and with y^{SW} and y^W, so we avoid reporting them. In conclusion, we see that the rankings do not change much, since the rank correlation remains below 0.3.

The second row of table 4 relates to the use of Gini as the inequality index, instead of the mean log deviation. Recall from section 3.3 that the particular characteristics of each index may affect the results, and also from section 2 that, more generally, every inequality index implies a normative choice (Atkinson 1970). This test returns rank correlations that offer a similar picture to the one obtained before, suggesting that the inequality index is not a major factor behind the differences across levels of analysis.

The third and final row of table 4 shows results with a different definition of household income. We use equivalized income instead of per capita income, because as discussed in section 3 this is a common approach too. Also, the conceptual analysis we have conducted in section 3.1 does not apply to it. Nevertheless we once again conclude that this test does not seem to entail a substantial difference with respect to our main methodology.

In sum, the rank correlations of table 4 suggest that none of the considered methodological

Table 5: Averaged Shapley-value decomposition of IOP with alternative methodologies

	y^H	y^{HE}	y^W	y^{SW}
Equal samples				
Gender	0.0002	-	-	0.0099
Immigrant status	0.0015	-	-	0.0011
Parental education	0.0047	-	-	0.0057
Parental occupation	0.0038	-	-	0.0043
Total	0.0102	-	-	0.0210
Gini				
Gender	0.0048	0.0060	0.0568	0.0438
Immigrant status	0.0089	0.0086	0.0067	0.0053
Parental education	0.0306	0.0313	0.0347	0.0327
Parental occupation	0.0282	0.0287	0.0294	0.0273
Total	0.0725	0.0746	0.1277	0.1091
Equivalent income				
Gender	0.0001	0.0010	0.0152	0.0098
Immigrant status	0.0015	0.0015	0.0019	0.0011
Parental education	0.0051	0.0052	0.0066	0.0057
Parental occupation	0.0040	0.0042	0.0054	0.0043
Total	0.0107	0.0118	0.0290	0.0209

Note: Contribution of each circumstance to total IOP, averaged across 30 countries (Sweden is excluded because it lacks information on parental occupation), with our four income aggregates and three methodological variations. Each alternative methodology is indicated in italics. The five rows of each specification show mean absolute contributions. Estimates refer to the year 2010. Cross-sectional files of the EU-SILC database.

features is responsible for much of the differences produced by the choice of the income aggregate. This provides further support to the hypothesis that the contribution of the circumstance gender is the main factor. To collect more evidence on this question we will analyze now the contribution of each individual circumstance to total IOP.

The second part of this robustness test revisits the decomposition of IOP. Just as we have done in the last part of section 4.2 we will now decompose IOP by means of the Shapley-value, but do so with estimates obtained following the alternative methodologies discussed above. The results are shown in table 5. This table presents the absolute contribution of each circumstance to overall IOP in each one of the three alternative settings we have chosen for the robustness test. The first five rows relate to using the same sample across all income aggregates, and as we explained when discussing table 4, with this test we report results from two aggregates only. A feature to notice is that the decomposition of IOP using y^{SW} has barely changed with respect to table 3, and that is because the procedure to employ equal samples with y^H and y^{SW} mostly consists og dropping observations from y^H's sample, since it is the biggest of the two (see section 3.2). Nonetheless, changes in the decomposition of IOP with y^H are almost negligible, what again belittles the importance of the samples' differences in explaining the role of the income aggregate.

The next five rows display results using the Gini index instead of the mean log deviation. Values are larger by construction, and obviously they are not directly comparable to the ones obtained using the mean log deviation. It is interesting to see that differences between total IOP with the four income aggregates are smaller now (for instance, total IOP with y^H is around 65% of total IOP with y^{SW}, while it was 50% using the mean log deviation). However we still find that the main difference across estimates is the contribution of gender, which remains much larger with aggregates at the individual level.

The last five rows of table 5 show results with household equivalent income, instead of per capita. Therefore, with respect to table 3 we only see changes in estimates with aggregates at the household level, y^H and y^{HE}—this robustness check does not affect results with y^W and y^{SW}. What we find is that choosing per capita or equivalent household income does not entail a meaningful difference, just as we advanced in section 3.

Summing up, all of the robustness tests performed point in the same direction: the contribution of the circumstance gender explains the lion's share of the differences between IOP estimates obtained with our four income aggregates, while the sample composition and the inequality index are of secondary importance.

5. Concluding remarks

We have seen that IOP estimates are highly sensitive to the choice of the income aggregate, and hence attention is due to this methodological aspect. Specifically, beyond the particular aggregate, the key feature is the level of analysis, either households or individuals. Considering a distribution of income at the household level returns IOP estimates that are generally lower than those obtained when we analyze individual incomes, and this difference is due almost entirely to the contribution of the circumstance gender. Indeed, gender is by far the most important source of discrepancy between IOP estimates at the household and individual level.

We have identified the mechanisms that operate behind this process, and tested their workings empirically. The results of the data analysis in section 4 match the mechanisms laid out in section 3, and also go in line with the previous literature regarding within-household redistribution and gender identity norms at play in the labor market. We conclude that, when measuring IOP, the income aggregate implies a normative choice regarding gender. In particular, if gender is considered a source of IOP an aggregate at the individual level must be employed.

This has relevant consequences. Let us discuss how this issue has affected IOP research. Brunori, Hufe, et al. (2020) propose the use of machine learning methods to reduce the risk of arbitrary selection of circumstances. In their illustration gender is considered a potential circumstance, but it is only shown as relevant—i.e., the algorithm deems it significant—in 1 of the 31 European countries for which they measure IOP. Hence, their analysis depicts gender as marginally important in the determination of IOP, even though this result is surely due to the fact that they consider household income as the outcome of interest. This is, the method they propose to avoid ad-hoc selection of circumstances is conditioned by the (arbitrary) choice of the income aggregate.

Suárez Álvarez and López Menéndez (2017) decompose total IOP in Spain and find that gender is nearly irrelevant, and no reference to the importance of the income aggregate is made. Again, from these results it could be concluded that gender is unimportant, although of course they are a product of choosing an aggregate at the household level.[38]

Another example is Equal Chances, the so-called world database on equality of opportunity and social mobility. As specified in its technical note (EqualChances 2018), even though gender is often included in the set of circumstances, in this project it is not because the outcome is defined at the household level (EqualChances 2018, p. 6). This is, the largest IOP database in the world presents gender as a dispensable dimension of IOP.

It is important to keep in mind that every income aggregate suffers from its own shortcomings and biases, and therefore they should be chosen according to the research question at hand and justified on that basis. Nonetheless, if income is the outcome of interest and gender is to be considered as a source of IOP, then an aggregate at the individual level must be considered. Indeed, focusing on households leads to downplaying the importance of gender as a source of IOP. If this field claims to embody the ideal of fairness that modern societies embrace, it better reassess this stance.

References

Andreoli, F. and A. Fusco, "The evolution of inequality of opportunity across Europe: EU-SILC evidence," in A.B. Atkinson, A.-C. Guio, and E. Marlier, eds., *Monitoring social inclusion in Europe*, 435–448, Publications Office of the European Union, Luxembourg, 2017.

Atkinson, A.B., "On the Measurement of Income Inequality," *Journal of Economic Theory,* 2, 244–263, 1970.

Becker, G.S., *A Treatise on the Family*, 1991st ed., Harvard University Press, Cambridge, 1981.

Bertrand, M., E. Kamenica, and J. Pan, "Gender Identity and Relative Income within Households," *Quarterly Journal of Economics,* 130, 571–614, 2015.

Bisin, A., E. Patacchini, T. Verdier, and Y. Zenou, "Ethnic identity and labour market outcomes of immigrants in Europe," *Economic Policy,* 26, 57–92, 2011.

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

Bourguignon, F., F.H.G. Ferreira, and M. Menéndez, "Inequality of Outcomes and Inequality of Opportunities in Brazil", 2003.

———, "Inequality of opportunity in Brazil," *Review of Income and Wealth,* 53, 585–618, 2007.

[38] However, these authors make a reference to this issue in Suárez Álvarez and López Menéndez (2018).

Brunori, P., P. Hufe, and D.G. Mahler, “The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees and Forests”, 2020.

Brunori, P., F. Palmisano, and V. Peragine, “Inequality of opportunity in sub-Saharan Africa,” *Applied Economics,* 51, 6428–6458, 2019.

Brunori, P., V. Peragine, and L. Serlenga, “Upward and downward bias when measuring inequality of opportunity,” *Social Choice and Welfare,* 52, 635–661, 2019.

Bursztyn, L., F. Thomas, and A. Pallais, “‘Acting Wife’: Marriage Market Incentives,” *American Economic Review,* 107, 3288–3319, 2017.

Checchi, D. and V. Peragine, “Inequality of opportunity in Italy,” *The Journal of Economic Inequality,* 8, 429–450, 2010.

Checchi, D., V. Peragine, and L. Serlenga, “Inequality of opportunity in Europe: Is there a role for institutions?,” in L. Cappellari, S.W. Polachek, and K. Tatsiramos, eds., *Research in Labor Economics*, vol. 43, 1–44, Emerald Group Publishing Limited, 2016.

Chetty, R., N. Hendren, P. Kline, and E. Saez, “Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States,” *Quarterly Journal of Economics,* 129, 1553–1623, 2014.

Corak, M., “Income Inequality, Equality of Opportunity, and Intergenerational Mobility,” *Journal of Economic Perspectives,* 27, 79–102, 2013.

Cowell, F.A., “Measurement of inequality,” in A.B. Atkinson and F. Bourguignon, eds., *Handbook of Income Distribution*, vol. 1, Handbook of Income Distribution, 87–166, Elsevier, 2000.

De La Rica, S., A. Glitz, and F. Ortega, *Immigration in Europe: Trends, policies, and empirical evidence*, 1st ed., vol. 1, Elsevier B.V., 2015.

Efron, B. and R. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, London, 1993.

EqualChances, “The World Database on Equality of Opportunity and Social Mobility,” *Technical note (beta version),* 2018.

Eurostat, *EU Statistics on Income and Living Conditions*, 2019.

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, 622–657, 2011.

Ferreira, F.H.G., J. Gignoux, and M. Aran, “Measuring Inequality of Opportunity with Imperfect Data: The Case of Turkey,” *Journal of Economic Inequality,* 9, 651–680, 2011.

Ferreira, F.H.G., C. Lakner, M.A. Lugo, and B. Özler, “Inequality of Opportunity and Economic Growth: How Much Can Cross-Country Regressions Really Tell Us?,” *Review of Income and Wealth,* 64, 800–827, 2018.

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*, 833–881, Oxford University Press, 2016.

Fleurbaey, M. and V. Peragine, “Ex Ante Versus Ex Post Equality of Opportunity,” *Economica,* 80, 118–130, 2013.

Fleurbaey, M., V. Peragine, and X. Ramos, “Ex post inequality of opportunity comparisons,” *Social Choice and Welfare,* 49, 577–603, 2017.

Foster, J.E. and N. Lustig, “Spotlight 3.2: Choosing an inequality index,” in P. Conceição, ed., *Human Development Report 2019*, 136–138, United Nations Development Programme, New York, 2019.

Foster, J.E. and A.A. Shneyerov, “Path Independent Inequality Measures,” *Journal of Economic Theory,* 91, 199–222, 2000.

Fremeaux, N. and M. Leturcq, “Inequalities and the Individualization of Wealth,” *Journal of Public Economics,* 184, 2–19, 2020.

Goedemé, T., “How much Confidence can we have in EU-SILC? Complex Sample Designs and the Standard Error of the Europe 2020 Poverty Indicators,” *Social Indicators Research,* 110, 89–110, 2013.

Goldin, C., “A grand gender convergence: Its last chapter,” *American Economic Review,* 104, 1091–1119, 2014.

Goldin, C. and L.F. Katz, “A most egalitarian profession: Pharmacy and the evolution of a family-friendly occupation,” *Journal of Labor Economics,* 34, 705–746, 2016.

Haddad, L. and R. Kanbur, “How Serious is the Neglect of Intra-Household Inequality?,” *The Economic Journal,* 100, 866–881, 1990.

Heckman, J.J., “The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models,” *Annals of Economic and Social Measurement,* 5, 475–492, 1976.

———, “Sample Selection Bias as a Specification Error,” *Econometrica,* 47, 153–161, 1979.

Hufe, P., A. Peichl, J.E. Roemer, and M. Ungerer, “Inequality of income acquisition: the role of childhood circumstances,” *Social Choice and Welfare,* 49, 499–544, 2017.

Jenkins, S.P. and P. Van Kerm, “The Measurement of Economic Inequality,” in B. Nolan, W. Salverda, and T.M. Smeeding, eds., *The Oxford Handbook of Economic Inequality*, Oxford University Press Inc., New York, 2009.

Kleven, H.J., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller, “Child Penalties across Countries: Evidence and Explanations,” *AEA Papers and Proceedings,* 109, 122–126, 2019.

Kleven, H.J., C. Landais, and J.E. Søgaard, “Children and Gender Inequality: Evidence from Denmark,” *American Economic Journal: Applied Economics,* 11, 181–209, 2019.

Lechene, V., K. Pendakur, and A. Wolf, “OLS estimation of the intra-household distribution of consumption”, 2019.

Lee, L.-f., “Self-Selection,” in B. Badi, ed., *A Companion to Theoretical Econometrics*, chap. 18, 383–409, John Wiley & Sons, Ltd, 2003.

Mahler, D.G. and X. Ramos, “Equality of Opportunity in Four Measures of Well-Being,” *Review of Income and Wealth,* 0, 1–28, 2019.

Marrero, G.A. and J.G. Rodríguez, “Inequality of Opportunity in Europe,” *Review of Income and Wealth,* 58, 597–621, 2012.

———, “Inequality of opportunity and growth,” *Journal of Development Economics,* 104, 107–122, 2013.

Milanovic, B., “Global Inequality of Opportunity: How Much of Our Income Is Determined by Where We Live?,” *The Review of Economics and Statistics,* 97, 452–460, 2015.

Niehues, J. and A. Peichl, “Upper bounds of inequality of opportunity: Theory and evidence for Germany and the US,” *Social Choice and Welfare,* 43, 73–99, 2014.

Palomino, J.C., G.A. Marrero, and J.G. Rodríguez, “Channels of Inequality of Opportunity: The Role of Education and Occupation in Europe,” *Social Indicators Research,* 143, 1045–1074, 2019.

Papageorge, N.W. and K. Thom, “Genes, Education, and Labor Market Outcomes: Evidence from the Health and Retirement Study,” *Journal of the European Economic Association,* 00, 1–49, 2019.

Puhani, P.A., “The Heckman correction for sample selection and its critique,” *Journal of Economic Surveys,* 14, 53–68, 2000.

Ramos, X. and D. Van de gaer, “Approaches To Inequality of Opportunity: Principles, Measures and Evidence,” *Journal of Economic Surveys,* 30, 855–883, 2016.

———, “Is Inequality of Opportunity Robust to the Measurement Approach?,” *Review of Income and Wealth,* 0, 1–19, 2020.

Roemer, J.E., “A Pragmatic Theory of Responsibility for the Egalitarian Planner,” *Public Affairs,* 22, 146–166, 1993.

———, *Equality of Opportunity*, Harvard University Press, Cambridge, 1998.

———, “Equal opportunity and intergenerational mobility: going beyond intergenerational income transition matrices,” in, trans., *Generational Income Mobility in North America and Europe*, 48–57, Cambridge University Press, 2004.

Roemer, J.E. and A. Trannoy, “Equality of Opportunity: Theory and Measurement,” *Journal of Economic Literature,* 54, 1288–1332, 2016.

Sauer, P., M. Rehm, and K. Mader, “Gender: The Hidden Dimension in the Measurement of Economic Inequality,” in, trans., *Handbook for Labour, Human Resources and Population Economics*, 2020.

Shorrocks, A.F., “The Class of Additively Decomposable Inequality Measures,” *Econometrica,* 48, 613–625, 1980.

———, “Decomposition procedures for distributional analysis: A unified framework based on the Shapley value,” *Journal of Economic Inequality,* 11, 99–126, 2013.

Singh, A., “Inequality of opportunity in earnings and consumption expenditure: The case of Indian men,” *Review of Income and Wealth,* 58, 79–106, 2012.

Suárez Álvarez, A. and A.J. López Menéndez, “Assessing Changes Over Time in Inequality of Opportunity: The Case of Spain,” *Social Indicators Research,* 139, 989–1014, 2017.

———, “Income inequality and inequality of opportunity in europe: Are they on the rise?,” *Research on Economic Inequality,* 26, 149–196, 2018.

———, “Trends in Inequality of Opportunity for Developing Countries: Does the Economic Indicator Matter?,” *Social Indicators Research,* 00, 1–37, 2020.

Teso, E., “The long-term effect of demographic shocks on the evolution of gender roles: Evidence from the transatlantic slave trade,” *Journal of the European Economic Association,* 17, 497–534, 2019.

Teyssier, G., “Inequality of Opportunity and the Parametric Decomposition of Inequality by Population Subgroups”, 2017.

Van de gaer, D. and X. Ramos, “Measurement of inequality of opportunity based on counterfactuals,” *Social Choice and Welfare,* 0 2020.

Appendix

The process determining participation in the labor market is not random. This causes samples used to estimate labor earnings to be censored, leading to biased results. James Heckman (1976, 1979) famously argued that this bias arises because of a missing data problem, and proposed a method to estimate the variables which when omitted in a regression give rise to the specification error.

Known as Heckman correction, the two-step or the limited information maximum likelihood (LIML), this method starts by modeling the sampling probability of each observation with a probit-type selection equation. The aim of this first step is to estimate the so-called inverse Mills ratio—also known as hazard ratio—, which will be included as a regressor in a subsequent equation to estimate the conditional expectation of the dependent variable in the second step. In our case, the sampling probability refers to the odds of participating in the labor market, and the dependent variable is the amount of individual income, conditional on certain personal characteristics. Actually, testing the null that the coefficient on the inverse Mills ratio is 0 equates to testing for the presence of a selection bias.

Estimating IOP using an outcome at the individual level is likely to suffer from this bias, which is why we explicitly model selection into employment. The selection equation is a probit regression with the following specification:

$$D_i = \beta_0 + \beta_1 female_i + \beta_2 couple_i + \beta_3 couple_i \times female_i + \beta_4 children_i + \beta_5 children_i \times female_i \\ + \beta_6 immigrant_i + \varepsilon_i, \quad (14)$$

where D_i is a dummy indicating whether individual i participates in the labor market, *female* refers to binary gender, *couple* to mating, *children* references the presence of underage individuals in the household, and *immigrant* refers to having being born outside the country of residence.

The kind of employment modeled, i.e., the labor market participation condition D_i, is full- and part-time employees and self-employed who spent in that situation at least 7 months of the reference period.[39] That allows us to study a relatively stable position in the labor market.

Some variables have been defined before, specifically *female* and *immigrant* (in section 4.1). Couple (variable RB240) is defined as couples living in the same household, either with legal or informal bindings. Age (RX020) is considered at the end of the reference period. The only variable that has not been described before is *children*, which takes the value 1 in the presence

[39] We used the variables "Number of months spent at full-time work" (PL070) and "Number of months spent at part-time work" (PL072) in waves prior to 2008. In waves from 2008 on these variables were updated in the survey design to "Number of months spent at full-time work as employee" (PL073), "Number of months spent at part-time work as employee" (PL074), "Number of months spent at full-time work as self-employed (including family worker)" (PL075), and "Number of months spent at part-time work as self-employed (including family worker)" (PL076).

Table 6: Selection model into employment

	ln Wage_i	D_i
Age	0.0478*	
	(0.0045)	
Age × Age	-0.0005*	
	(0.0001)	
Personal education		
Secondary	-0.0703*	
	(0.0092)	
Tertiary or more	0.3474*	
	(0.0111)	
Personal occupation		
Skilled worker	0.2793*	
	(0.0119)	
Professional	0.7635*	
	(0.0129)	
Female		0.0468*
		(0.0135)
Couple		0.2873*
		(0.0126)
Female × Couple		-0.3969*
		(0.0164)
Children		0.3612*
		(0.0116)
Female × Children		-0.3466*
		(0.0147)
Immigrant status		-0.0957*
		(0.0113)
Inverse Mills ratio	-1.5244*	
	(0.0330)	
Constant	8.5951*	0.5331*
	(0.1006)	(0.0098)
Observations	169,111	
Censored values	42,534	

Note: This table shows the results of the selection model regressions on the pooled data of our sample in 2010. D_i is the censoring condition. The reference values of personal education and occupation are their lowest categories, "Primary or less" and "Elementary occupations". Robust standard errors clustered by country in parenthesis. * $p < 0.01$. Cross-sectional files of the EU-SILC database.

of underage household members and 0 otherwise.

The outcome equation, for estimating labor earnings, is as follows:

$$\ln wage_i = \beta_0 + \beta_1 age_i + \beta_2 age_i^2 + \beta_3 education_i + \beta_4 occupation_i + \varepsilon_i, \tag{15}$$

where *education* and *occupation* are the highest educational level and the occupational category of individual i, respectively. Age is self-explanatory. Personal education (variable PE040) and occupation (PL050) are defined in exactly the same way as the variables parental education and occupation (coded in three levels following ISCED-97 and ISCO-88 classifications, respectively—see section 4.1).

Equation (15) does not include any variable already present in eq. (14) in order ot avoid possible collinearity problems, which can be a severe issue with this technique (Lee 2003; Puhani 2000). Indeed, with this specification we find no evidence of correlation between regressors of eq. (15) and the inverse Mills ratio. We test this by calculating the R^2 of a regression of the inverse Mills ratio on the regressors of eq. (15), which is 0.0186 (Puhani 2000). However, we do find evidence of a selection bias, as indicated by the significance of the inverse Mills ratio—see table 6.

Since our sample is composed by 31 countries space restrictions prevent us from reporting results of the selection model for each case. Nevertheless, following Checchi, Peragine, and Serlenga (2016) we do report results on the pooled data of all countries in table 6, in order to provide with a general idea of the results. The inverse Mills ratio is significant (as it is in all countries individually—not reported), bringing evidence of a selection bias. Also, as indicated above, we find no evidence of collinearity.