23rd Public Economics Meeting, University of Vigo
February 4 and 5, 2016

Inequality of Opportunity: a Parametric Approach

José María Sarabia[1], Vanesa Jordá, Faustino Prieto

Department of Economics, University of Cantabria, Avda de los Castros s/n, 39005-Santander, Spain

Abstract

In this paper we measure the extent of inequality of opportunity using a new parametric approach. We rely on the idea that differences in earning can be ethically fair if they come from choices that individuals are completely responsible for, called *efforts*. Unacceptable disparities are those come from external factors called *circumstances*. We use a flexible parametric form to model the joint distribution of circumstances and efforts. Our model avoids the need to break continuous variables of circumstances into groups, while also allowing for the use of categorical variables such as gender or race. Using data from Living Conditions Survey, we estimate the level of inequality of opportunity in Spain in 2005. Our results are robust to different parametric functional forms and indicators of parental education.

JEL Classification: D31, D63, J62

Key Words: Multivariate distributions; Inequality of opportunity; Theil indices; Decompositions

1 Introduction

Income inequality has been of great interest in the economic literature, being also a sort of concern as regards redistribution policies. It is however argued that disparities driven by different levels of efforts of individuals are

[1]E-mail addresses: sarabiaj@unican.es (J.M. Sarabia).

less objectionable than those arise from differences in the characteristics of these individuals which are exogenous to them. Among the last kind of features we can highlight race, gender, familiar background or region, among others. In this sense, the distribution of an specific outcome such as earnings, income or wealth is not relevant itself. Instead, the differences due to predefined circumstances are of main interest in terms of social justice. The resulting inequalities once the circumstances have been equalized are driven by individual choices of which individuals are totally responsible, and hence outside the scope of policy makers.

The idea that different achievements can be ethically fair if they come from individual choices has been repeatedly pointed out in the literature (see e.g. Rawls, 1971; Dworkin, 1981; Cohen, 1989). Unacceptable disparities are those come from external factors that individuals cannot control for. Roemer (1993, 1998) formalizes of this philosophical distinction by defining 'equality of opportunity' as the the independence of the distribution of particular outcomes to the exogenous characteristics of the individual. The distribution of individual achievements can be decomposed into two different factors, namely circumstances and efforts, which would lead to two different components of inequality of the overall outcome of interest: inequality of efforts and inequality of opportunity.

According to Ferreira and Gignoux (2011), there are at least three reasons why the study of equality of opportunity is of main interest in the economic field. First of all, as regards policy design, the actions should be focused on equalizing circumstances or at least compensating worse-off individuals concerning determinants beyond their responsibility. On the other hand, inequality driven by differences in efforts are socially accepted while disparities due to circumstances beyond the individual scope are ethically unfair. Therefore, policies targeted to reduce inequalities due to this kind of factors would receive and stronger social support. Last but not least, inequality of opportunity seems to be more correlated with the economic performance of the countries than overall disparities. In fact, this kind of inequality has a negative impact on economic growth while differences fostered by effort seem to be positively correlated with growth (Marrero and Rodriguez, 2013).

Based on the previous definition of inequality of opportunity, the distributions of income for individuals with homogeneous characteristics must be the identical. Taking this interpretation literally, inequality of opportunity can be studied by comparing the distributions of income conditional on the type using first order and second order dominance tests (Lefranc et al., 2008;

2009). If no stochastic dominance relationships can be achieved, the society is equal in terms of opportunities, so differences in outcomes would be consequence of the efforts exerted by individuals. The main limitation of this methodology is the impossibility to obtain enough large samples to estimate conditional income distributions. This requirement limits the definition of groups to a reduced number, making the *types* quite general groups of individuals. Note that once the types have been defined any differences in income within types are attributed to effort or luck. If the set o individuals within a given type presents heterogeneity in terms of circumstances, we would be overestimating the level of 'ethically acceptable' inequality. Moreover, stochastic dominance does not inform about the degree of inequality of opportunity and hence it would be not possible to rank different distributions according to this concept of disparities.

This is the reason why several studies have focused on an alternative while less general methodology of constructing indices of inequality of opportunity. Within this framework two approaches can be distinguished. On the one hand, nonparametric techniques have been used to decompose overall inequality into differences due to external circumstances of individuals and a second component that would represent the morally fair inequalities. The procedure is based on the decomposition of inequality in the within and the between components of mutually exclusive groups defined by the considered *types* of individuals (see Checchi and Peragine, 2010). The results are again heavily dependent on the number of types defined. Indeed, continuous variables such as parental education measured by years of schooling need to be discrtized in order to apply this methodology, thus resulting in an underestimation of inequality of opportunity. In fact, this bias would be higher as the number of groups decreases. In contrast, the consideration of too many groups can lead non-representative results due to the low frequency of observations in some of them.

An alternative approach followed by Bourguignon et al. (2007) would be appealing when samples are not fully representative of all types, although at cost of imposing a linear parametric form. In particular, this methodology estimates a linear model to assess the impact of 'effort' and 'circumstance' variables on earnings. These estimates are used to build counterfactual distributions that equalize circumstances across types. The comparison of inequality levels of the actual and the counterfactual distribution yield a quantification of inequality of opportunity. The main advantage of this methodology is that it permits to test the impact of specific circumstances controlling by the other factors. Its flexibility has make this approach very popular, being one of the most widely used in empirical investigations (Bourguignon et al. 2007; Ferreira and Gignoux, 2011; Marrero and Rodriguez, 2012; 2013). The hypothesis of a linear parametrization has been relaxed by Pistolesi (2009) by estimating a semiparametric specification based on duration models.

This paper proposes an alternative parametric approach to assess the level of inequality of opportunity. We use a flexible functional form to model the joint distribution of circumstances and efforts. Our model avoids the need to break continuous variables of circumstances into groups, while also allowing for the use of categorical variables such as gender or race. The use of parametric functional forms for the study of earnings and income distributions has been well documented in the literature (see Cowell, 2000; Slottje, 1990), albeit it has not yet been applied to quantify inequality of opportunity. As a model for earnings distribution, we use the generalized beta of the second kind (GB2) (see McDonald, 1984). The GB2 is a wide family which includes many well-known models as special or limiting cases, thus providing an excellent description of income distributions with few parameters. We obtain closed expressions of inequality measures of the multivariate distribution to evaluate the degree of inequality of opportunity.

The contents of this paper are as follows. In section 2 we present the theoretical framework and the proposed parametric model used to assess inequality of opportunity. The data used and the estimation methods are described in Section 3. An illustration of the methodology presented in the paper is given in Section 4. Finally, section 5 concludes.

2 Measuring inequality of opportunity

Within the framework of equal opportunities, differences in what was called originally by Roemer (1998) *advantages* (income, earnings or consumption) can come from two different sources. On the one hand, individuals exert different levels of effort which would result in differences in income as a kind of *natural reward*. In contrast, inequality can be driven by determinants behind individual's control called circumstances for which the individual should be compensated (Fleurbaey, 1995) since this kind of differences are not ethically acceptable.

Following Bourguignon et al. (2007), we will focus in this paper on the distribution of earnings denoted by w, which are a function of factors exogenous to individuals given by circumstances (C), the level of effort (e) and unobservable determinants that we denote by u,

$$w_i = f(C_i, e_i, u_i).$$

The level of effort is considered to be a continuous variable, while circumstances are made up of discrete variables which divide population into mutually exclusive groups of homogeneous individuals. By definition, C_i are exogenous in the model since the individual cannot choose among the set of circumstances. Efforts are arguably influenced by circumstances, so the previous equation can be expressed as,

$$w_i = f(C_i, e_i(C_i, v_i), u_i).$$

According to Roemer (1998), there is not inequality of opportunity if the distribution of earnings is independent to the circumstances, $F(w|C) = F(w)$, where $F(.)$ stands for the cumulative distribution function (CDF) of earnings. This condition is known as strong equality of opportunity (Lefranc et al., 2008). Indeed, this is a special case of equality of opportunity that relies on first order stochastic dominance (FSD) criterion when all points coincide.
Let c and c' be two different circumstances, we say that c FSD-dominates c' iff $F(w|c) \leq F(w|c') \forall w \in \Re_+$ (Lefranc et al., 2008). If so, all individuals would prefer to be characterized by the vector of circumstances c and the distribution of w would present inequality of opportunity. Instead, if the conditional CDF on c and *c'* cross, any set of circumstances is preferred over the other, we say that there is equality of opportunity. A less demanding criterion based on second order stochastic dominance (SSD) is also considered by Lefranc et al. (2008). The circumstances c SSD-dominates *c'* iff $\int_0^x F(w|c)dw \leq \int_0^x F(w|c')dw, \ \forall x \in \Re_+$. Again, the absence of dominance relationships based on SSD implies that the individual is not able to rank all possible circumstances which would correspond to equality of opportunity.

The previous approach to characterize inequality of opportunity presents two main limitations. The methodology based on stochastic dominance is not able to rank situations when there is inequality of opportunity. On the other hand, this methodology presents some restrictions on the empirical ground, given that the size of the sample of the conditional distributions decreases as the number of circumstances increases. Alternatively, we can rely on a weaker condition to identify differences across different vectors of circumstances using summary statistics of conditional distributions. In this sense, Van de Gaer (1993) proposed the equality of conditional means as a necessary condition of equal opportunities, being also congruent with the original definition given by Roemer (1998). To assess the extent of unfair disparities, inequality measures are computed on the distribution of conditional means, which is equivalent to evaluate the amount of inequality that would exist in the economy in there were no differences within each type.

We should be however cautious as regards the interpretation of these measures. It has been repeatedly pointed out that they would represent a lower bound of the actual level of inequality of opportunity. Ferreira and Gignoux (2011) proved that, as long as there are unobserved circumstances that are not taken into account in our analysis, we will be underestimating the disparities driven by these exogenous factors. This reasoning is also extended to the partition of circumstances into more categories. Consider, for instance, parental education. This variable can be categorized into different educational levels or can be expressed in years, a structure that would lead no less inequality of opportunity than the categorical one. Consequently, assessments of inequality of opportunity are heavily dependent on the structure of the data. In the nonparametric approach, only categorical circumstances are considered in order to classify individuals into Roemerian types that brake population into mutually exclusive groups. Continuous variables are then categorized in order to apply this approach. While the number of categories is totally arbitrary, it has a strong impact on the evaluation of ethically not acceptable inequalities. Ideally, we should respect the continuous nature of these circumstances in order to avoid any possible bias due to discretization.

However, this methodology cannot be implemented in the case of nonparametric techniques because it would result in the definition of infinite mutually exclusive groups.

This limitation is avoided by the parametric approach proposed by Bourguignon et al. (2007), which allows for the consideration of discrete and continuous variables to model the set of circumstances. This advantage comes with the price of imposing a linear parametric structure that relates the distribution of a particular advantage with effort and circumstance variables. In the same spirit, we also use a parametric model to estimate the extent of inequality of opportunity, including continuous and categorical circumstances. In contrast to previous studies, we rely on multivariate distributions of effort and circumstance variables to model the joint distribution of earnings.

We start considering the joint distribution of earnings (w), effort(e) and circumstances (C), which is given by $f(w,e,C)$, where w and e are univariate variables and C is a vector of J variables. The the marginal distribution of earnings is given by,

$$f(w) = \int_{\Re^J}\int_e f(w,e,C)dedC.$$

Using conditional densities, the previous equation can be rewritten in the following form,

$$f(w) = \int_{\Re^J}\int_e f(w|e,C)g(e|C)h(C)dedC.$$

The previous equation is the general expression when all variables are continuous. Consider now that we divide our vector of circumstances into P discrete (C_d) and $J-P$ continuous (C_c) variables. The P discrete variables can be summarized into a single one that includes all the possible combinations (say k) between the categories of the P characteristics. Then the previous equation can be expressed as,

$$f(w) = \sum_{j=1}^{k}\Pr(C_{d_j})\int_{\Re^{J-P}}\int_e f(w,e,C_c|C_d)dedC_c.$$

where $\Pr(C_{d_j})$ is the probability of being characterized by the vector of circumstances C_{d_j}.

Our methodology relies on parametric distributions to model $f(w,e,C_c|C_d)$. In the standard nonparametric framework, all possible circumstances are included in C_d, which defines mutually exclusive groups (*types*) and then $f(w) = \sum_{j=1}^{k}\Pr(C_{d_j})f(w,e|C_d)$, where the conditional distribution of effort on the circumstances is independent across types. Our methodology permits the inclusion of continuous variables in the analysis, which makes the conditional distribution of effort on discrete circumstances no longer independent.

Let $X_i, i=1,\ldots,P$ be the distribution of earnings for a given type i, i.e. conditioned to belong to some vector of discrete characteristics. Our methodology relies on the GB2 family to model the conditional distributions,

$$X_i \sim \mathcal{GB}2(a_i, p_0, q_i, \sigma_i), \quad i = 1,2,\ldots,P.$$

A description of the different parametric distributions used in the paper is included in the Appendix.

This family seems to be particularly suitable to model income distributions given that it contains several parametric models used to fit income variables, including the beta 2, the Singh-Maddala and the Dagum distributions (Bresson, 2009; Hajargasht et al., 2012, Chotikapanich et al., 2012).

The conditional distributions of earnings on discrete circumstances are driven by two different components: the distribution of effort within that type (e_i) and the distribution of continuous circumstances (c_c), which are common to all types. To consider this structure, we use the following representation of the GB2 distribution,

$$X_i = \sigma_i \left(\frac{c_c}{e_i} \right)^{1/a_i}, \quad i = 1, 2, \ldots, P, \tag{1}$$

where $c_c, e_1, \ldots, e_P$ are mutually independent gamma random variables with distributions $c_c \sim \mathcal{G}a(p_0)$ and $e_i \sim \mathcal{G}a(q_i)$, $i = 1, 2, \ldots, P$. Note that the continuous circumstances are common to all types, but their influence on the conditioning distribution (X_i) may differ across groups due to the different value of the parameter a_i.

The common variable c_c introduces the dependence in the model, so we can model the multivariate distribution of earnings of all types as follows,

$$(X_1, X_2, \ldots, X_P)^\top = \left(\sigma_1 \left(\frac{c_c}{e_1} \right)^{1/a_1}, \sigma_2 \left(\frac{c_c}{e_2} \right)^{1/a_2}, \ldots, \sigma_m \left(\frac{c_c}{e_P} \right)^{1/a_P} \right)^\top, \tag{2}$$

where $a_i, \sigma_i > 0$, $i = 1, 2, \ldots, P$.

These classes of multivariate distributions will be constructed using "variables in common" techniques (see, Balakrishnan and Lai, 2009). The idea of this methodology is to construct pairs of dependent random variables from three or more random variables[2]. In many situations these initial random variables are independent, but occasionally they may be dependent. In our case, the functions connecting these random variables are given by (19), where all the pairs of random variables share the same numerator or denominator.

[2]This methodology has been used recently for constructing multivariate dependent beta (Olkin and Liu, 2003), Student t (Fang et al., 1990; Jones, 2002), Marshall-Olkin (Sarhan and Balakrishnan, 2007) and F (El-Bassiouny and Jones, 2009) distributions (see Sarabia and Gómez-Déniz, 2008).

Theorem 1 *The joint probability density function (PDF) of the multivariate random variable (2) is given by,*

$$f_{X_1,\ldots,X_P}(x_1,\ldots,x_P) = \frac{\Gamma\left(p_0 + \sum\limits_{i=1}^{P} q_i\right)}{\Gamma(p_0)\prod\limits_{i=1}^{P}\Gamma(q_i)} \cdot \frac{\prod\limits_{i=1}^{P} \frac{a_i}{\sigma_i} \cdot \left(\frac{x_i}{\sigma_i}\right)^{-a_i q_i - 1}}{\left[1 + \sum\limits_{i=1}^{P}\left(\frac{x_i}{\sigma_i}\right)^{-a_i}\right]^{p_0+q_1+\cdots+q_P}}, \quad (3)$$

if $x_i > 0$, $i = 1, 2, \ldots, P$ *and 0 elsewhere.*

Proof: See Appendix. ■

It is possible to define a second version of a multivariate GB2 distribution, where now the shape parameter q_0 is fixed. the conditional distributions of each type are given by,

$$X_i \sim \mathcal{GB}2(a_i, p_i, q_0, \sigma_i), i = 1, 2, \ldots, P.$$

Accordingly, we define the multivariate P-dimensional random variable,

$$(X_1, X_2, \ldots, X_P)^\top = \left(\sigma_1\left(\frac{e_1}{c_c}\right)^{1/a_1}, \sigma_2\left(\frac{e_2}{c_c}\right)^{1/a_2}, \ldots, \sigma_P\left(\frac{e_P}{c_c}\right)^{1/a_P}\right)^\top, \quad (4)$$

where $e_i \sim \mathcal{G}a(p_i)$, $i = 1, 2, \ldots, P$ and $c_c \sim \mathcal{G}a(q_0)$. The common random variable c_c introduces the dependence in the multivariate random variable.

Theorem 2 *The joint PDF of (4) is given by,*

$$f_{X_1,\ldots,X_P}(x_1,\ldots,x_P) = \frac{\Gamma\left(q_0 + \sum\limits_{i=1}^{P} p_i\right)}{\Gamma(q_0)\prod\limits_{i=1}^{P}\Gamma(p_i)} \cdot \frac{\prod\limits_{i=1}^{P} \frac{a_i}{\sigma_i} \left(\frac{x_i}{\sigma_i}\right)^{a_i p_i - 1}}{\left[1 + \sum\limits_{i=1}^{P}\left(\frac{x_i}{\sigma_i}\right)^{a_i}\right]^{q_0+p_1+\cdots+p_m}}, \quad (5)$$

if $x_i > 0$, $i = 1, 2, \ldots, P$ *and 0 elsewhere.*

Proof: See Appendix. ■

To investigate the level of inequality of opportunity we need to distinguish between differences due to effort and whose are driven by circumstances.

With this aim, we need to use an additively decomposable inequality measure by population subgroups in order to distinguish these two components of overall inequality independently. Among the range of inequality measures only the Generalized Entropy family has this characteristic. In line with previous studies, we rely on the Theil-0, also called MLD index, which in addition of being additively decomposable, presents several appealing properties for the study of inequality of opportunity[3]. This inequality index can be generally expressed as,

$$T_0(X) = -\mathbb{E}\left[\log\left(\frac{X}{\mu_X}\right)\right], \tag{6}$$

where $\mu_X = \mathbb{E}(X)$. For the decomposition of earnings inequality into inequality of opportunity and inequality of efforts, we will use the following lemmas, which presents the Theil 0 index for the GB2, the GG and the inverted GG distributions.

Lemma 1 *Let $X \sim \mathcal{GB}2(a,p,q,b)$ be a GB2 distribution. Then, the Theil index is given by ($q > 1/a$)*

$$T_0(X) = -\frac{\psi(p) - \psi(q)}{a} + \log\frac{\Gamma(p+\frac{1}{a})\Gamma(q-\frac{1}{a})}{\Gamma(p)\Gamma(q)}, \tag{7}$$

where $\psi(z) = \Gamma'(z)/\Gamma(z)$ denotes the digamma function.

The Theil-0 index for the GG and the inverted GG distributions are given in the following result.

Lemma 2 *Let $X \sim \mathcal{GG}(a,p,b)$ be a GG distribution. Then, the Theil index is given by,*

$$T_0(X) = -\frac{\psi(p)}{a} + \log\frac{\Gamma(p+\frac{1}{a})}{\Gamma(p)}, \tag{8}$$

Now, let $\tilde{X} \sim \mathcal{IGG}(a,p,b)$ be a inverted GG distribution. If $p > 1/a$, the Theil index is given by,

$$T_0(\tilde{X}) = \frac{\psi(p)}{a} - \log\frac{\Gamma(p-\frac{1}{a})}{\Gamma(p)}, \tag{9}$$

where $\psi(z) = \Gamma'(z)/\Gamma(z)$ denotes the digamma function.

[3]The main advantage of this measure for the study of inequality of opportunity is its property of path-independent decomposition (Foster and Shneyerov, 2000).

The proof of Lemma 1 can be found in Jenkins (2009) and the proof of Lemma 2 in Sarabia et al. (2015).

Using the stochastic representations (2 and 4), we decompose earning inequality into four components.

Theorem 3 *Assume that earnings w can be decomposed into P mutually independent groups $X_1, \ldots, X_P$, such that $f_W(w) = \sum_{i=1}^{P} \pi_i f_{X_i}(w)$, where $\pi_i \geq 0$ and $\sum_{i=1}^{P} \pi_i = 1$. For the random variables X_i, assume the common factor model,*

$$X_i = \sigma_i \left(\frac{c_{p_0}}{e_{q_i}} \right)^{1/a_i}, \quad i = 1, 2, \ldots, P,$$

where $c_{p_0} \sim \mathcal{G}(p_0)$, $e_{q_i} \sim \mathcal{G}(q_i)$, $i = 1, 2, \ldots, P$ are independent gamma random variables with $p_0 > 0$ and $\sigma_i, q_i > 0$, $i = 1, 2, \ldots, P$.

Then, the Theil-0 index of w can be decomposed in the following as,

$$\begin{aligned} T_0(X) &= \sum_{i=1}^{P} \pi_i T_0(e_{q_i}^{-1/a_i}) + \sum_{i=1}^{P} \pi_i T_0(c_{p_0}^{1/a_i}) - \\ &\quad - \sum_{i=1}^{P} \pi_i \log \frac{\mathbb{E}(c_{p_0}^{1/a_i})}{\mathbb{E}(c)} - \sum_{i=1}^{P} \pi_i \log \frac{\mathbb{E}(e_{q_i}^{-1/a_i})}{\mathbb{E}(\tilde{Y})}, \end{aligned} \tag{10}$$

where $T_0(e_{q_i}^{-1/a_i})$ and $T_0(c_{p_0}^{1/a_i})$ represent the Theil-0 indices of a $\mathcal{IGG}(a_i, q_i, 1)$ and a $\mathcal{GG}(a_i, p_0, 1)$ random variables, given in (9) and (8) respectively, and

$$\mathbb{E}(c) = \sum_{i=1}^{P} \pi_i \mathbb{E}(c_{p_0}^{1/a_i}) = \sum_{i=1}^{P} \pi_i \frac{\Gamma(p_0 + \frac{1}{a_i})}{\Gamma(p_0)},$$

and

$$\mathbb{E}(\tilde{e}) = \sum_{i=1}^{P} \pi_i \mathbb{E}(e_{q_i}^{-1/a_i}) = \sum_{i=1}^{P} \pi_i \frac{\Gamma(q_i - \frac{1}{a_i})}{\Gamma(q_i)}.$$

Proof: See Appendix. ∎

The first term in (10) represent the "inequality of efforts", which does not consider any disparities driven from the circumstances. Assuming that effort can be modeled as a residual component, in the sense that it represents the part of earnings variability that is not captured by the circumstances, this term informs about the inequality within each type once the effect of circumstances is removed. The second and the third terms in (10) represent the part of "inequality of opportunity" driven by the continuous circumstances. The first component informs about earning disparities due to inequality in the continuous circumstance within types. the second component, instead, corresponds to earning inequality derived by differences in opportunities between types. The last term in (10) informs about the "inequality of opportunity" due to discrete circumstances, once the effect of continuous circumstance has been removed.

3 Data and estimation strategy

We use data from Living Conditions Survey (LCS) collected by the Spanish Statistical Office. This survey is conducted annually since 2004 collecting data on socio-demographic variables. Unfortunately, the year 2005 is the only one that contains information about individual's background which is used in the literature of inequality of opportunity to model the exogenous circumstances [4]. The LSC sample size in 2005 was 30375, but we restrict our sample to individuals between 25 and 65 (both included) who had positive earnings during 2004, which leads a sample of 16447 individuals. Due to missing observations in other relevant variables related with circumstances (gender, parental education and country of birth) our sample is further restricted to 8282 individuals.

Following previous studies, we use annual gross earnings as a measure of *advantage* (see Checchi and Peraigne, 2010; Sapata,2009). As regards circumstance variables, the data set has available tree kind of circumstances: (i) gender, (ii) country of birth and (iii) parental education. Table 1 presents information about the categories of each of these variables and their associated proportion.

The first two variables belong to the purely discrete set of circumstances (C_d). Parental education also presents a discrete structure, but is usually considered as a continuous variable in the parametric approach. The cate-

[4] In 2011, a new wave of the special module of inter-generational poverty is also available. However, the structure of the data regarding the categories of parental education and other variables of circumstances is fairly different from the data in 2005 and the results would be substantially affected by the different categorization of circumstances (Ferreira and Gignoux, 2011).

Table 1: Descriptive statistics

Variable	Description	Mean / proportion
Earnings	Annual gross earnings	21264.0311
Gender	Male	0.5717
	Female	0.4283
Mother's Education	Less than primary	0.2370
	Primary	0.6073
	Lower Secondary	0.0697
	Upper Secondary	0.0442
	Post-secondary	0.0027
	Tertiary	0.0391
Father's Education	Less than primary	0.1941
	Primary	0.5529
	Lower Secondary	0.0676
	Upper Secondary	0.0601
	Post-secondary	0.0051
	Tertiary	0.0923
Highest educational level of the couple	Less than primary	0.5739
	Primary	0.0762
	Lower Secondary	0.0694
	Upper Secondary	0.0060
	Post-secondary	0.0060
	Tertiary	0.1079
Country of birth	Spain	0.9367
	Other country	0.0633

gorical variable is transformed into a cardinal indicator by giving the same numbers of years of schooling to individuals that belong to the same category (see e.g. Bourguignon et al., 2007). The main limitation is the assumption of a linear relationship between the advantages and the circumstances [5]. Our analysis relies on an alternative parametric specification which provides us with an statistical framework that explicitly addresses the dependence between circumstances and advantages. We use the representation of the GB2 as a ratio of gamma distributions (Eq.(1)), modeling parental educational achievements as a continuous circumstance. The relationship between efforts and circumstances is explicitly addressed by this representation and no subjective choices need to be made. However, two requirements must to be satisfied to ensure the accuracy of the estimates. On the one hand, the GB2 family should be an adequate model for distribution of earnings of each type. Indeed, this distribution is particularly well-suited to model income and earnings distributions (see Kleiber and Koltz, 2003; McDonald, 1984). Secondly, the Gamma distribution should fit adequately the educational outcomes of the country. This statement has been barely investigated in the literature. The distribution of education has been principally studied relying on non-parametric techniques, using the official duration of each educational stage to transform educational achievements into a numerical indicators. On the other hand, duration models have been used to investigate the probability of dropping out school (see e.g. Arulampalam et al., 2004 Plank et al., 2005). We take a similar approach in this study, but instead of focusing only on dropout patterns, we investigate the time until leaving the educational system independently of the reason. Then, education is represented by the variable time that individuals attend the school until leaving the educational system. This approach is applied by Jordá (2015), using a general model, which includes the gamma distribution as special case. The results point out an outstanding performance of the continuous approach which fits particularly well the distribution of educational outcomes.

For the estimation purposes, denote earnings by w. The individuals are classified into four mutually exclusive groups: Spanish males X_1; foreign males X_2; Spanish females X_3 and foreign females X_4. Then, we can write $f(w|C_d) = \sum_{i=1}^{4} \pi_i f_{X_i}(x|C_d)$, where $\pi_i \geq 0$, $i = 1, 2, 3, 4$ and $\sum_{i=1}^{4} \pi_i = 1$ is the proportion of the ith type. The sample size of each category will be

[5]An alternative nonlinear specification based on semiparametric duration models is assumed by Pistolesi (2009)

$n \times \pi_i = n_i$, $i = 1, 2, 3, 4$.

The second kind of circumstances is given by parental education. Given that we have information for both parents, we have considered three different possibilities: father's years of schooling (Lefranc et al., 2008), mother's years of schooling and the maximum years of education in the couple (Checchi and Peragine, 2010). According to previous discussions, the education variable is modeled as a parametric continuous random variable given by the gamma distribution. The gamma distribution will be denoted by $c_c \sim \mathcal{G}(\alpha, \sigma)$, where $\alpha \geq 0$ is the shape parameter and $\sigma \geq 0$ the scale parameter.

Then, the data consist in pairs of observations $(x_{ij}, w_{ij})|c_c$, with $j = 1, 2, \ldots, n_i$, $i = 1, 2, 3, 4$, where x_{ij} are the annual gross earnings of the jth individual belonging to group ith, with sample weights w_{ij} and conditional to the educational attainment c_c.

We proceed in two steps: estimation of parental education and estimation of earnings for different types.

Step 1 We estimate the distribution of parental education by weighted non-linear least squares (WNLS) fitting,

$$(\alpha, \sigma)|c_c = \underset{\alpha,\sigma}{\operatorname{argmin}} \sum_{i=1}^{4} \sum_{j=1}^{n_i} \pi_i w_{ij} \left\{ \tilde{F}_{ij} - F_{\mathcal{G}(\alpha,\sigma)}(x_{ij}; \alpha, \sigma|c_c) \right\}^2, \tag{11}$$

where $\tilde{F}_{ij}$ is the value of the empirical CDF and

$$F_{\mathcal{G}(\alpha,\sigma)}(x_{ij}; \alpha, \sigma|c_c) = \int_0^{x_{ij}|c_c} \frac{t^{\alpha-1} e^{-t/\sigma}}{\sigma^\alpha \Gamma(\alpha)} dt$$

is the theoretical CDF of the gamma distribution. The initial estimates of the parameters are,

$$\begin{aligned} \hat{\alpha}_{init} &= m_{ij}^2 / s_{ij}^2, \\ \hat{\sigma}_{init} &= s_{ij}^2 / m_{ij}, \end{aligned}$$

where m_{ij} and s_{ij}^2 are the sample mean and variance, respectively.

Step 2 Estimation of earnings for each type. We assume that $X_i|c_c \sim \mathcal{GB}2(a, p_0, q, b)$ and $X_i|c_c \sim \mathcal{GB}2(a, p_0, q, b)$, where p_0 is estimated from the education distribution in Step 1. For the first model, the log-likelihood is given by,

$$\log \ell(a,q,b|c_c) = \sum_{j=1}^{n_i} w_{ij} \log f_{\mathcal{GB}2}(x_{ij}; a, p_0, q, b|c_c), \quad i = 1,2,3,4, \tag{12}$$

which can be written as,

$$\begin{aligned}\log \ell(a,q,b|c_c) &= w_{i+}(\log\Gamma(p_0+q) + \log a) + (ap_0 - 1)\sum_{j=1}^{n_i} w_{ij}\log x_{ij} \\ &\quad - w_{i+}(ap_0\log b + \log\Gamma(p_0) + \log\Gamma(q)) - \\ &\quad -(p_0+q)\sum_{j=1}^{n_i} w_{ij}\log\left(1 + (x_{ij}/b)^a\right),\end{aligned}$$

where $w_{i+} = \sum_{j=1}^{n_i} w_{ij}$.

If we denote by $\boldsymbol{\theta} = (a,q,b)$ the vector of the parameters and the derivative of $\log\Gamma(x)$ by $\psi(x)$, the digamma function, the components of the score vector $U(\boldsymbol{\theta})$ are given by,

$$U_a(\boldsymbol{\theta}) = \frac{w_{i+}}{a} + p_0\sum_{j=1}^{n_i} w_{ij}\log\left(\frac{x_{ij}}{b}\right) - (p_0+q)\sum_{j=1}^{n_i} w_{ij}\log\left(\frac{x_{ij}}{b}\right)\left[\left(\frac{b}{x_{ij}}\right)^a\right]^{-1} \tag{13}$$

$$U_q(\boldsymbol{\theta}) = w_{i+}\psi(p_0+q) - w_{i+}\psi(q) - \sum_{j=1}^{n_i} w_{ij}\log\left[1 + \left(\frac{x_{ij}}{a}\right)^a\right], \tag{14}$$

$$U_b(\boldsymbol{\theta}) = -\frac{w_{i+}ap_0}{b} + \frac{(p_0+q)a}{b}\sum_{j=1}^{n_i}\left[1 + \left(\frac{a}{x_{ij}}\right)^a\right]^{-1}. \tag{15}$$

The initial estimates can be obtained from the moment estimates of the Fisk distribution.

The Fisher information matrix of the GB2 with complete information was obtained by Brazaukas (2002), and can be adapted easily to our model. Note that this distribution is a regular family in terms of the maximum likelihood estimation and we may use the expression,

$$I(\boldsymbol{\theta}) = -\mathbb{E}\left(\frac{\partial^2 \log \ell}{\partial\theta_i \partial\theta_j}\right)_{ij}.$$

For the estimation of the model $X_i|c_c \sim \mathcal{GB}2(a,p,q_0,b)$ we proceed in a similar way.

4 Results

The results corresponding to the estimation of the models based on the $\mathcal{GB}2(a, p_0, q, b)$ (p_0 fixed) and $\mathcal{GB}2(a, p, q_0, b)$ (q_0 fixed) distribution are given in Tables 2 and 3. For each two models, we have fitted the earnings distribution X (with p_0 or q_0 fixed respectively) according to the four types X_1, X_2, X_3 and X_4, conditionally to the three different definitions of parental education (father's years of schooling, mother's years of schooling and combined). These tables include the parameter estimates and the standard errors (in parentheses) using the previous parametric models.

We have compared both models using the Akaike information criterion (AIC), defined by (Akaike 1974),

$$AIC = -2\log\ell + 2d,$$

where $\log\ell = \log\ell(\hat{\theta})$ is the log-likelihood of the model evaluated at the maximum likelihood estimates and d is the number of parameters. We chose the model with the smallest value of AIC statistic. We have also included the BIC statistics, and in this case we chose that model with highest value.

For all the estimations, the parameters of the two GB2 are significant. If we compare both models, the $\mathcal{GB}2(a, p_0, q, b)$ distribution outperforms the $\mathcal{GB}2(a, p, q_0, b)$ distribution, in terms of the statistics AIC and BIC (see Tables 2 and 3). On the other hand, if we compare the estimation in terms of the three education indicators, the estimations based on the years of education of the father outperforms the other two estimations.

In order to check graphically the adequacy of the GB2 models to the data sets, we have obtained the corresponding PP-plots for each set of data. Figure 4 corresponds to the plots of the empirical CDF of earnings versus the theoretical cumulative distribution from GB2b model (with $q0$ fixed), considering fathers years of schooling for each type: Spanish males (X_1); foreign males (X_2); Spanish females (X_3) and foreign females (X_4), in the year 2005.[6]. The results reveal an outstanding performance of the GB2 model in terms of goodness of fit. The accuracy is specially high for the types Spanish males and Spanish females.

The results of inequality decomposition for the two parametric functional forms considered in this study are presented in Table 5. We include the four

[6]The rest of the PP-plots (with $p0$ fixed, and also considering mothers years of schooling or the maximum years of education in the couple), are available upon request

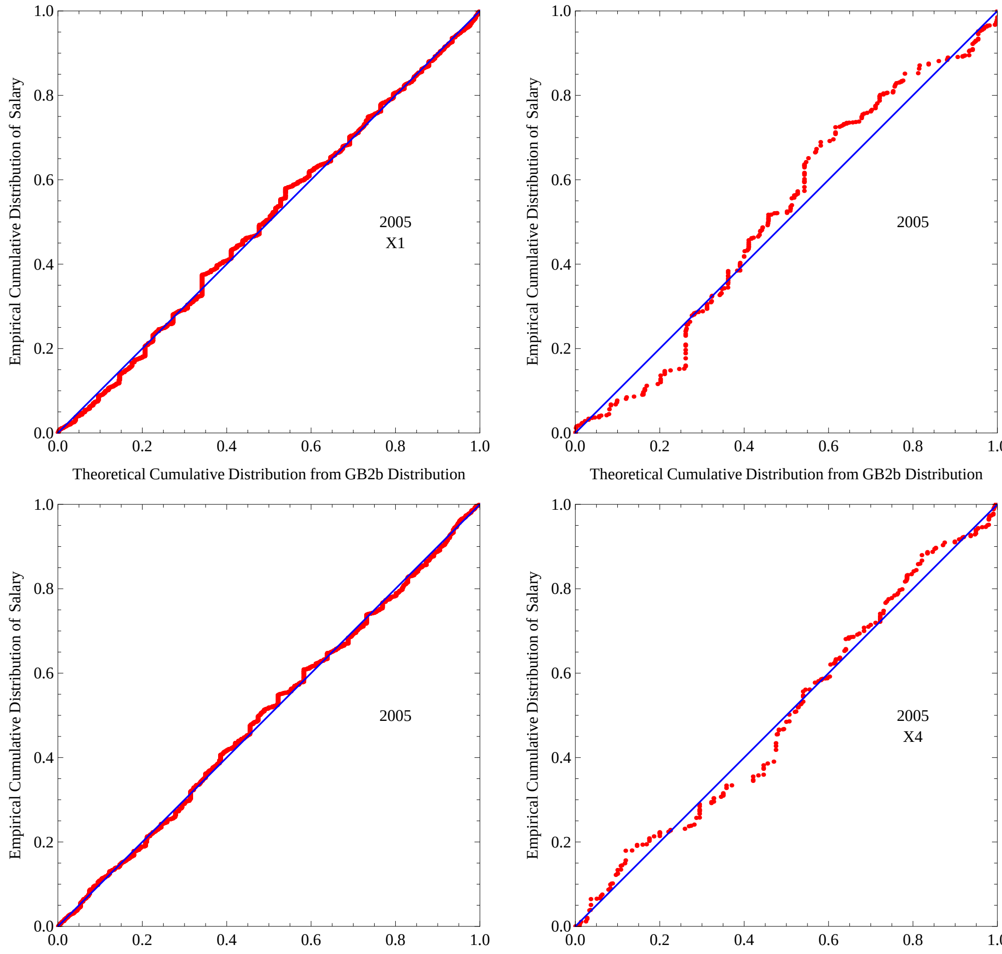
Table 2: Estimation of the generalized beta distribution $\mathcal{GB}2(a, p_0, q, b)$ (p_0 fixed) according to the four groups X_1, X_2, X_3 and X_4 and the three education classifications. Parameter estimates, standard errors (in parentheses) and AIC and BIC statistics. Year 2005

	Father Education			
	X_1	X_2	X_3	X_4
a	3.33198	6.03276	1.80824	2.94472
$STD(a)$	(0.00247)	(0.01213)	(0.0014)	(0.00632)
q	0.99560	0.435511	3.21570	1.31758
$STD(q)$	(0.00158)	(0.00133)	(0.01112)	(0.00699)
b	17701.0	12021.8	25614.0	10557.5
$STD(b)$	(11.4531)	(8.06013)	(65.9501)	(26.0693)
p_0	1.47485	1.47485	1.47485	1.47485
AIC	$1.116 \cdot 10^8$	$1.330 \cdot 10^7$	$7.745 \cdot 10^7$	$9.407 \cdot 10^6$
BIC	$-5.583 \cdot 10^7$	$-6.651 \cdot 10^6$	$-3.873 \cdot 10^7$	$-4.704 \cdot 10^6$
	Mother Education			
	X_1	X_2	X_3	X_4
a	3.05599	5.56383	1.62810	2.75148
$STD(a)$	(0.00228)	(0.01086)	(0.00128)	(0.00578)
q	1.11573	0.48552	3.98653	1.51699
$STD(q)$	(0.00182)	(0.00146)	(0.01509)	(0.00814)
b	17361.1	11849.1	28073.8	10718.6
$STD(b)$	(11.8464)	(7.91982)	(85.5624)	27.3864
p_0	1.70240	1.70240	1.70240	1.70240
AIC	$1.138 \cdot 10^8$	$1.393 \cdot 10^7$	$7.947 \cdot 10^7$	$1.053 \cdot 10^7$
BIC	$-5.688 \cdot 10^7$	$-6.965 \cdot 10^6$	$-3.974 \cdot 10^7$	$-5.263 \cdot 10^6$
	Combined			
	X_1	X_2	X_3	X_4
a	3.34048	5.99607	1.79763	3.03034
$STD(a)$	(0.00245)	(0.01175)	(0.00138)	(0.00621)
q	0.98767	0.44639	3.24677	1.31537
$STD(q)$	(0.00155)	(0.00133)	(0.01109)	(0.00662)
b	17627.1	12180.7	25779.0	10736.9
$STD(b)$	(11.1699)	(8.03352)	(65.9045)	(24.3667)
p_0	1.47887	1.47887	1.47887	1.47887
AIC	$1.152 \cdot 10^8$	$1.411 \cdot 10^7$	$8.024 \cdot 10^7$	$1.075 \cdot 10^7$
BIC	$-5.759 \cdot 10^7$	$-7.056 \cdot 10^6$	$-4.012 \cdot 10^7$	$-5.377 \cdot 10^6$
Sample size	4448	286	3309	238

Table 3: Estimation of the generalized beta distribution $\mathcal{GB}2(a, p, q_0, b)$ (p_0 fixed) according to the four groups X_1, X_2, X_3 and X_4 and the three education classifications. Parameter estimates, standard errors (in parentheses) and AIC and BIC statistics. Year 2005

	Father Education			
	X_1	X_2	X_3	X_4
a	2.67179	2.40897	2.67861	2.76935
$STD(a)$	(0.00157)	(0.00358)	(0.00247)	(0.00683)
p	1.89199	7.26929	0.87302	1.58725
$STD(p)$	(0.00333)	(0.06962)	(0.00167)	(0.01070)
b	18580.3	7920.2	20122.5	10732.7
$STD(b)$	(16.7252)	(40.2710)	(19.5431)	(35.8548)
q_0	1.47485	1.47485	1.47485	1.47485
AIC	$1.117 \cdot 10^8$	$1.337 \cdot 10^7$	$7.746 \cdot 10^7$	$9.407 \cdot 10^6$
BIC	$-5.585 \cdot 10^7$	$-6.683 \cdot 10^6$	$-3.873 \cdot 10^7$	$-4.704 \cdot 10^6$
	Mother Education			
	X_1	X_2	X_3	X_4
a	2.43584	2.19008	2.46942	2.61408
$STD(a)$	(0.00143)	(0.00322)	(0.00229)	(0.00598)
p	2.17530	9.76710	0.96155	1.77616
$STD(p)$	(0.00394)	(0.10902)	(0.00187)	(0.01128)
	18517.1	6969.9	20916.3	11086.4
$STD(b)$	18.2028	(44.3017)	21.0745	35.5396
q_0	1.70240	1.70240	1.70240	1.70240
AIC	$1.138 \cdot 10^8$	$1.400 \cdot 10^7$	$7.947 \cdot 10^7$	$1.053 \cdot 10^7$
BIC	$-5.690 \cdot 10^7$	$-6.999 \cdot 10^6$	$-3.974 \cdot 10^7$	$-5.264 \cdot 10^6$
	Combined			
	X_1	X_2	X_3	X_4
a	2.65758	2.45217	2.67249	2.85975
$STD(a)$	(0.00154)	(0.00354)	(0.00242)	(0.00633)
p	1.92142	6.75630	0.87083	1.56561
$STD(p)$	(0.00335)	(0.05954)	(0.00163)	(0.00935)
b	18463.3	8352.4	20178.3	11016.8
$STD(b)$	(16.5382)	(38.6060)	(19.2018)	(31.5152)
q_0	1.47887	1.47887	1.47887	1.47887
AIC	$1.152 \cdot 10^8$	$1.418 \cdot 10^7$	$8.024 \cdot 10^7$	$1.075 \cdot 10^7$
BIC	$-5.761 \cdot 10^7$	$-7.089 \cdot 10^6$	$-4.012 \cdot 10^7$	$-5.377 \cdot 10^6$
Sample size	4448	286	3309	238

Table 4: PP-plot: Empirical Cumulative Distribution of Salary versus Theoretical Cumulative Distribution from GB2b model (q_0 fixed), considering fathers years of schooling, for Spanish males (X1); foreign males (X2); Spanish females (X3) and foreign females (X4). Year 2005.



components of Eq. (10): IE denotes inequality of efforts which would correspond to the within-group inequality of the specific component. IO(1) is the between-type earnings inequality due to parental education, IO(2) informs about the amount of inequality driven by mean differences in parental education among the four types considered in this study and IO(3) is the part of inequality driven from differences in earnings between types. The results point out that the proportion of inequality due to individuals responsibility varies between 33 to 38 percent. This proportion seems to be rather insensitive to the variable used to measure parental education, specially for the model given in (4). The main contribution to inequality of opportunity is given by differences in parental education which account for half of the total inequality of earnings. Finally, inequality of opportunity due to differences in gender and country of birth represent a residual part which varies between 7 and 10 percent depending on the model considered.

5 Conclusions

Within the framework of inequality of opportunity, disparities driven by factors beyond the individuals' control are ethically unfair. On the other hand, inequality that comes from factors that individuals are responsible for are socially accepted. According to this distinction, The distribution of individual achievements can be decomposed into two different factors, namely circumstances and efforts, which would lead to two different components of inequality of the overall outcome of interest: inequality of efforts and inequality of opportunity. In this paper we present an alternative parametric model, that relies on a flexible functional form to model the joint distribution of circumstances and efforts. The model avoids the need to break continuous variables of circumstances into groups, while also allowing for the use of categorical variables. Moreover, the proposed methodology provides us with an statistical framework that explicitly addresses the dependence between circumstances and advantages.

In this paper, we have focused on the distributions of earnings in Spain in 2005, considering three exogenous circumstances: gender, country of birth and parental education. Our results point out that the proportion of inequality of efforts varies between 33 to 38 percent. This proportion is rather robust to the consideration of different variables to measure parental education. Differences in parental education which account for half of the total

Table 5: Decomposition of the inequality in inequality of efforts (IE) and inequality of opportunity (IO) based on the Theil 0 index, according to three different definitions of parental education. The models $\mathcal{GB}2(a, p_0, q, b)$ and $\mathcal{GB}2(a, p, q_0, b)$. Year 2005.

Based on the $\mathcal{GB}2(a, p_0, q, b)$ model						
	Father Education		Mother Education		Combined	
Component	Index value	%	Index value	%	Index value	%
IE	0.0788	38.70	0.0772	33.05	0.0787	38.38
IO (1)	0.0701	34.43	0.0712	30.46	0.0701	34.16
IO (2)	0.0338	16.57	0.0591	25.28	0.0350	17.05
IO (3)	0.0210	10.30	0.0262	11.21	0.0214	10.41
Based on the $\mathcal{GB}2(a.p.q_0.b)$ model						
	Father Education		Mother Education		Combined	
Component	Index value	%	Index value	%	Index value	%
IE	0.0694	37.67	0.0706	37.01	0.0693	37.92
IO (1)	0.0769	41.71	0.0749	39.24	0.0767	41.93
IO (2)	0.0241	13.09	0.0312	16.37	0.0228	12.49
IO (3)	0.0139	7.53	0.0141	7.38	0.0140	7.67

inequality of earnings. Finally, inequality of opportunity driven by the discrete circumstances, gender and country of birth, represent a residual part which varies between 7 and 10 percent depending on the model considered.

Appendix

Parametric distributions

The Gamma and the Generalized Gamma distributions

A random variable X is said to have a gamma distribution if its PDF is given by $(\alpha > 0)$,

$$f_X(x) = \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}, \quad x > 0,$$

and $f_X(x) = 0$ if $x < 0$. We will represent $X \sim \mathcal{G}(\alpha)$.

Now, the generalized gamma distribution (GG) is defined as $Y = bX^{1/a}$, where $X \sim \mathcal{G}(p)$ and $a, b, p > 0$. The PDF of Y is,

$$f_Y(y) = \frac{ay^{ap-1}e^{-(y/b)^a}}{b^{ap}\Gamma(p)}, \quad y > 0, \tag{16}$$

and $f_Y(y) = 0$ if $y < 0$. A GG distribution will be denoted by $Y \sim \mathcal{GG}(a, p, b)$.

If $Y \sim \mathcal{GG}(a, p, b)$ with $a, p, b > 0$ and $p + r/a > 0$ then,

$$\mathbb{E}(Y^r) = b^r \cdot \frac{\Gamma(p + \frac{r}{a})}{\Gamma(p)}.$$

The Generalized Beta of the Second Kind Distribution

The Generalized Beta of the Second Kind (GB2) distribution is defined in terms of the PDF,

$$f_X(x; a, p, q, \sigma) = \frac{a(x/\sigma)^{ap-1}}{\sigma B(p, q)\left[1 + (x/\sigma)^a\right]^{p+q}}, \quad x > 0, \tag{17}$$

where $a, p, q, \sigma > 0$, $B(s, t) = \Gamma(s)\Gamma(t)/\Gamma(s + t)$ and $\Gamma(\cdot)$ is the gamma function. The parameters a, p, q are shape parameters and σ is a scale parameter. A random variable with pdf (17) will be represented by $X \sim \mathcal{GB}2(a, p, q, \sigma)$.

The GB2 distributions contains important income distribution as special or limiting case. The classical Singh-Maddala distribution is obtained when $p = 1$ (Singh and Maddala, 1976), and will be represented by $\mathcal{SM}(a, q, \sigma)$; the three-parameter Dagum distribution (Dagum, 1977) corresponds to the choice $q = 1$ and will be represented by $\mathcal{D}(a, p, \sigma)$ and the second kind beta distribution is obtained by setting $a = 1$, and will be represented by

$\mathcal{B}2(p,q,\sigma)$. Fisk distribution is obtained for $p=q=1$ and classical Pareto II distribution (Arnold, 1983) for $a=p=1$. The generalized gamma distribution (McDonald, 1984) appears as limiting case setting $\sigma = q^{1/a}\tilde{\sigma}$, and $q \to \infty$. In consequence, classical gamma and Weibull distributions are also limiting cases of the GB2 distribution.

The shape parameters control the tail behavior of the model. In particular, the GB2 density is regularly varying at infinity with index $-aq-1$, and regularly varying at the origin with index $-ap-1$. The rth moment of the GB2 is,

$$\mathbb{E}(X^r) = \sigma^r \cdot \frac{B\left(p+\frac{r}{a}, q-\frac{r}{a}\right)}{B(p,q)}, \tag{18}$$

and exists when $-ap < r < aq$.

The different multivariate extensions are based on the following representations of the GB2 distribution. Let $Y_p \sim \mathcal{G}a(p)$ and $Y_q \sim \mathcal{G}a(q)$ be independent gamma random variables with shape parameters p and q, and let $a>0$. The GB2 distribution can be represented as,

$$X = \sigma \left(\frac{Y_p}{Y_q}\right)^{1/a} \sim \mathcal{GB}2(a,p,q,\sigma). \tag{19}$$

Previous stochastic representation permits to simulate samples of the GB2 distribution from independent gamma random variables.

Proof of Theorem 1

If we denote $Y_0 \sim \mathcal{G}(p_0)$ and $Y_i \sim \mathcal{G}(q_i)$, $i=1,2,\ldots,P$, the joint CDF is given by,

$$\begin{aligned}
& \Pr\{X_i \le x_i;\ 1 \le i \le P\} = \\
= & \Pr\left\{\sigma_i \left(\frac{Y_0}{Y_i}\right)^{1/a_i} \le x_i;\ \ 1 \le i \le P\right\} \\
= & \int_0^\infty \Pr\left\{\sigma_i \left(\frac{Y_0}{Y_i}\right)^{1/a_i} \le x_i;\ \ 1 \le i \le P | Y_0 = y_0\right\} dF_{Y_0}(y_0) \\
= & \int_0^\infty \prod_{i=1}^{P} G_{Y_i}\left\{y_0 \left(\frac{x_i}{\sigma_i}\right)^{-a_i}\right\} dF_{Y_0}(y_0),
\end{aligned}$$

where $G_{Y_i}(\cdot)$ represents the survival function of the gamma distribution. Taking partial derivatives with respect z_i we obtain the joint PDF,

$$\frac{\partial^P \Pr\{X_i \leq x_i;\ 1 \leq i \leq P\}}{\partial x_1 \cdots \partial x_P} = f_{X_1,\ldots,X_P}(x_1,\ldots,x_P) =$$

$$= \int_0^\infty \prod_{i=1}^P \frac{a_i y_0}{\sigma_i} \left(\frac{x_i}{\sigma_i}\right)^{-a_i-1} f_{Y_i}\left\{y_0 \left(\frac{z_i}{\sigma_i}\right)^{-a_i}\right\} dF_{Y_0}(y_0).$$

Finally, substituting by the expressions of the PDF of the gamma random variable and integrating we obtain (3).

Proof of Theorem 2

Using previous arguments to those used in the previous section and denoting $Y_i \sim \mathcal{G}(p_i)$, $i = 1, 2, \ldots, P$ and $Y_0 \sim \mathcal{G}(q_0)$, the joint CDF is given by,

$$\Pr\{X_i \leq x_i;\ 1 \leq i \leq P\} = \int_0^\infty \prod_{i=1}^P F_{Y_i}\left\{y_0 \left(\frac{x_i}{a_i}\right)^{a_i}\right\} dF_{Y_0}(y_0).$$

Taking partial derivatives with respect $x_1, \ldots, x_P$ we obtain the joint PDF,

$$f_{X_1,\ldots,X_P}(x_1,\ldots,x_P) = \int_0^\infty \prod_{i=1}^P \frac{a_i y_0}{\sigma_i} \left(\frac{x_i}{a_i}\right)^{a_i-1} f_{Y_i}\left\{y_0 \left(\frac{x_i}{a_i}\right)^{a_i}\right\} dF_{Y_0}(y_0).$$

Substituting by the expressions of the PDF of the gamma random variables and integrating we obtain (5).

Proof of Theorem 3

The proof of Theorem 3 is based in a double decomposition of the Theil-0 index. First, we have the usual decomposition of the Theil index in within and between factors,

$$T_0(X) = \sum_{i=1}^P \pi_i T_0(X_i) - \sum_{i=1}^P \pi_i \log \frac{\mathbb{E}(X_i)}{\mathbb{E}(X)},$$

where $\mathbb{E}(X) = \sum_{i=1}^P \pi_i \mathbb{E}(X_i)$. Now, because of the Theil index does not depends of change of scale we can assume $\sigma_i = 1$ for all i. We have,

$$X_i = Y_{p_0}^{1/a_i} \times Y_{q_i}^{-1/a_i},\ \ i = 1, 2, \ldots, P$$

where $Y_\alpha \sim \mathcal{G}(\alpha)$ and then $Y_{p_0}^{1/a_i} \sim \mathcal{GG}(a_i, p_0)$ and $Y_{q_i}^{-1/a_i} \sim \mathcal{IGG}(a_i, q_i)$. Now using Lemmas (1) and (2) and decomposing each term in other two components, we completed the proof.

Acknowledgements

The authors thanks to the Ministerio de Economía y Competitividad, project ECO2013-48326-C2-2-P, for partial support of this work.

References

Arnold, B.C. (1983). *Pareto Distributions.* International Cooperative Publishing House, Fairland, MD.

Arulampalam, W., Naylor, R.A., Smith, J.P. (2004). A hazard model of the probability of medical school drop-out in the UK, *Journal of the Royal Statistical Society Series A*, **167**, 157-178.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automatic Control*, **19**, 716-723.

Balakrishnan, N., Lai, C.-D. (2009). *Continuous Bivariate Distributions*, Second Edition. Springer, New York.

Bourguignon, F., F. Ferreira, and M. Walton (2007). Equity, Efficiency and Inequality Traps: A Research Agenda, *Journal of Economic Inequality*, **5**, 235-56.

Brazauskas, V. (2002). Fisher information matrix for the Feller-Pareto distribution. *Statistics and Probability Letters*, **59**, 159-167.

Bresson, F. (2009). On the estimation of growth and inequality elasticities of poverty with grouped data. *Review of Income and Wealth*, **55**, 266-302.

Checchi, D. and V. Peragine (2010). Inequality of Opportunity in Italy, *Journal of Economic Inequality*, **8**, 429-50.

Chotikapanich, D., Griffiths, W.E., Prasada Rao, D.S., Valencia, V. (2012). Global Income Distributions and Inequality, 1993 and 2000: Incorporating Country-Level Inequality Modeled with Beta Distributions, *The Review of Economics and Statistics*, **94**, 52-73.

Cohen, G. A. (1989). On the Currency of Egalitarian Justice, *Ethics*, **99**, 906-44.

Cowell, F.A. (2000). Measurement of Inequality, in A.B. Atkinson and F. Bourguignon (eds), *Handbook of Income Distribution*, Vol. 1, Elsevier, Amsterdam, 87-166.

Dagum, C. (1977). A new model of personal income distribution: specification and estimation. *Economie Appliquée*, **30**, 413-437.

Dworkin, R. (1981). What is Equality? Part 1: Equality of Welfare, *Philosophy & Public Affairs*, **10**, 185-246.

El-Bassiouny, A.H., Jones, M.C. (2009). A bivariate F distribution with marginals on arbitrary numerator and denominator degrees of freedom, and related bivariate beta and t distributions. *Statistical Methods and Applications*, **18**, 465-481.

Fang, K.T., Kotz, S., and Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London.

Ferreira, F. and J. Gignoux (2011). The Measurement of Inequality of Opportunity: Theory and an Application to Latin America, *Review of Income and Wealth*, **57**, 622-57.

Fleurbaey, M. (2008). *Fairness, Responsibility, and Welfare*, Oxford University Press, Oxford.

Foster, J.E. and A. Shneyerov (2000). Path Independent Inequality Measures, *Journal of Economic Theory*, **91**, 199-222.

Hajargasht, G., Griffiths, W.E., Brice, J., Prasada Rao, D.S., Chotikapanich, D. (2012). Inference for Income Distributions Using Grouped Data, *Journal of Business & Economic Statistics*, **30**, 563-575.

Jenkins, S.P. (2009). Distributionally-sensitive indices and the GB2 distribution. *Review of Income and Wealth*, **55**, 392-398.

Jones, M.C. (2002). A dependent bivariate t distribution with marginals on different degrees of freedom. *Statistics and Probability Letters*, **56**, 163-170.

Jordá, V. (2015). The Global Distribution of Education: 1970-2010. *Mimeo.*

Kleiber, C., Kotz, S. (2003). *Statistical Size Distributions in Economics and Actuarial Sciences.* John Wiley, New York.

Lefranc, A., N. Pistolesi, and A. Trannoy (2008). Inequality of Opportunities vs. Inequality of Outcomes: Are Western Societies All Alike? *Review of Income and Wealth*, **54**, 513-46.

Lefranc, A., N. Pistolesi, and A. Trannoy (2008). Inequality of Opportunities vs. Inequality of Outcomes: Are Western Societies All Alike? *Review of Income and Wealth*, **54**, 513-46.

Lefranc, A., N. Pistolesi, and A. Trannoy (2009). Equality of Opportunity and Luck: Definitions and Testable Conditions, with an Application to Income in France, *Journal of Public Economics*, **93**, 1189-207.

Marrero, G., Rodriguez, J.A. (2012). Inequality of opportunity in Europe. *Review of Income and Wealth*, **58**, 597-621.

Marrero, G.A. and J.G. Rodriguez (2013). Inequality of Opportunity and Growth, *Journal of Development Economics*, **104**, 107-122.

McDonald, J.B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, **52**, 647-663.

Olkin, I., Liu, R. (2003). A bivariate beta distribution. *Statistics and Probability Letters*, **62**, 407-412.

Plank, S.B., DeLuca, S., Estacion, A. (2008). High School Dropout and the Role of Career and Technical Education: A Survival Analysis of Surviving High School. *Sociology of Education*, 81, 345-370.

Pistolesi, N. (2009). Inequality of Opportunity in the Land of Opportunities, 19682001, *Journal of Economic Inequality*, **7**, 411-33.

Rawls, J. (1971). *A Theory of Justice*, Harvard University Press, Cambridge, MA.

Roemer, J. E. (1993). A Pragmatic Approach to Responsibility for the Egalitarian Planner, *Philosophy & Public Affairs*, **10**, 146-66.

Roemer, J. E. (1998). *Equality of Opportunity*, Harvard University Press, Cambridge, MA.

Sapata, C. (2009). Inequality of Income and Inequality of Opportunity In Spain. *Working Paper*, Autonomous University of Barcelona, Spain.

Sarhan, A.M., and Balakrishnan, N. (2007). A new class of bivariate distributions and its mixture. *Journal of Multivariate Analysis*, **98**, 1508-1527.

Sarabia, J.M., Gómez-Déniz, E. (2008). Construction of multivariate distributions: a review of some recent results (with discussion). *SORT, Statistics and Operations Research Transactions*, **32**, 3-36.

Sarabia, J.M., Jordá, V., Remuzgo, L. (2015). Explicit Expressions for the Theil Indices in Parametric Families of Income Distributions. *Submitted.*

Singh, S.K., Maddala, G.S. (1976). A function for the size distribution of incomes. *Econometrica*, **44**, 963-970.

Slottje, D.J. (1990). Using grouped data for constructing inequality indices: Parametric vs. non-parametric methods. *Economics Letters*, **32**, 193-197.

Van de Gaer, D. (1993). *Equality of Opportunity and Investment in Human Capital*, PhD Dissertation, Catholic University of Leuven.