Equality of Opportunity (EOp), the Great Gatsby Curve, and Polarization

Giovanni Bernardo Steven Durlauf Andros Kourtellos Chih Ming Tan[#]

Preliminary and Incomplete. Please Do Not Circulate.

Version: *March 22, 2023*

Abstract

We formally relate a central ideal of distributive justice, equality of opportunity (EOp; Roemer (1993, 1998)), to inequality, economic mobility, and political polarization. We extend a standard family investment model (Becker and Tomes, 1986; Solon, 2004) with a public good (schooling) and majority voting over the provision of the public good to provide structural justifications for (i) the canonical EOp regression that underlies popular ex-ante approaches to the measurement of EOp, (ii) a link between EOp measures with measures of economic mobility derived from the canonical Solon IGE regression and income inequality (i.e., the "Great Gatsby Curve"), and, (iii) the impact of EOp on the redistribution preferences of the median voter. We investigate the implications of the model using data for 29 countries from the 2011 wave of EU-SILC. We argue that substantive normative disagreement in the literature regarding which factors constitute circumstances imply substantial model uncertainty in empirical specifications and address this using model averaging methods. We find that reductions in EOp are associated with higher inequality, lower economic mobility, lower public expenditures on education (as a percentage of GDP), reduced confidence in public institutions, and lower social cohesion.

I. Introduction

Following Roemer (1993, 1996, 1998), the concept of equality of opportunity (EOp) has increasingly become the central ideal of distributive justice. The EOp perspective distinguishes

[#] Bernardo: Department of Law, University of Naples "Federico II", Via Marina Nuova 22, Naples 80133, Italy. Durlauf: Harris School of Public Policy, University of Chicago, 1155 East 60th Street, Chicago, IL 60637,USA. Kourtellos: Department of Economics, University of Cyprus, P.O. Box 537, CY 1678 Nicosia, Cyprus. Tan (corresponding author): Department of Economics & Finance, Nistler College of Business and Public Administration, University of North Dakota, Nistler Hall 330P, 3125 University Ave, Stop 8369, Grand Forks, ND 58202-8369, USA. Email: chihming.tan@und.edu.
Tan thanks the Greg and Cindy Page Faculty Distribution Fund for financial support.

between the determinants of an individual's outcomes; such as income, education, or health; see, survey by Pignataro (2012), that are viewed as being under the control of the individual ("effort") and those that are not under her control ("circumstances").

The fundamental question that the EOp perspective poses is the extent to which individual agency, decision-making, choices, and effort – those things that are within the individual's control – can overcome the confines of circumstances in determining individual outcomes. The societal aggregation of these outcomes then begets notions of inequality. Inequality in outcomes is justifiable when it is the result of individuals who share the same circumstances making different choices (exerting different levels of effort). However, inequalities that derive from differences in factors that are outside of the individual's control (i.e., circumstances) are not justified and should be mitigated by or compensated for via public policy.

While this approach has deep normative appeal because of its conceptual clarity, in practice, substantial issues arise when taking the concept to data. First, the extent to which the determinants of outcomes are observable may vary across contexts, and that may determine the principles that govern public compensatory efforts. The EOp literature distinguishes between two main perspectives, namely ex-ante and ex-post.[1] From the ex-ante perspective, there is equality of opportunity if differences in circumstances do not matter for an individual's opportunity set. The focus of this perspective is therefore centered on inequality that arises from differences in circumstances and to enact policies that compensate for disadvantages that arise from such differences. Among the ex-ante perspective, Ferreira and Peragine (2016) distinguish two different approaches, "between-types inequality" (see for instance, Bourguignon et al. (2007), Checchi and Peragine (2010)) and "direct unfairness" (Fleurbaey and Schokkaert (2009); Hufe et al. (2018)).

From the ex-post perspective, there is equality of opportunity if individuals that exert the same level of effort receive the same outcomes. Compensatory policies would then need to properly identify different degrees of effort and then be targeted at equalizing the outcomes of individuals exerting the same degree of effort. The ex-post approach can be divided into "within-tranches inequality" that is applied, for instance in Checchi and Peragine (2010) and Aaberge et al. (2011), while the "unfair inequality" is proposed by Fleurbaey and Schokkaert (2009). It is clear that the ex-ante and ex-post perspectives have different information requirements. The ex-ante perspective only requires the researcher to observe circumstance variables, but the ex-post perspective requires both circumstances and effort variables to be observed. The extent to which either case can be fulfilled depends on the context (and data quality). Like much of the empirical literature, this paper will focus

[1] See, Brunori et al. (2013), Ferreira and Peragine (2016), and Roemer and Trannoy (2016) for excellent surveys of the recent literature.

on an applied example from the ex-ante perspective. Nevertheless, the methodological critiques that follow apply equally to both cases.

The first contribution of this paper is to nest the discussion of EOp within a standard family investment model (Becker and Tomes, 1986). Doing so achieves a number of objectives. First, we demonstrate that the canonical approach that underlies popular ex-ante approaches to the measurement of EOp – one that relies on a projection of outcomes on circumstances – can be structurally justified within the set of standard assumptions that also delivers the canonical linear intergenerational elasticity of earnings (IGE) regression. The literature has utilized the IGE regression to provide estimates for the intergenerational persistence of earnings outcomes and as such has been interpreted as a measure of economic mobility.

As highlighted in Durlauf et al. (2022), such measures of economic mobility are mechanically related to standard measures of cross-sectional earnings inequality. This visual relationship between economic mobility and inequality has been described as the Great Gatsby Curve (GGC; Krueger (2012); Corak (2013)) and has featured prominently in policy debates that center on, for example, the nature of meritocracy in the United States and elsewhere. Our second objective is, therefore to add to this debate by explicitly relating concepts of social justice (EOp) with the Great Gatsby Curve.

Finally, we consider an extension of the family investment model to include a public good (public schooling; Solon (2004)) and consider a standard approach for determining the provision of the public good (majority voting). Because our data has comprehensive information on parental characteristics, we also allow for assortative mating in the model. Doing so allows us to examine the impact of EOp on the redistribution preferences of the median voter, and thereby address issues related to the potential polarization of such preferences.

The second contribution of this paper is methodological and concerns the proper estimation of measures of EOp. We point out that a now standard critique of empirical work appears to be particularly relevant to the EOp context. A key issue with the standard approach to estimating measures of EOp (or, more accurately, inequality of opportunity) ignores model uncertainty and bases inference on a single model with a particular set of circumstances, assuming that the model was correctly specified. This is a particularly unrealistic assumption in the EOp context because there exists substantive normative disagreement in the literature regarding which factors constitute circumstances and which constitute effort. As Ferreira and Peragine (2016; pp. 5) quoting Barry (1991) point out:

> "In general, a conservative (or "right-wing") interpretation of the theory tends to limit the scope of circumstances, and attribute much to individual choice, whereas a

> progressive (or "left-wing") interpretation tends to recognize the importance of individual responsibility in principle, but to downplay the ability of individuals to make real choices: in this view, individual choices tend be considered as completely determined by social circumstances (Barry, 1991)."

And, it may well be the case that these disagreements are entirely reasonable or even founded in theory. Take, for instance, the following suggestion by Roemer and Trannoy (2016)[2]:

> "…if one views all actions of the child as due to either nature or nurture, both of which are beyond his/her control, by hypothesis, before the age of consent, then one should simply take the child's educational accomplishments at the age of consent as a circumstance with respect to determining outcomes in later life."

They argue that a child should not be able to claim accomplishments (positive or otherwise) below a certain age (18 years being the age of consent), and that all such factors should comprise the child's circumstances. While standard models in the social interactions literature (see, Durlauf and Ioannides (2010) for a recent survey) that seek to identify neighborhood or peer group effects explicitly posit interdependence between circumstances (the choices made by members of one's social group) and individual decision-making, the literature in no way suggests a complete lack of agency on the part of the individual for her outcomes. Instead, the dependencies generate probabilistic outcomes, not deterministic ones driven by either nature or nurture as suggested above. Not all children in disadvantaged environments choose to drop out of school, engage in teen pregnancy, etc. The assertion of determinism, in our view, unjustly downplays the often courageous exertion of individual agency in overcoming the true underlying circumstances that are present in determining things like educational accomplishments.

It is unclear if disagreements over which factors that affect an individual's outcomes are within their control and which are not (and to what extent) will ever be satisfactorily resolved. Methodologically, what these disagreements highlight is a need to obtain estimates for EOp that are sensitive to differences in normative perspectives. We should note that even if these differences could be conceptually resolved, there remains the issue in applied work that each of these high-level circumstances may potentially be proxied by multiple measures, and the choice of which measures to employ and include (or exclude) in any given empirical exercise may deliver (potentially sizeable)

[2] Hufe et al. (2017) similarly argue that a child should not be held responsible for any of their accomplishments below 16 years of age.

differences in estimates for inequality of opportunity. In the limit, estimates for inequality of opportunity based on any given ad hoc model may even be potentially "fragile" (in the sense of Leamer (1983)).

Our overriding contention is that the above issues rooted in the substantial normative disagreements over which factors constitute circumstances imply the existence of model uncertainty in the estimation of measures of inequality of opportunity. The first contribution of this paper is therefore to highlight the need to obtain estimates for inequality of opportunity that are robust to model uncertainty. We propose to address systematically the issue of model uncertainty using Bayesian model averaging approaches (e.g., Raftery et al. (1997)).[3] Unlike the standard approach that conditions on a single model, model averaging approaches construct estimates conditional on model space which is composed of the set of all possible combinations of all possible predictors suggested by all the relevant theories of equality of opportunity. The idea of model averaging is to "integrates out" the uncertainty over models by computing the average of model-specific estimates weighted by the posterior model probabilities that capture the evidentiary support of each model given the data.

In our application, we further envision organizing the set of observable factors into classes of variables according to particular *theories* of outcome determinacy; e.g., respondent's (direct) characteristics, family structure, parental characteristics, and general socioeconomic status. We implement a hierarchical model prior structure to query the extent to which each of these theories are responsible for inequality of opportunity in outcomes. As we detail in Section III, careful handling of such considerations will be important in our proposed approach to deriving robust answers to such questions. We investigate the implications of the model using data for 29 countries from the 2011 wave of the European Union Statistics on Income and Living Conditions (EU-SILC). We find that reductions in EOp are associated with higher inequality, lower economic mobility, lower public expenditures on education (as a percentage of GDP), reduced confidence in public institutions, and lower social cohesion.

The rest of the sections of the paper are organized as follows. We present and discuss the implications of a Becker-Tomes family investment model with public schooling and assortative mating in Section II. We discuss our Bayesian model averaging approach to generating robust estimates for inequality of opportunity in Section III. We describe the EU-SILC data we employ in Section IV and present our findings in Section V. Section VI concludes.

[3] Model averaging approaches have been extensively applied in the economic growth literature; see, for example, Fernandez et al. (2001) and Sala-i-Martin et al. (2004).

II. Family Investment Model with Public Schooling and Assortative Mating

II.1 Model Setup

Following Solon (2004), we include taxes and public investment in schooling in the standard Becker-Tomes family investment model. We further modify this model by allowing for assortative mating in family formation.

In this economy, there exists I family dynasties; denote by i = 1, 2, …, I. Each individual within a dynasty lives for just 2 periods: (i) childhood (c), and (ii) adulthood (a). The index t denotes the time of birth; i.e., denotes the "birth cohort" that each individual belongs to. A mating model would require a pairing for parents and for each household to produce two kids for population growth to be stable. For analytical convenience, we will assume that in each family, there are two parents and one offspring. This assumption is mostly without loss of generality in the sense that the outcome would be the same if the dynasty produced two observationally equivalent (to the parents) children in each generation.

Parents pay taxes that are proportional to their combined household income to fund public investment in their offspring's human capital. The tax rate is given by τ. From the point of view of parents, the level of public investment that accrues to their child is taken as exogenously given. The household budget constraint is therefore given by

$$(1-\tau)Y_{i,a,t}^{HH} = C_{i,a,t} + I_{i,c,t} \qquad (1)$$

where household income, $Y_{i,a,t}^{HH}$, is defined as the sum of the log of the individual parental incomes; i.e., i's income, $Y_{i,a,t}$, and her partner's, $Y_{i,a,t}^{p}$ (the superscript "p" denotes the "partner" of the offspring):

$$\log\left(Y_{i,a,t}^{HH}\right) \equiv \log\left(Y_{i,a,t}\right) + \log\left(Y_{i,a,t}^{p}\right) \qquad (1')$$

Parents allocate disposable household income towards their consumption, $C_{i,a,t}$, and private investments in their child, $I_{i,c,t}$.

Educational outcomes for the child are then determined by a human capital production function that employs both private investments by the parents as well as public investment ($G_{i,c,t}$):

$$E_{i,c,t} = \theta \log\left(I_{i,c,t} + G_{i,c,t}\right) + \zeta_{i,c,t} + \varepsilon_{i,c,t} \quad (2)$$

where $\zeta_{i,c,t}$ is an index that captures the impact of observed *Circumstances* heterogeneity across children on human capital outcomes. Such circumstances could take the form of race/ethnicity, gender/gender identity, religiosity, family structure, etc. The error term, $\varepsilon_{i,c,t}$, could be interpreted as unobserved *Effort* and/or capture other latent abilities of the child.[4]

The offspring's income (when she becomes an adult) depends on the human capital she accumulated as a child as well as unobserved *Effort* and *Luck* (Wagstaff and Kanbur (2015)):

$$\log\left(Y_{i,a,t+1}\right) = c + \widetilde{w} E_{i,c,t} + l_{i,a,t+1} \quad (3)$$

where $\widetilde{w}$ is the returns to education (real wage rate) and $l_{i,a,t+1}$ captures idiosyncratic unobserved *Effort* and *Luck* (both "option luck" and "brute luck").

Finally, to get to the next generation, the offspring forms a family by assortatively mating based on educational outcomes[5] so that the expected human capital of her partner is given by,

$$E_t(E^p_{i,c,t} | E_{i,c,t}) = \alpha_{AM} E_{i,c,t} + (1 - \alpha_{AM})\, \bar{E}_t \quad (4)$$

where $0 \leq \alpha_{AM} \leq 1$ is the (strength of) *assortative mating* parameter. Note that, with assortative mating, parental investments in their offspring not only improve the latter's individual incomes by improving their human capital, but, will also potentially improve their household incomes by securing for them a better (more highly educated) match/mate.

Analogous to (1'), the household income for the offspring is given by

$$\log\left(Y^{HH}_{i,a,t+1}\right) \equiv \log\left(Y_{i,a,t+1}\right) + \log\left(Y^{p}_{i,a,t+1}\right)$$

---

[4] Solon (2004) allows for the possibility that latent abilities exhibit intergenerational persistence, so that,

$$\varepsilon_{i,c,t+1} = \delta + \lambda \varepsilon_{i,c,t} + \upsilon_{i,c,t+1}$$

where $0 < \lambda < 1$ and $\upsilon_{i,c,t+1}$ is white noise. We suppress such persistence for the moment because the main complications it raises is in the evaluation of the linear projection of intergenerational incomes. We will discuss this issue later in the text.

[5] The specification for assortative mating based on human capital follows Holmlund (2008). Using data from the Health and Retirement Study (HRS), Domingue et al. (2014) provides evidence for assortative mating based on education attainment as opposed to genes.

The expected (at time *t*) household income for the child is therefore,

$$E_t\left(Y_{i,a,t+1}^{HH}|E_{i,c,t}\right) = 2c + \widetilde{w}(1+\alpha_{AM})E_{i,c,t} + \widetilde{w}(1-\alpha_{AM})\,\bar{E}_t \qquad (5)$$

II.2. Parents' Decision Problem

The decision problem of parents is to maximize utility given by

$$U_{i,a,t} = (1-\pi)\log\left(C_{i,a,t}\right) + \pi E_t\left(\log\left(Y_{i,a,t+1}^{HH}\right)\right) \qquad (6)$$

subject to the budget constraint (1) given their knowledge of (2) and (5), and where $0 \leq \pi \leq 1$ is an altruism parameter.

Proposition 1: The optimal allocations for parental investments in offspring and consumption are then given by

$$I_{i,c,t}^{*} = \left[\frac{\pi w\theta}{1-\pi(1-w\theta)}\right](1-\tau)Y_{i,a,t}^{HH} - \left[\frac{1-\pi}{1-\pi(1-w\theta)}\right]G_{i,c,t} \qquad (7)$$

$$C_{i,c,t}^{*} = \left[\frac{1-\pi}{1-\pi(1-w\theta)}\right](1-\tau)Y_{i,a,t}^{HH} + \left[\frac{1-\pi}{1-\pi(1-w\theta)}\right]G_{i,c,t} \qquad (8)$$

where $w = w(\alpha_{AM}) = \widetilde{w}(1+\alpha_{AM})$. Hence, optimal private investments increase with parental altruism, with the efficiency of human capital investment, and decreases with public investment.

II.3. Equilibrium Law-of-Motion for Income

In the aggregate, therefore, the optimal allocations from parental decisions imply an equilibrium law of motion for household income that is given by

$$E_t\left(\log\left(Y_{i,a,t+1}^{HH}\right)\right) = 2c + w\theta\log\left\{\left[\frac{\pi w\theta}{1-\pi(1-w\theta)}\right]\left((1-\tau)Y_{i,a,t}^{HH} + G_{i,c,t}\right)\right\} + w\zeta_{i,c,t} + \widetilde{w}(1-\alpha_{AM})\,\bar{E}_t + w\varepsilon_{i,c,t} \quad (9)$$

As in Solon (2004), we assume that public investments in children's human capital are made according to the following rule:

$$\frac{G_{i,c,t}}{(1-\tau)Y_{i,a,t}^{HH}} = \kappa - \gamma\log\left(Y_{i,a,t}^{HH}\right) \quad (10)$$

where the progressivity parameter is given by $\gamma > 0$.[6]

Proposition 2: The equilibrium intergenerational income process is given by

$$\log\left(Y_{i,a,t+1}^{HH}\right) = \eta + \rho t + \beta\log\left(Y_{i,a,t}^{HH}\right) + w\zeta_{i,c,t} + \varepsilon_{i,a,t+1} \quad (11)$$

where the regression error $\varepsilon_{i,a,t+1}$ captures both unobserved *Effort* and "*Luck*" (cannot be separately identified), $\eta = \left\{2c + w\theta\kappa + w\theta\log\left[\frac{\pi w\theta}{1-\pi(1-w\theta)}\right] + w\theta\log(1-\tau)\right\}$ is a constant term, the time trend captures $\widetilde{w}(1-\alpha_{AM})\,\bar{E}_t$, and $\beta = w(1-\gamma)\theta = \widetilde{w}(1+\alpha_{AM})(1-\gamma)\theta$.

The canonical IGE regression in the Great Gatsby Curve literature is a linear projection of household income of one generation, $Y_{i,a,t}^{HH}$, on that of the next, $Y_{i,a,t+1}^{HH}$. The projection parameter, $\hat{\beta}$; i.e., the intergenerational elasticity of household earnings (IGE), is then interpreted as a measure of intergenerational mobility.

Circumstances therefore influence measured economic mobility to the extent that they are correlated with the offspring's socioeconomic status (as proxied by $Y_{i,a,t}^{HH}$) and to the greater degree when the returns to education are higher. Recall also that $w = w(\alpha_{AM}) = \widetilde{w}(1+\alpha_{AM})$; hence, the assortative mating parameter, α_{AM}, functions opposite to the progressivity parameter, γ. Just as

[6] Like Solon (2004), we assume (for analytical convenience) that (10) is small.

increased progressivity increases economic mobility (i.e., decreases β), increased assortative mating decreases economic mobility (i.e., increases β).

II.4. Connection with EOp Regressions

The standard linear EOp specification employed by the ex-ante compensation approach for estimating inequality of opportunity is based on a linear projection of outcomes on circumstances (Roemer (1993, 1996, 1998)). The set of significant circumstances is then employed to define *types* based on all combinations of the support of those circumstance variables. The distribution of outcomes within types then defines the ex-ante *opportunity set* faced by individuals under those circumstances. An *inequality index* is then constructed to describe differences/inequality in opportunity sets faced by different types of individuals in the economy. As discussed in Ferreira and Peragine (2016), the ordinal approach is to employ concepts of stochastic dominance to compare the *type*-dependent distribution of outcomes/opportunity sets that different types of individuals face. The more popular cardinal approach, which we focus on in this paper, employs a summary statistic (in this case, the conditional mean) to characterize the opportunity set for each type of individuals.

Proposition 3: The family investment model rationalizes the standard linear EOp specification. Lagging (3) by one period for the expression for parent's income, and then substituting into (11) above, we obtain

$$\log\left(Y_{i,a,t+1}^{HH}\right) = \tilde{\eta} + \rho t + \tilde{\rho}(t-1) + w\beta E_{i,c,t-1} + w\zeta_{i,c,t} + \nu_{i,a,t+1}, \quad (12)$$

where $\nu_{i,a,t+1}$ is a regression error. Hence, household income of the offspring depends on the following circumstances characteristics: the educational attainment of the parents and own predetermined characteristics.

II.5. Political Economy: Preference for Redistribution

Suppose that parents vote on the progressivity of the public school system (*public policy* of interest, in this case). How would we characterize their preferences? By substituting the optimal allocations (7) and (8) into the objective function (6), we can express the optimal utility of parents in terms of the progressivity parameter; i.e., $U_{i,a,t}^{*}(\gamma)$. The optimal – i.e., preferred – level of

progressivity for parents (i.e., *redistribution preference*) is then obtained by optimizing $U^*_{i,a,t}(\gamma)$ with respects to γ yielding

$$\gamma^*_{i,a,t} = \frac{(1+\kappa)\pi w\theta+(1-\pi)}{\pi w\theta \cdot \log\left(Y^{HH}_{i,a,t}\right)} \tag{13}$$

Unsurprisingly, as parental income increases, the household optimally prefers a lower level of progressivity. Hence, the *distribution* of redistribution preferences will depend on the cross-sectional distribution of household incomes (i.e., inequality).

If the realized level of progressivity is determined via majority voting, then, the focus would need to rest on the redistribution preferences of the median voter. Analytical solutions for characterizing the redistribution preferences of the median voter are not generally obtainable. However, we offer a stylized example to provide intuition.

Proposition 4: We first express (13) above as

$$Z^*_{i,a,t} = \Psi.\log\left(Y^{HH}_{i,a,t}\right) \tag{14}$$

where $Z^*_{i,a,t} = \left(\gamma^*_{i,a,t}\right)^{-1}$ is the inverse of the progressivity parameter; i.e., higher values mean lower progressivity, and $\Psi = \frac{\pi w\theta}{(1+\kappa)\pi w\theta+(1-\pi)}$. We next suppose that the distribution of $\log\left(Y^{HH}_{i,a,t}\right)$ is given by a mixture of Normal distributions with different means and standard deviations:

$$F_Y = p.\Phi_A + (1-p)\Phi_B$$

where (with loss of generality) we set $p = 0.5$. Then, the median value, m, for the inverse of redistribution preferences $Z^*_{i,a,t}$ is given by:

$$m = \Psi.\left(\frac{\mu_A\sigma_B+\mu_B\sigma_A}{\sigma_A+\sigma_B}\right) = \left(\frac{\pi w\theta}{(1+\kappa)\pi w\theta+(1-\pi)}\right)\left(\frac{\mu_A\sigma_B+\mu_B\sigma_A}{\sigma_A+\sigma_B}\right) \tag{15}$$

where $\Phi(.)$ is the standard Normal distribution.

Therefore, from (15), we find that the median voter's preferences, m, will be determined not just by the "polarization" parameters from the income distribution; i.e., μ_j and σ_j for $j = A$ and B, but also the structural parameters in $\Psi = \frac{\pi w \theta}{(1+\kappa)\pi w\theta + (1-\pi)}$, including the assortative mating parameter, α_{AM}, since $w = w(\alpha_{AM}) = \tilde{w}(1 + \alpha_{AM})$. It can be shown that as α_{AM} increases, m increases so that an increase in assortative mating increases the *inverse* of the voted upon progressivity parameter, and therefore reduces realized progressivity. Also, writing $\mu_B = \mu_A + \Delta$ so that Δ is a measure of income *polarization* (i.e., how far the means of the component populations are apart) we see from (15) that as Δ increases, m increases, and therefore realized progressivity is reduced.

II.6. Discussion

In summary, the model presented makes several predictions regarding economic mobility, inequality, redistribution preferences, and public goods provision. Equation (11) suggests that both progressivity and circumstances play direct roles in determining economic mobility, while equation (12) demonstrates that the distribution of circumstances influences the distribution of household income, or inequality. Additionally, equation (13) predicts that increased polarization in income distribution will lead to greater polarization in redistribution preferences. Finally, equation (15) shows that greater polarization over redistribution preferences could potentially lead to lower progressivity in public goods provision. Together, these equations demonstrate the complex interplay between factors such as economic mobility, inequality, and public policy preferences.

III. Methodology

We now turn to the empirical implementation of the ex-ante compensation approach for estimating inequality of opportunity. As briefly discussed above, the standard parametric approach (Bourguignon et al., 2007; Ferreira and Gignoux, 2011) is to analyze a cross-sectional linear projection that links child's income as an adult Y_i with a p-dimensional vector of circumstances $C_i = \left(C_{1i}, C_{2i}, \dots, C_{pi}\right)'$

$$Y_i = C_i'\beta + e_i, \tag{16}$$

where e_i is assumed to include the effort (and luck) and $i = 1,2, \dots, N$. Following Roemer (1993, 1996, 1998), effort e_i is treated as unobserved.

Based on the realizations of individual circumstances C_i, we can partition the sample into M mutually exclusive types $T_1, T_2, \dots, T_M$ where observations i and j are members of the same type if they share the same circumstances. In particular, $C_i = C_j, \forall i, j | i \in T_m, j \in T_m, \forall m$. Further, the distribution of the outcome Y is then given by $F(y) = \sum_{m=1}^{M} \pi_m F_m(y)$ where π_m is the population share of type T_m.

We focus on ex-ante approaches to EOp. In such approaches the opportunity set of any type, T_m, may be summarized by information on the type's outcome distribution $F_m(y)$; e.g., by the expectation of its outcomes, $E(Y_i|C_i; i \in T_m)$. Inequality of opportunity is then at a minimum when all types face opportunity sets with the same value; i.e., $E(Y_i|C_i; i \in T_m) = \mu_{Y,m} = \mu_Y$. However, if that is not the case, then an appropriate measure of inequality of opportunity (an inequality index) would need to be defined across the counterfactual distribution where each individual's outcome is represented by the value of their opportunity set.

In the next sections, we detail how we estimate robust Inequality of Opportunity (IOp) indices based on the Gini coefficient (Lefranc et al., 2009) and the mean logarithmic deviation (Checci and Peragine, 2010; Ferreira and Gignoux, 2011).[7]

III.1 Bayesian Model Averaging (BMA)

As discussed in the Introduction, we propose to address systematically the issue of model uncertainty in equation (16) using Bayesian model averaging (BMA) approaches (e.g., Raftery et al. (1997)). Unlike the standard approach that conditions on a single model $\mathcal{M}_m \in \mathcal{M}$, model averaging approaches construct estimates conditional on model space $\mathcal{M}$, which is composed of the set of all possible combinations of all possible predictors suggested by all the relevant theories of equality of opportunity. The idea of model averaging is to "integrates out" the uncertainty over models by computing the average of model-specific estimates weighted by the posterior model probabilities that capture the evidentiary support of each model given the data. Model-specific estimates require the specification of parameter priors which we discuss in the next subsection.

The posterior distribution of β give the data $\wp$ is given by

$$\hat{\mu}(\beta|\wp) = \sum_{m \in \mathcal{M}} \hat{w}_m \hat{\mu}(\beta|\mathcal{M}_m, \wp) \qquad (17)$$

[7] The inequality index should satisfy two properties: (i) measured inequality should increase (decrease) if transfers are made from individuals from low (high) expected outcome types to those from high (low) expected outcome types, and, (ii) measured inequality should be unaffected if transfers are made between individuals of the same type. Given these properties, the choice of inequality indices reflects inequality aversion at different points of the distribution of opportunities. In this paper, we employ two standard choices in the literature: the Gini coefficient and the mean logarithmic deviation.

where the weight $\widehat{w}_m = \hat{\mu}(\mathcal{M}_m|\wp)$ is the empirical posterior distribution of $\mathcal{M}_m$. Each weight can be computed via the Bayes' rule as the product of the integrated likelihood of the data given a model and the prior model probability (which we discuss below). The model averaging estimator of β is given by the posterior mean

$$\hat{\beta}_{BMA} = \sum_{m\in\mathcal{M}} \widehat{w}_m \hat{\beta}_{\mathcal{M}_m,\wp}, \tag{18}$$

where $\hat{\beta}_{\mathcal{M}_m,\wp}$ is the least-squares estimator of β based on a single model $\mathcal{M}_m$. The standard way to conduct inference in the context of BMA is to reference the posterior probability of inclusion (PIP) for each regressor. The PIP is computed as the sum of the posterior probabilities of all the models that contain that variable. Following Eicher, Henn, and Papageorgiou (2011), a PIP value less than 50% implies insufficient evidence for an effect, while PIP values ranging from 50% to 75% suggest weak evidence for an effect. Similarly, PIP values ranging from 75% to 95% indicate positive evidence for an effect, while PIP values ranging from 95% to 99% imply strong evidence for an effect. Finally, PIP values ranging from 99% to 100% indicate decisive evidence for an effect. Following Sala-i Martin et al. (2004) and Durlauf (2008) we also report BMA posterior t statistics for coefficient estimates.

III.1.1 Parameter Priors

Following Fernandez, Ley, and Steel (2001) we assume that the prior distribution is given by $\mu(\beta_m|\sigma^2) = N(0, \sigma^2(gX_m'X_m)^{-1})$, where the variance of the prior distribution includes a shrinkage factor (or hyperparameter) *g* known as "Zellner's g prior". The hyperparameter *g* captures the prior belief of the econometrician on the parsimony of the model. A large *g* corresponds to the belief that the size of the true model is small; that is, many coefficients are indeed zero, while a small *g* means that a larger model is more likely to be the true model. When $g \rightarrow 0$ we obtain the least squares (LS) estimator of the full model.

Effectively, the choice of the prior distribution of β has been transformed into the simple choice of a single parameter. Different choices of *g* correspond to different prior structures. Following Fernandez, Ley, and Steel our baseline prior is based on their benchmark priors (g-BRIC), which correspond to $g = max(n, k^2)$. Following the recent literature, we investigate the robustness of this prior structure using the Unit Information Prior (g-UIP), which corresponds to $g = n$.

Following Liang et al. (2008) we also investigate the robustness of our main findings using a more flexible prior structure that assumes a hyper-prior on g; i.e., $g \sim Beta(1, (a-1)/2)$ with two choices for a: "hg-UIP" that corresponds to a = 2 + 2/n and "hg-BRIC" that corresponds to a = 2 + $2/max(n, k^2)$.

III.1.2 Model Priors

Our chosen baseline model prior is the Uniform Model Prior, which assumes that the probability of any variable being included in the true model is equally likely with a value of $\pi = 0.5$. Despite its simplicity, this prior has demonstrated superior performance as argued by Eicher, Papageorgiou, and Raftery (2011) who investigated the predictive performance of several parameter prior specifications.

Despite its superior performance in some cases, a uniform model prior may not be appropriate in all settings. As Brock and Durlauf (2001) argue, this type of prior can create a problem similar to the irrelevance of independent alternatives (IIA) in the discrete-choice literature, as it ignores interrelations between different variables. In our study, where we aim to evaluate the relative importance of different theories of spatial mobility rather than individual proxy variables, it is essential that our model priors capture non-informativeness (i.e., agnosticism) across theories. The uniform prior fails to do so, as a researcher can easily manipulate the prior weights across theories by selectively choosing "redundant" proxy variables for each theory. This issue is also related to the common statistical problem of multicollinearity, which arises when the regressors are highly correlated and can result in imprecise inference. Therefore, it is necessary to consider alternative model priors that account for the interdependencies between variables and promote agnosticism across theories in our context.

To address the issues with the uniform model prior, we conduct two sets of robustness exercises using hierarchical priors and dilution priors. Firstly, we follow the approach suggested by Brock and Durlauf (2001) and use a tree structure to construct prior probabilities. We classify the set of variables in M_j into T theories and define priors across theories and over variables within theories. The prior probability that a particular theory is included in the "true" model is assumed to be 0.5 to reflect a non-informative prior across theories. This prior specification assumes that theories are independent, meaning that the inclusion of one theory in a model does not affect the probability of another theory being included. Secondly, we employ a tessellation-defined dilution prior proposed by George (1999, 2010) to dilute the prior model probabilities of clusters of similar models. This ensures that the inclusion of similar variables does not increase the model probabilities. The dilution

prior assigns uniform probabilities to neighborhoods of models, rather than to individual models, with the neighborhoods defined by appropriate regions of the surface of a high-dimensional sphere forming Voronoi tessellations; see, Moser and Hofmarcher (2014).

IV. Data

Our empirical analysis is based on the European Union Statistics on Income and Living Conditions (EU-SILC) dataset. This dataset provides harmonized survey data on income, poverty, and living conditions, making it a valuable resource for statistics that compare income distribution and social inclusion from both cross-sectional and longitudinal perspectives. It covers a representative sample of households from 31 different European countries. We specifically used the 2011 wave of EU-SILC as it includes a form that provides detailed information on the individual characteristics related to the education, employment, and income of the respondent's parents. This wave serves as a crucial resource for our research, allowing for a comprehensive examination of the impact of parental background and household financial situation on children's income.

We focused on the child's income as an adult as the outcome variable in equation (16) above. Consistent with the literature, we used equivalized disposable household income as a proxy for the child's income, adjusted for the square root of household size. This income measure includes various sources, such as farm and business income, wages, bonuses, overtime commissions, professional practice, and income from roomers and boarders or businesses, profit from capital investments, interests, dividends, and other income. Taxes included in our analysis are regular taxes on wealth, regular inter-household cash transfers paid, tax on income, and social insurance contributions.

We examine three sets of circumstances related to the households of the respondents when they were 14 years old (see, Table 1) that then comprise the vector of circumstances $C_i = \left(C_{1i}, C_{2i}, \dots, C_{pi}\right)'$ in (16) above: (i) Father characteristics, including country of origin, citizenship, educational level, occupational status, and type of occupation; (ii) Mother characteristics, including country of origin, citizenship, educational level, and occupational status; and (iii) Household characteristics, including tenancy status, household financial situation, ability to meet end-of-month expenses, family type, number of adults and children, and number of workers.

To capture the contribution of individual characteristics and mitigate problems arising from variables lacking intrinsic ordering of categories, variables with multiple categories (e.g., level of education or employment status) were transformed into binomial dummies. Variables related to managerial position and occupation type of mothers were excluded from our analysis due to a high prevalence of missing values across several countries. Instead, following Brunori et al. (2021), we

grouped employment types into three macro-categories: high-skilled non-manual (managers, professionals, technicians, and associate professionals), low-skilled non-manual (clerical support workers, service and sales workers, armed forces occupations), and skilled manual and elementary occupations (skilled agricultural, forestry, and fishery workers, craft-related trades workers, plant and machine operators, assemblers, and elementary occupations). The unemployed, unknown, and dead were grouped separately.

In the next section, we present the results of our analysis using data from 29 out of 31 countries covered by EU-SILC. We excluded Ireland and Slovenia due to insufficient observations to implement the methodology proposed in this paper. We first present our robust estimates for the Inequality of Opportunity (IOp) index described in Section III above. We then proceed to consider a set of potential correlates with our robust IOp index including a measure of assortative mating, public education both at non-tertiary and tertiary level, trust in the government, trust of others, rule of law, lack of corruption, regulatory quality, and political instability. A detailed description of the data and sources is provided in Appendix Table A1. Summary statistics are provided in Appendix Table A2.

V. Results

V.1 Robust Inequality of Opportunity (IOp) Index

As discussed in Section 3, we apply BMA to (16) to construct measures of IOp that are robust to uncertainty in model specification. We first present results comparing the IOp indices obtained under different parameter and model prior specifications. Table 2 reports the correlation matrix between the IOp indices computed using ten model-parameter prior combinations (as described in Section 3 above); six of the specifications employ fixed parameter priors while four utilize flexible g-priors. The results show that the IOp measures are strongly correlated with one another, and there are no qualitative changes when different specifications are considered. These results show that BMA estimates for IOp indices are robust to the choice of priors. As is consistent with the literature, we select the g-BRIC as the benchmark model.

Next, we also compare our benchmark model to particular models in the model space. The idea is to see what would have happened if instead of engaging in model averaging (leading to our benchmark model), we had instead attempted model selection based on the evidentiary (posterior) weight of models. We do so under the context of the relationship between the IOp index with a measure of (outcome) inequality (the Gini coefficient). In Figure 1a, we compare results using the IOp measures based on ten model-parameter prior combinations previous discussed. In Figure 1b, we

compare the benchmark IOp index with the largest model in the model space (the "kitchen sink" model). In Figure 1c, we compare our benchmark model with the top 15 models in terms of posterior model probability in the model space. Finally, in Figure 1d, we compare results using the IOp index based on our benchmark with alternative IOp indices related to four different threshold levels of posterior model probabilities.

We affirm yet again that there were no significant differences between the relationship estimated by our benchmark model and the other models based on different model-parameter prior specifications. The top 15 models also do not show any significant differences from our benchmark model, while the kitchen sink model tends to understate the relationship between the IOp index and the Gini coefficient. Our findings therefore argue for caution in employing the canonical approach in the EOp empirical literature which is typically based on the kitchen sink model (and ignores model uncertainty). The kitchen sink model has low posterior model probability (outside the top 15 models) and would not have been selected based on the posterior evidence, but its predictions are qualitatively different from those of the top 15 models as well as the benchmark BMA model (which predictions are in agreement).

Which circumstance variables are most important in determining the benchmark IOp index? Do these vary across our set of countries? In the Appendix, Tables A.3, A.4, and A.5 are heatmaps that show the degree of PIP for the circumstances variables for each country based on the benchmark BMA specification. On average, the most important factors in determining the income level of children are the educational level of both the mother and father, as well as the type of occupation of the father, particularly if it is a non-manual high-skilled occupation. However, the coefficients show greater variability across countries, highlighting the fact that there is no single model that explains inequality of opportunity, and emphasizing once again the need to account for model uncertainty in estimating the IOp index. Displayed in Figure 2 is the geographic distribution of IOp across the 29 countries being analyzed. As anticipated, the Mediterranean and Eastern European nations exhibit the greatest levels of inequality, whereas the Nordic and Central European countries demonstrate the highest levels of opportunity equality.

V.2 Cross-country correlations

We now turn to the empirical evidence for the predictions of the model discussed in Section 2.6 above. Given the limited sample size, we show descriptive evidence in the form of bivariate correlations.

V.2.1 Great Gatsby Curve

We present evidence for the presence of the GGC. As the discussion in Section 2.6 suggests, we should find correlations between the IOp index and GGC variables. We begin our analysis by demonstrating the significant positive correlation between the IOp index and the Gini coefficient in Figure 3. This positive association shows that higher levels of cross-sectional income inequality are associated with higher degrees of inequality of opportunity. This pattern mimics the GGC since equality of opportunity promotes mobility by enabling individuals to reach their full potential, regardless of their background or circumstances. In Figures 4(a)-(c), we illustrate this positive association between the IOp index and various measures of intergenerational mobility. In particular, Figure 4(a) shows the relationship between the IOp index and the Intergenerational Elasticity of Earnings (IGE) measure by Corak (2013). The IGE, which measures the intergenerational persistence of earnings outcomes, is a standard measure of intergenerational mobility. Figure 4(b) shows the relationship between the IOp index and a measure of social mobility constructed by the World Economic Forum (2020). We observe a significant positive correlation between the IOp index and intergenerational mobility in all cases. Interestingly, the "GGC stylized fact" remains consistent across different levels of income inequality aggregation. In particular, Figure 5 illustrates a noteworthy correlation between the IOp index and the Gini coefficient based on the regional differences in household income. This suggests that as the gap between household incomes across different regions becomes more pronounced, the IOp increases.

V.2.2 Assortative Mating

Next, we consider bivariate correlations between the IOp index on the vertical axis and various determinants on the horizontal axis. In Figure 6, we examine how assortative mating is related to inequality of opportunity. As predicted in Section 2.6, countries with higher degrees of assortative mating tend to exhibit less equality of opportunity. The model predicts that higher degrees of assortative mating results in a stronger incentive for parents to invest in their children and also reduces their desire for increased progressivity in the provision of educational public goods. The result is then that the educational level of parents more strongly determines their children's outcomes in societies with higher degrees of assortative mating. This prediction is borne out in the data. The persistence of this pattern can lead to long-term consequences, perpetuating social advantages and disadvantages and reinforcing income inequality across generations (Greenwood et al. (2016) and Eika et al. (2014)).

V.2.3 Education

Figures 7(a)-(b) show a negative correlation between the IOp index and the share of GDP allocated to public education, both at non-tertiary and tertiary levels. Across countries, those with the highest levels of inequality, such as Greece, Italy, and Spain, tend to spend less on public education. Conversely, the more equal countries, such as Denmark, Norway, and Sweden, are among the highest spenders on both non-tertiary and tertiary public education. This finding affirms the prediction of the model from equation (13) and Proposition 4. The results suggest that exogenous interventions in public educational expenditures – something that may not be achieved under majority voting – may be necessary to reduce inequality of opportunity and promote social mobility. The availability of quality education to children from low-income families increases their likelihood of success and enables them to escape the constraints of poverty, leading to greater mobility. In contrast, countries with low levels of public investment in education tend to perpetuate inequality and disadvantage. Thus, prioritizing of progressivity in public education is essential for building more equal societies, where individuals have the opportunity to succeed regardless of their background, but this is not something that may be achieved in an already unequal democracy.

V.2.4 Social Trust and Confidence in Institutions

As equation (13) suggests, increasing income inequality potentially leads to polarization in redistribution preferences. Proposition 4 then implies that enacted public policies – as determined by the median voter – may be far apart from the preferences of the extremes of the distribution. Such polarization could lead to increasing distrust of fellow citizens as well as reduced confidence in public institutions. The data appears to bear out this concern. In Figures 8(a)-(b) we find negative correlations between the IOp index and trust (both in the government and interpersonally). Countries that have higher inequality of opportunity exhibit lower levels of social trust. Figures 9(a)-(d) show negative associations between the IOp index and various governance indicators and institution quality measures. Specifically, societies with higher levels of inequality of opportunity evidence weaker rule of law, higher levels of corruption, lower regulatory quality, and greater political instability.

VI Conclusion

This paper makes two primary contributions. The first contribution is to nest existing approaches to constructing ex-ante measures of inequality of opportunity within a standard family investment model. By doing so, we provide a structural rationalization for a popular approach in the EOp literature. Second, we argue that issues of model uncertainty are endemic in the EOp literature because of substantive normative disagreements over the nature of what constitute circumstances as well as how to proxy for such determinants. We employ Bayesian model averaging methods to provide estimates for an index of inequality of opportunity (IOp) that are robust to model uncertainty.

By employing data for 29 countries from the 2011 wave of the European Union Statistics on Income and Living Conditions (EU-SILC), we find, as predicted by our model, that higher levels of inequality of opportunity are associated with higher inequality, lower economic mobility, and lower public expenditures on education (as a percentage of GDP). These findings are exacerbated in societies with higher degrees of assortative mating. Higher levels of inequality of opportunity are also associated with lower levels of social trust and reduced confidence in public institutions.

Our overall findings suggest a nexus between inequality of opportunity and societal polarization with one feeding into the other. Rebuilding stability within polarized democracies may require exogenous interventions to rebalance access to economic opportunities. However, achieving this goal may be challenging in a society where the majority heavily influences public policy decisions.

Table 1 – Set of Circumstances

(i) Father's characteristics	(ii) Mother's characteristics
Country of birth and citizenship • Born/citizen present country of residence • Born/citizen of another EU-27 country • Born/citizen of another European country • Born/citizen of a country outside Europe	Country of birth and citizenship • Born/citizen present country of residence • Born/citizen of another EU-27 country • Born/citizen of another European country Born/citizen of a country outside Europe
Highest level of education attained • Don't know/Dead • Neither read nor write in any language • Low level (primary education or lower secondary education) • Medium level (upper secondary education and post-secondary non-tertiary education) • High level (first stage of tertiary education and second stage of tertiary • education)	Highest level of education attained • Don't know/Dead • Neither read nor write in any language • Low level (primary education or lower secondary education) • Medium level (upper secondary education and post-secondary non-tertiary education) • High level (first stage of tertiary education and second stage of tertiary • education)
Occupational activity status • Don't know/Dead • Employed • Self-employed (including family worker) • Unemployed • In retirement or early retirement • Fulfilling domestic tasks and care responsibilities • Other inactive person	Occupational activity status • Don't know/Dead • Employed • Self-employed (including family worker) • Unemployed • In retirement or early retirement • Fulfilling domestic tasks and care responsibilities • Other inactive person
Type of occupation • Don't know/Dead/Unemployed • Managers • Professional • Technicians and associate professionals • Clerical support workers • Service and sales workers • Skilled agricultural, forestry and fishery workers • Craft-related trades workers • Plant and machine operators, and assemblers • Elementary occupations • Armed forces occupations	
iii) Household characteristics	
The financial status of household • Don't know • Very bad • Bad • Moderately bad • Moderately good • Good • Very good	Ability to meet the end of the month • Don't know • With great difficulty • With difficulty • With some difficulty • Fairly easily • Easily • Very easily
Tenancy status of the house • Don't know • Owner • Tenant • Public housing	Family type • Lived with both parents (or persons considered as parents) • Lived with father only (or person considered as a father) • Lived with mother only (or person considered as a mother) • Lived in a private household without any parent • Lived in a collective household or institution
Number of adults	Number of working adults
Number of children	

Table 2 – This table reports the correlation matrix between the Inequality of Opportunity based on different specification of priors. The first 6 specifications correspond to fixed priors structures while the remaining four are based on hyper-g priors.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Uniform g-BRIC	1.000									
(2) Uniform g-UIP	1.000	1.000								
(3) Hierarchical g-BRIC	1.000	1.000	1.000							
(4) Hierarchical g-UIL	1.000	1.000	1.000	1.000						
(5) Tesselation g-BRIC	1.000	1.000	1.000	1.000	1.000					
(6) Tesselation_g-UIP	1.000	1.000	1.000	1.000	1.000	1.000				
(7) Uniform_hg-BRIC	0.993	0.992	0.993	0.992	0.993	0.992	1.000			
(8) Uniform_hg-UIP	0.994	0.993	0.994	0.993	0.994	0.993	1.000	1.000		
(9) Hierarchical_hg-BRIC	0.992	0.991	0.991	0.990	0.992	0.991	0.998	0.998	1.000	
(10) Hierarchical hg-UIP	0.992	0.991	0.992	0.991	0.992	0.991	0.998	0.998	1.000	1.000

Figure 1 – Model selection. The panels show a comparison of different models in the model space.

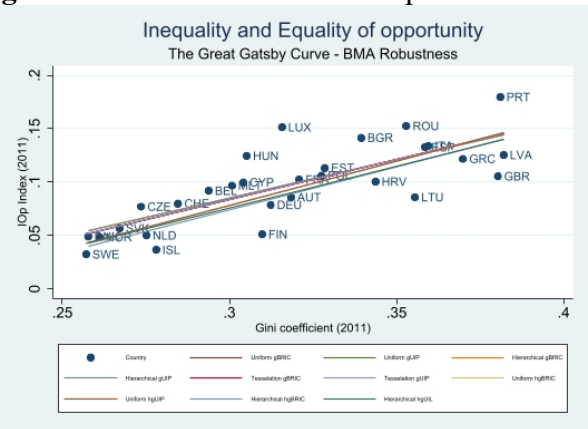

Figure 1.a

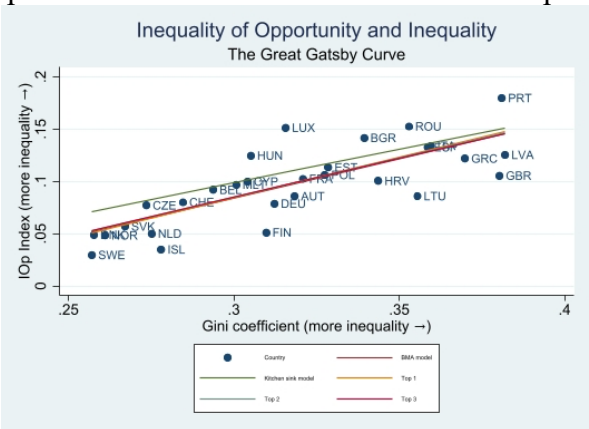

Figure 1.b

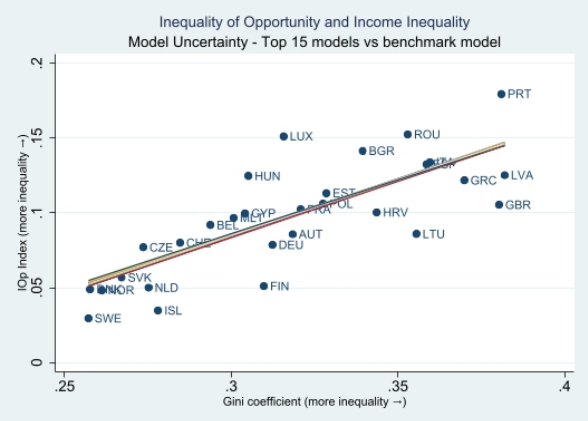

Figure 1.c

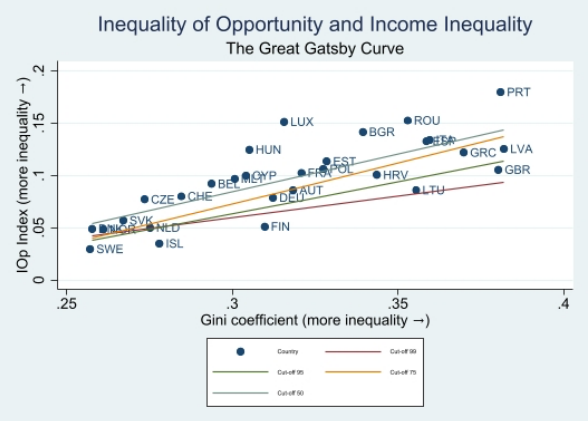

Figure 1.b

Figure 2 – Inequality of Opportunity across European Countries

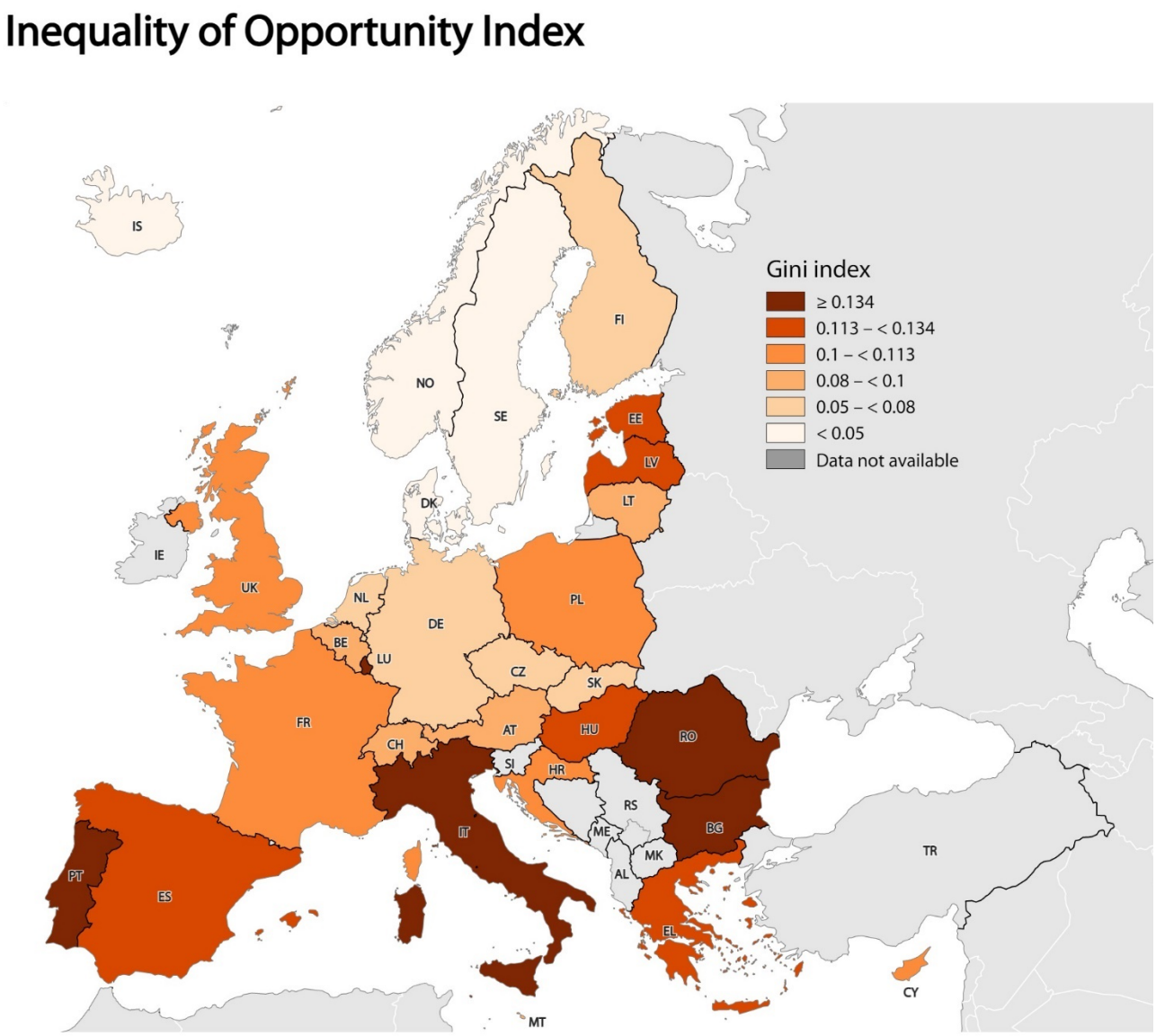

Figure 3 – The Great Gatsby Curve. The panels displays the relationship between Inequality of Opportunity (y-axis) and Income Inequality (x-axis).

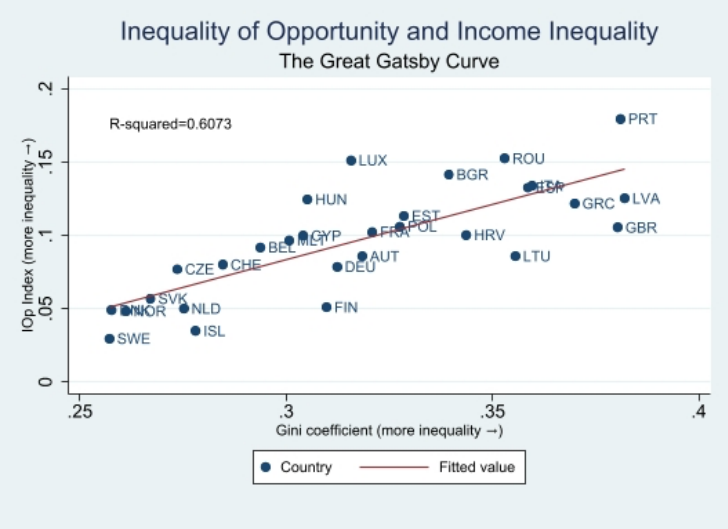

Figure 4 – Mobility. The panels displays the relationship between Inequality of Opportunity (y-axis) and different measure of mobility (x-axis). The measures of mobility are respectively the intergenerational earnings elasticity β (fig. 4.a) and the social mobility index (fig. 4.b). See the Appendix for the data sources.

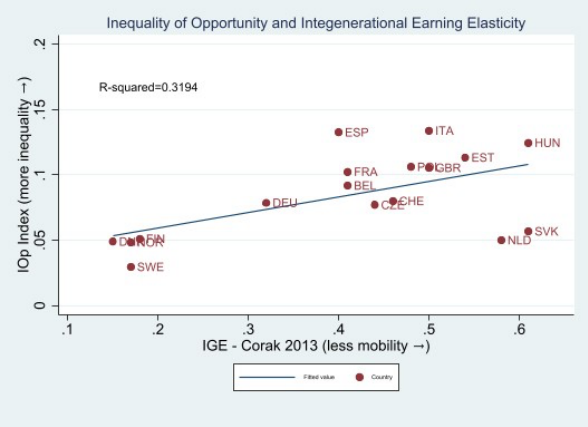

Figure 4.a

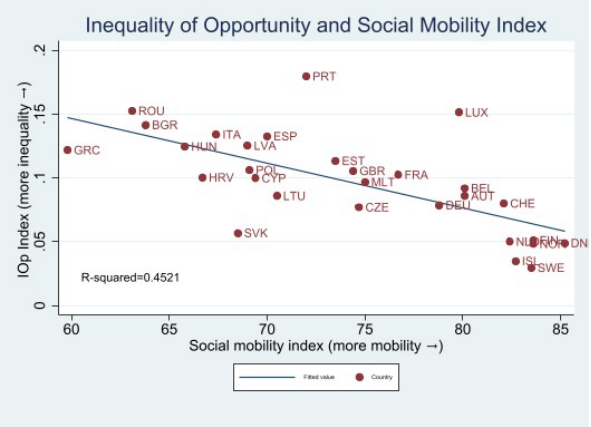

Figure 4.b

Figure 5 – Regional Disparities. The panel displays the relationship between Inequality of Opportunity (y-axis) and a measure of Regional Household Income Disparities (x-axis) based on the Gini coefficient of households' disposable income within countries.

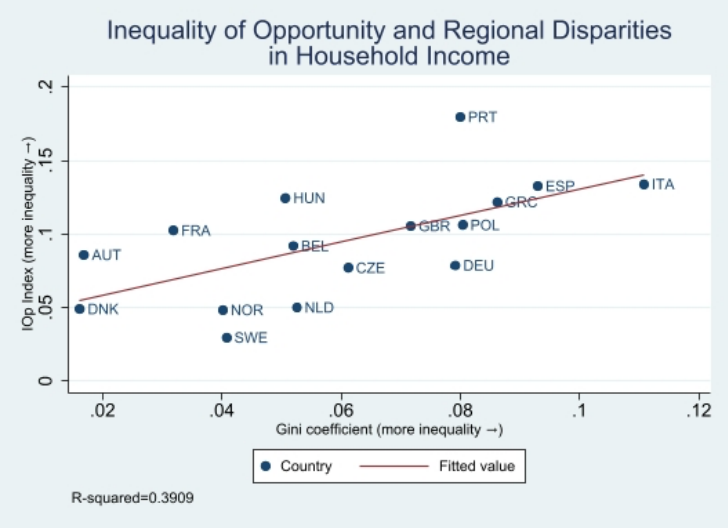

Figure 6 – Assortative Mating. The panel shows the relationship between Inequality of Opportunity and assortative mating measured as the Pearson correlation between the educational level of household partners at the country level.

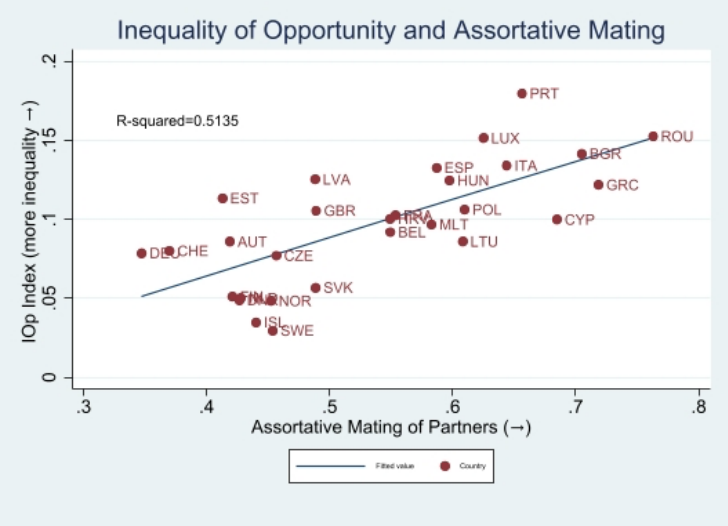

Figure 7 – These panels show the relation between Inequality of Opportunity and public expenditures in education. The left panel shows expenditures on non-tertiary education while the right refers to expenditures on tertiary education (in terms of % of GDP).

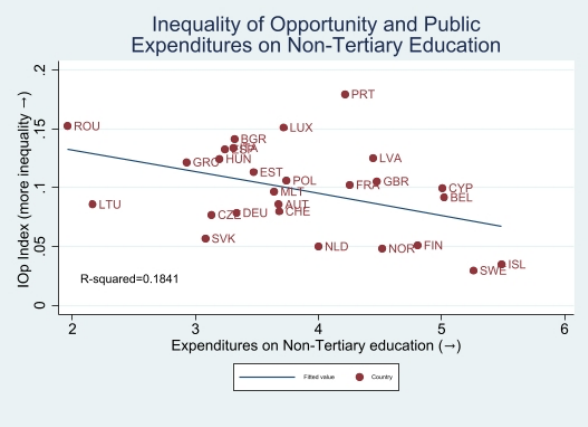

Figure 7.a

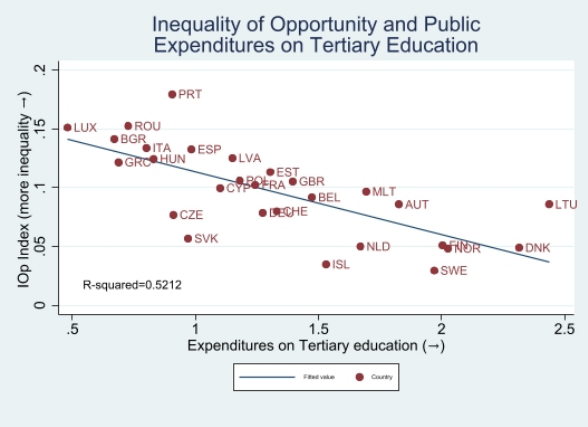

Figure 7.b

Figure 8 – Social Trust. The panels display the relationship between Inequality of Opportunity and trust which is measured as trust in the national Government on the left and interpersonal trust on the right.

Figure 8.a

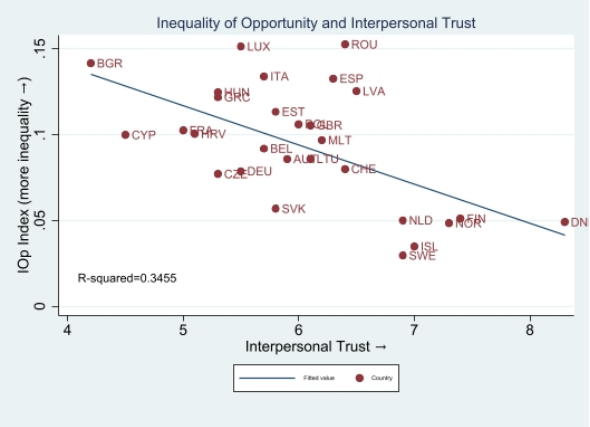

Figure 8.b

Figure 9 – Institutional Quality and Stability. The upper panels show the relationship between Inequality of Opportunity with the rule of law (left) and the corruption perception index (right). The bottom panels show the relationship between Inequality of Opportuniyt with the regulatory quality index and the political stability index. See Appendix B for the data sources.

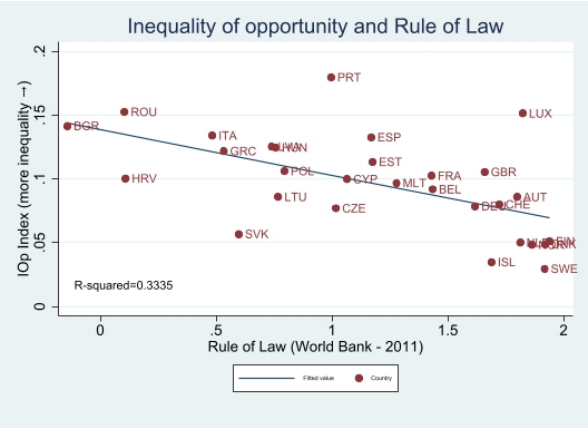

Figure 9.a

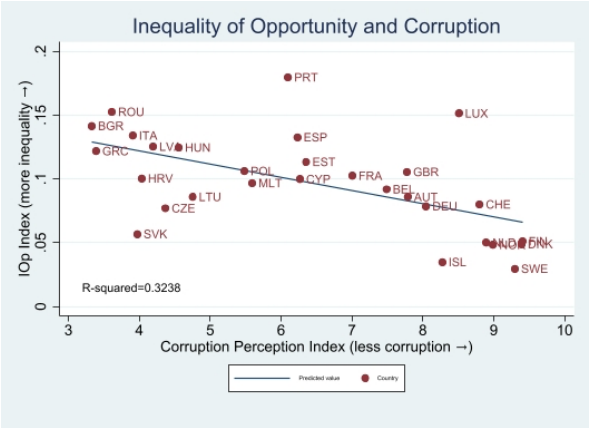

Figure 9.b

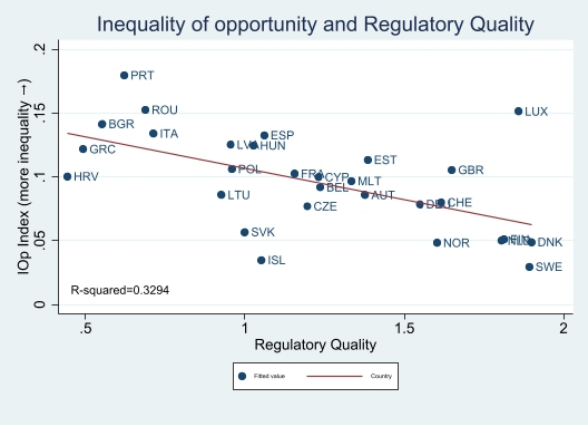

Figure 9.c

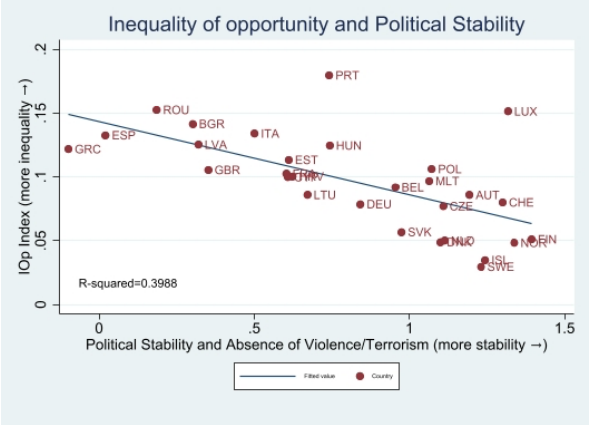

Figure 9.d

Bibliography

Aaberge, R., Mogstad, M. and Peragine, V. (2011) "Measuring Long-term Inequality of Opportunity", *Journal of Public Economics*, 95(3–4), pp. 193-204.

Barry, B. (1991) "Theories of Justice." Berkeley: University of California Press.

Becker G., Tomes N. (1986) "Human Capital and the Rise and Fall of Families. *J. Labor Econ.* 4:S1–39.

Bourguignon, F., Ferreira, F.H.G. and Menéndez, M. (2007) "Inequality of Opportunity in Brazil," *Review of Income and Wealth*, 53: 585-618.

Brock, W., and Durlauf, S. N. (2001) "Growth Empirics and Reality," *World Bank Economic Review*, 15, 229–272.

Brunori, P. and Neidhöfer, G. (2021) "The Evolution of Inequality of Opportunity in Germany: A Machine Learning Approach," *Review of Income and Wealth*, 67: 900-927.

Checchi, D., Peragine, V. (2010) "Inequality of opportunity in Italy," *J Econ Inequal* **8,** 429–450.

Corak, M. (2013). "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives*, 27 (3): 79-102.

Domingue, B., Fletcher, J., Conley, D., Boardman J. (2014) "Genetic and Educational Assortative Mating among US Adults," *Proceedings of the National Academy of Sciences*, 111, 7996–8000.

Durlauf, S. N., Kourtellos, A., Tan, C. M. (2008) "Are Any Growth Theories Robust?" *Economic Journal*, 118(527): 329–46.

Durlauf, S. N., Kourtellos, A., Tan, C. M. (2022) "The Great Gatsby Curve," *Annual Review of Economics*, 14, p. 571-605.

Durlauf, S. N., Ioannides, Y. M. (2010) "Social Interactions," *Annual Review of Economics*, 2(1), 451–478.

Eicher, T., Papageorgiou, C., Raftery, A. E. (2011) "Default Priors And Predictive Performance In Bayesian Model Averaging, With Application To Growth Determinants," *Journal of Applied Econometrics*, 26, 30–55.

Eika, L., Mogstad, M., Zafar, B. (2019) "Educational Assortative Mating and Household Income Inequality," Journal of Political Economy, 127(6), 2795-2835.

Fernandez, C., E. Ley, Steel, M. (2001) "Benchmark Priors for Bayesian Model Averaging," *Journal of Econometrics*, 100, 381–427.

Ferreira, F., Gignoux, J. (2011) "The Measurement of Inequality of Opportunity: Theory and an application to Latin America," *Rev. Income Wealth*, 57, 622–657.

Ferreira, F. H., & Peragine, V. (2016) "Individual Responsibility and Equality of Opportunity." In The Oxford Handbook of Well-Being and Public Policy.

Fleurbaey, M., Schokkaert, E. (2009) “Unfair Inequalities in Health and Healthcare,” *J. Health Econ.* 28, 73–90.

George, E. (1999) “Discussion of Bayesian Model Averaging and Model Search Strategies,” M. A. Clyde., in J. Bernardo, J. Berger, A. David, and A. Smith, ed.: Bayesian Statistics 6. pp. 157–177 (Clarendon Press: Oxford).

George, E. (2010) “Dilution Priors: Compensating for Model Space Redundancy,” in J. Berger, T. Cai, and I. Johnstone, ed.: Borrowing Strength: Theory Powering Applications - A Festschrift for Lawrence D. Brown . pp. 165–185 (Institute of Mathematical Statistics: Beachwood, OH).

Greenwood, J., Guner, N., Kocharkov, G., Santos, C. (2016) “Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment, and Married Female Labor-Force Participation,” *American Economic Journal: Macroeconomics*, American Economic Association, 8(1), 1-41.

Holmlund, H. (2008) “Intergenerational Mobility and Assortative Mating: Effects of an Educational Reform,” CEE DP 91, London School of Economics.

Hufe, P., Peichl, A., Roemer, J. *et al.* (2017) “Inequality of Income Acquisition: the Role of Childhood Circumstances.” *Soc Choice Welf,* 49**,** 499–544.

Hufe, P., Kanbur, P., Peichl, A. (2018), “Measuring Unfair Inequality: Reconciling Equality of Opportunity and Freedom from Poverty” IZA DP No. 11601.

Krueger A. (2012) “The Rise and Consequences of Inequality in the United States.” Speech, Cent. Am. Prog., Washington, DC.

Leamer, E. (1983) “Let’s Take the Con Out of Econometrics,” *American Economic Review*, 73, 31-43.

Lefranc, A., Pistolesi, N., and Trannoy. A. (2009) “Equality of Opportunity and Luck: Definitions and Testable Conditions, with an Application to Income in France,” *Journal of Public Economics*, 93 (11-12): 1189-1207.

Liang, F., Paulo, R., Molina, G., Clyde, M. A., Berger, J. O. (2008) “Mixtures of g Priors for Bayesian Variable Selection,” *Journal of the American Statistical Association*, 103, 410–423.

Moser, M., Hofmarcher, P. (2014) “Model Priors Revisited: Interaction Terms in BMA Growth Applications,” *Journal of Applied Econometrics*, 29, 344–347.

Pignataro, G. (2012) “Equality of Opportunity: Policy and Measurement Paradigms,” *Journal of Economic Surveys*, 26: 800-834.

Raftery, A. E., Madigan, D., Hoeting, J. A. (1997) “Model Selection and Accounting for Model Uncertainty in Linear Regression Models,” *Journal of the American Statistical Association,* 92:179–191.

Roemer, J. (1993) “A Pragmatic Theory of Responsibility for the Egalitarian Planner.” *Philosophy & Public Affairs* 10: 146–66.

Roemer, J. (1996) “Theories of Distributive Justice.” Cambridge MA: Harvard University Press.

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

Roemer, J., Trannoy, A. (2016) “Equality of Opportunity: Theory and Measurement,” *Journal of Economic Literature*, 54 (4): 1288-1332.

Sala-i-Martin, X., Doppelhofer, G., Miller, R. (2004) “Determinants of Long-term Growth: a Bayesian Averaging of Classical Estimates (BACE) Approach,” *Am. Econ. Rev.*, 94, 813–835.

Solon, G. (2004) “A Model of Intergenerational Mobility Variation over Time and Place.” Chapter 2 in Miles Corak (ed.), *Generational Income Mobility in North America and Europe*. Cambridge, UK: Cambridge University Press.

Wagstaff, A., and Kanbur, R. (2015) “Inequality of Opportunity: The New Motherhood and Apple Pie?” *Health Econ.*, 24, 1243– 1247.

World Economic Forum. "The global social mobility report 2020: Equality, opportunity and a new economic imperative." Cologny: World Economic Forum, 2020.

A. Appendix

Table A.1 – Aggregate variables

Variables	Source	Year	Description
Corruption Index	World Bank	2011	Control of Corruption captures perceptions of the extent to which public power is exercised for private gain. Estimate gives the country's score on the aggregate indicator i.e. rating in units of a standard normal distribution from -2.5 to 2.5.
Public Expenditure on Non-Tertiary Education (% of GDP)	Eurostat	2011	Government expenditure on education as a percentage of GDP
Public Expenditure on Tertiary Education (% of GDP)	Eurostat	2011	Government expenditure on education as a percentage of GDP
Assortative mating	EU-SILC	2011	The Pearson correlation coefficient was calculated to assess the relationship between the educational levels of partners within the same household, at the country level.
Inequality Index (Equalized disposable Income - Gini)	EU-SILC	2011	$EDI = \frac{Diposableincome}{\sqrt{Family\ Size}}$ Index rating from 0 to 1
Intergenerational Earnings Elasticity	Corak (2013)		Index rating from 0 to 1
Interpersonal Trust	Eurostat	2013	Average rating of trust in others (rating from 0 to 10 in EU-SILC survey)
Political Instability Index	World Bank	2011	Political Stability and Absence of Violence/Terrorism measures perceptions of the likelihood of political instability and/or politically-motivated violence, including terrorism. Country's score on the aggregate indicator, i.e. rating from -2.5 to 2.5.
Regional Income Disparity	OECD	2009	"Gini index of inequality of GDP per capita across TL3 regions, 1995 and 2010", in OECD (2013)
Regulatory Quality Index	World Bank	2011	Regulatory Quality captures perceptions of the ability of the government to formulate and implement sound policies and regulations that permit and promote private sector development. Estimate gives the country's score on the aggregate indicator i.e. rating from -2.5 to 2.5
Role of Law	World Bank	2011	Rule of Law captures perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence. Estimate gives the country's score on the aggregate indicator, in units of a standard normal distribution, i.e. rating from -2.5 to 2.5.
Social Mobility Index	World Economic Forum	2011	Rating from 0 to 100
Trust in Government	Eurostat	2013	Average rating of trust in the National Goverment (rating from 0 to 10 in EU-SILC survey).

Table A.2 – Descriptive statistics of variables

Table A.2.1 – Descriptive statistics of Father's variables

	Birth area		Citizenship		Education				Activity		Occupation		
	Native	Other Country	Resid.	No Resid.	Illiterate	low	Medium	High	Empl.	Self-Employed	Manual workers	Low-skills non manual	High-skills non manual
Austria	0.80	0.20	0.84	0.16	0.01	0.01	0.56	0.36	0.77	0.23	0.59	0.17	0.22
Belgium	0.81	0.81	0.82	0.18	0.02	0.50	0.23	0.21	0.79	0.21	0.49	0.16	0.35
Bulgaria	0.99	0.01	1.00	0.00	0.02	0.51	0.36	0.10	0.97	0.03	0.71	0.08	0.14
Croatia	0.83	0.17	0.84	0.16	0.01	0.50	0.36	0.07	0.85	0.13	0.69	0.17	0.12
Cyprus	0.82	0.18	0.83	0.17	0.05	0.66	0.19	0.09	0.59	0.40	0.68	0.16	0.14
Czech Republic	0.92	0.08	0.96	0.04	0.00	0.66	0.23	0.09	0.98	0.02	0.66	0.24	0.08
Denmark	0.95	0.05	0.96	0.04	0.00	0.32	0.46	0.22	0.71	0.29	0.52	0.16	0.32
Estonia	0.79	0.21	0.82	0.18	0.00	0.37	0.40	0.16	0.97	0.01	0.71	0.03	0.21
Finland	0.96	0.04	0.96	0.04	0.03	0.49	0.26	0.22	0.74	0.26	0.45	0.22	0.27
France	0.82	0.18	0.88	0.12	0.03	0.71	0.08	0.11	0.80	0.18	0.56	0.12	0.29
Germany	0.88	0.12	0.94	0.06	0.00	0.09	0.52	0.26	0.84	0.13	0.47	0.12	0.33
Greece	0.90	0.10	0.93	0.07	0.05	0.61	0.12	0.07	0.42	0.56	0.70	0.13	0.12
Hungary	0.98	0.02	0.99	0.01	0.02	0.62	0.25	0.09	0.93	0.05	0.72	0.15	0.07
Iceland	0.93	0.07	0.93	0.07	0.00	0.32	0.50	0.16	0.66	0.34	0.54	0.34	0.12
Italy	0.93	0.07	0.94	0.06	0.03	0.75	0.17	0.05	0.71	0.29	0.62	0.21	0.15
Latvia	0.77	0.33	0.80	0.20	0.00	0.46	0.37	0.12	0.97	0.01	0.73	0.04	0.19
Lithuania	0.93	0.07	0.96	0.04	0.01	0.61	0.24	0.09	0.94	0.01	0.70	0.09	0.16
Luxembourg	0.41	0.59	0.43	0.57	0.04	0.49	0.33	0.13	0.82	0.18	0.59	0.09	0.30
Malta	0.96	0.04	0.97	0.03	0.16	0.57	0.19	0.06	0.77	0.22	0.50	0.22	0.19
Netherlands	0.93	0.07	0.96	0.04	0.00	0.37	0.32	0.22	0.80	0.20	0.43	0.13	0.41
Norway	0.90	0.10	0.92	0.08	0.00	0.30	0.39	0.31	0.74	0.26	0.46	0.08	0.44
Poland	0.97	0.03	1.00	0.07	0.00	0.48	0.45	0.06	0.71	0.29	0.79	0.07	0.13
Portugal	0.96	0.04	0.97	0.03	0.21	0.70	0.03	0.03	0.71	0.28	0.71	0.14	0.14
Romania	1.00	0.00	1.00	0.00	0.02	0.77	0.10	0.03	0.65	0.25	0.67	0.03	0.07
Slovak Republic	0.98	0.02	0.99	0.01	0.00	0.34	0.55	0.09	0.97	0.02	0.66	0.07	0.21
Spain	0.90	0.10	0.91	0.09	0.04	0.78	0.07	0.09	0.62	0.19	0.62	0.15	0.19
Sweden	0.84	0.16	0.88	0.12	0.00	0.51	0.16	0.22	0.81	0.19	0.53	0.15	0.30
Switzerland	0.69	0.31	0.71	0.29	0.03	0.19	0.54	0.17	0.69	0.30	0.45	0.12	0.40
United Kingdom	0.84	0.16	0.90	0.10	0.02	0.51	0.25	0.17	0.85	0.15	0.51	0.12	0.35

Table A.2.2 – Descriptive statistics of Mother's variables

	Birth area		Citizenship		Education				Activity		
	Native	Other Country	Resid.	No Resid.	Illiterate	low	Medium	High	Empl.	Self-Employed	Housework
Austria	0.80	0.2	0.84	0.16	0.01	0.56	0.36	0.05	0.39	0.17	0.43
Belgium	0.81	0.19	0.83	0.17	0.03	0.55	0.23	0.15	0.35	0.13	0.51
Bulgaria	0.99	0.01	0.99	0.01	0.02	0.49	0.37	0.11	0.89	0.03	0.05
Croatia	0.83	0.17	0.85	0.15	0.01	0.68	0.21	0.04	0.35	0.07	0.54
Cyprus	0.82	0.18	0.83	0.17	0.09	0.67	0.17	0.06	0.32	0.17	0.50
Czech Republic	0.93	0.07	0.96	0.04	0	0.68	0.27	0.04	0.91	0.01	0.07
Denmark	0.95	0.05	0.96	0.04	0	0.51	0.31	0.18	0.62	0.07	0.29
Estonia	0.79	0.21	0.83	0.17	0	0.36	0.41	0.19	0.92	0.01	0.07
Finland	0.96	0.04	0.97	0.03	0.03	0.49	0.28	0.20	0.66	0.25	0.06
France	0.83	0.17	0.90	0.10	0.05	0.73	0.09	0.08	0.45	0.09	0.42
Germany	0.88	0.12	0.94	0.06	0.01	0.26	0.53	0.10	0.50	0.05	0.42
Greece	0.90	0.10	0.93	0.07	0.08	0.59	0.12	0.04	0.17	0.32	0.49
Hungary	0.98	0.02	0.99	0.01	0.03	0.65	0.25	0.06	0.73	0.02	0.22
Iceland	0.92	0.08	0.94	0.06	0.00	0.61	0.28	0.09	0.61	0.10	0.27
Italy	0.93	0.07	0.94	0.06	0.00	0.80	0.14	0.03	0.25	0.10	0.64
Latvia	0.78	0.22	0.81	0.19	0.01	0.41	0.42	0.14	0.89	0.01	0.07
Lithuania	0.94	0.06	0.97	0.03	0.01	0.56	0.30	0.10	0.84	0.02	0.11
Luxembourg	0.40	0.60	0.42	0.58	0.06	0.61	0.24	0.08	0.32	0.12	0.54
Malta	0.96	0.04	0.97	0.03	0.15	0.66	0.15	0.03	0.07	0.01	0.90
Netherlands	0.93	0.07	0.97	0.03	0	0.52	0.32	0.10	0.62	0.06	0.30
Norway	0.89	0.11	0.91	0.09	0	0.34	0.44	0.21	0.63	0.10	0.24
Poland	0.97	0.03	1.00	0	0	0.54	0.41	0.05	0.48	0.31	0.18
Portugal	0.95	0.05	0.97	0.03	0.28	0.64	0.03	0.03	0.35	0.20	0.41
Romania	1.00	0	1.00	0	0.02	0.76	0.12	0.02	0.35	0.23	0.30
Slovak Republic	0.98	0.02	0.99	0.01	0.00	0.42	0.52	0.05	0.86	0.01	0.10
Spain	0.90	0.10	0.91	0.09	0.07	0.82	0.05	0.04	0.17	0.07	0.75
Sweden	0.83	0.17	0.88	0.12		0.50	0.21	0.24	0.72	0.06	0.20
Switzerland	0.67	0.33	0.70	0.30	0.05	0.38	0.46	0.06	0.38	0.15	0.44
United Kingdom	0.84	0.16	0.90	0.10	0.03	0.67	0.11	0.14	0.60	0.05	0.26

Table A.2.3 – Descriptive statistics of Household variables

	Parents in HH		HH Composition			Financial situation		Tenancy status	
	Both	Lone	Adults	Workers	Child	bad	good	Owner	Rentier
Austria	0.92	0.06	2.71	1.80	2.54	0.52	0.48	0.60	0.27
Belgium	0.92	0.07	2.38	1.66	2.72	0.24	0.56	0.79	0.19
Bulgaria	1	0	2.50	2.09	2.05	0.24	0.56	0.91	0.05
Croatia	0.95	0.3	2.62	1.48	2.33	0.63	0.37	0.89	0.05
Cyprus	0.95	0.3	2.66	1.70	2.75	0.5	0.5	0.80	0.11
Czech Republic	0.94	0.5	2.17	2.01	2.24	0.41	0.69	0.64	0.03
Denmark	1	0	2.35	2.48	2.26	0.30	0.7	0.81	0.19
Estonia	0.91	0.08	2.25	1.93	2.26	0.37	0.63	0.85	0.09
Finland	0.98	0..01	2.47	2.30	2.29	0.25	0.75	0.84	0.15
France	0.90	0.08	2.47	1.70	1.74	0.36	0.64	0.65	0.30
Germany	0.88	0.10	2.25	1.72	2.23	0.20	0.80	0.52	0.45
Greece	0.95	0.02	2.35	1.63	2.37	0.59	0.41	0.84	0.11
Hungary	0.90	0.06	2.18	1.79	2.29	0.63	0.37	0.84	0.12
Iceland	0.94	0.05	2.42	1.94	2.60	0.74	0.36	0.90	0.08
Italy	1	0	2.62	1.66	2.35	0.48	0.52	0.71	0.04
Latvia	0.97	0.02	2.17	1.96	2.33	0.34	0.66	0.50	0.28
Lithuania	0.93	0.05	2.37	2.08	2.57	0.52	0.48	.70	0.22
Luxembourg	0.92	0.05	2.55	1.67	2.70	0.39	0.61	0.75	0.22
Malta	0.97	0.01	3.02	1.86	2.68	0.42	0.58	0.59	0.37
Netherlands	0.96	0.03	2.11	1.58	3.06	0.20	0.8	0.63	0.36
Norway	0.99	0.01	2.08	1.82	1.87	0.24	0.76	0.93	0.06
Poland	0.97	0.02	2.79	2.05	2.48	0.54	0.46	0.68	0.07
Portugal	0.95	0.03	2.75	2.25	2.70	0.67	0.33	0.57	0.08
Romania	0.99	0.01	2.79	1.92	2.22	0.68	0.32	0.84	0.03
Slovak Republic	0.98	0.02	2.53	2.13	2.37	0.51	0.49	0.67	0.01
Spain	0.96	0.01	2.91	2.13	2.42	0.40	0.60	0.82	0.03
Sweden	0.86	0.11	2.10	1.84	2.38	0.32	0.68	0.76	0.23
Switzerland	0.90	0.07	2.51	1.89	2.49	0.36	0.64	0.56	0.44
United Kingdom	0.89	0.08	2.35	2.32	2.39	0.41	0.59	0.68	0.29

Table A.2.4 – Descriptive statistics of Country-level Covariates

VARIABLES	N	mean	sd	min	max
Inequality of Opportunity	29	0.097	0.0380	0.0295	0.180
Gini index	29	0.318	0.0392	0.257	0.382
IGE Index	17	0.408	0.157	0.150	0.610
Social Mobility Index	29	74.18	7.252	59.80	85.20
HDI	29	0.879	0.0418	0.794	0.944
Regional Income Disparity (Gini)	16	0.0602	0.0273	0.0162	0.111
Assortative Mating	29	0.535	0.112	0.347	0.763
Public Expenditure on Non-Tertiary Education	27	3.817	0.879	1.960	5.490
Public Expenditure on Tertiary Education	28	1.317	0.520	0.480	2.437
Trust in National Government	27	57.72	18.63	24.30	94.30
Trust in Others	29	5.990	0.893	4.200	8.300
Rule of Law	29	1.175	0.612	-0.142	1.940
Corruption Index	29	1.090	0.859	-0.333	2.393
Regulatory Quality	29	1.210	0.444	0.443	1.899
Political Stability	29	0.808	0.419	-0.0987	1.393

Tab A.3: List of dummies variables

Variable	Definition	Explanation
DCitiz_F1	Citizenship Father	Dummy =1 if Born in the respondent's present country of residence (i.e. country of the survey)
DCitiz_F2	Citizenship Father	Dummy =1 if Born in another EU-27 country
DCitiz_F3	Citizenship Father	Dummy =1 if Born in another European country
DCitiz_F4	Citizenship Father	Dummy =1 if Born outside Europe
DCitiz_M1	Citizenship Mother	Dummy =1 if Born in the respondent's present country of residence (i.e. country of the survey)
DCitiz_M2	Citizenship Mother	Dummy =1 if Born in another EU-27 country
DCitiz_M3	Citizenship Mother	Dummy =1 if Born in another European country
DCitiz_M4	Citizenship Mother	Dummy =1 if Born outside Europe
Dcountry_F1	Country of Birth - Father	Dummy =1 if Born in the respondent's present country of residence (i.e. country of the survey)
Dcountry_F2	Country of Birth - Father	Dummy =1 if Born in another EU-27 country
Dcountry_F3	Country of Birth - Father	Dummy =1 if Born in another European country
Dcountry_F4	Country of Birth - Father	Dummy =1 if Born outside Europe
Dcountry_M1	Country of Birth - Mother	Dummy =1 if Born in the respondent's present country of residence (i.e. country of the survey)
Dcountry_M2	Country of Birth - Mother	Dummy =1 if Born in another EU-27 country
Dcountry_M3	Country of Birth - Mother	Dummy =1 if Born in another European country
Dcountry_M4	Country of Birth - Mother	Dummy =1 if Born outside Europe
DEdu_F1	Highest level of education attained by the father	Dummy =1 if don't know
DEdu_F2	Highest level of education attained by the father	Dummy =1 if could neither read nor write in any language
DEdu_F3	Highest level of education attained by the father	Dummy =1 if Low level (pre-primary, primary education or lower secondary education)
DEdu_F4	Highest level of education attained by the father	Dummy =1 if Medium level (upper secondary education and post-secondary non-tertiary education)
DEdu_F5	Highest level of education attained by the father	Dummy =1 if High level (first stage of tertiary education and second stage of tertiary education)
Dedu_M1	Highest level of education attained by the mother	Dummy =1 if don't know
Dedu_M2	Highest level of education attained by the mother	Dummy =1 if could neither read nor write in any language
Dedu_M3	Highest level of education attained by the mother	Dummy =1 if Low level (pre-primary, primary education or lower secondary education)
Dedu_M4	Highest level of education attained by the mother	Dummy =1 if Medium level (upper secondary education and post-secondary non-tertiary education)
Dedu_M5	Highest level of education attained by the mother	Dummy =1 if High level (first stage of tertiary education and second stage of tertiary education)
DOcc_M1	ACTIVITY STATUS OF THE MOTHER	Dummy =1 if don't know
DOcc_M2	ACTIVITY STATUS OF THE MOTHER	Dummy =1 if Other inactive person
DOcc_M3	ACTIVITY STATUS OF THE MOTHER	Dummy =1 if Fulfilling domestic tasks and care responsibilities
DOcc_M4	ACTIVITY STATUS OF THE MOTHER	Dummy =1 if Unemployed
DOcc_M5	ACTIVITY STATUS OF THE MOTHER	Dummy =1 if In retirement or in early retirement or had given up business
DOcc_M6	ACTIVITY STATUS OF THE MOTHER	Dummy =1 if Employed
DOcc_M7	ACTIVITY STATUS OF THE MOTHER	Dummy =1 if Self-employed (including family worker)
DOcc_F1	ACTIVITY STATUS OF THE Father	Dummy =1 if don't know
DOcc_F2	ACTIVITY STATUS OF THE MOTHER	Dummy =1 if Other inactive person
DOcc_F3	ACTIVITY STATUS OF THE Father	Dummy =1 if Fulfilling domestic tasks and care responsibilities
DOcc_F4	ACTIVITY STATUS OF THE Father	Dummy =1 if Unemployed
DOcc_F5	ACTIVITY STATUS OF THE Father	Dummy =1 if In retirement or in early retirement or had given up business
DOcc_F6	ACTIVITY STATUS OF THE Father	Dummy =1 if Employed
DOcc_F7	ACTIVITY STATUS OF THE Father	Dummy =1 if Self-employed (including family worker)
DMOcc_F1	Main Occupation - Father	Dummy =1 if unemployed/unknown/dead
DMOcc_F2	Main Occupation - Father	Dummy =1 if low-skilled non-manual
DMOcc_F3	Main Occupation - Father	Dummy =1 if skilled manual and elementary occupation
DMOcc_F4	Main Occupation - Father	Dummy =1 if high-skilled non-manual
Dfin_sit_a_141	Financial situation of the household	Dummy =1 if don't know
Dfin_sit_a_142	Financial situation of the household	Dummy =1 if Very bad
Dfin_sit_a_143	Financial situation of the household	Dummy =1 if bad
Dfin_sit_a_144	Financial situation of the household	Dummy =1 if Moderately bad
Dfin_sit_a_145	Financial situation of the household	Dummy =1 if Moderately good
Dfin_sit_a_146	Financial situation of the household	Dummy =1 if Good
Dfin_sit_a_147	Financial situation of the household	Dummy =1 if Very good
Dfin_sit141	Ability to make ends meet	Dummy =1 if don't know
Dfin_sit142	Ability to make ends meet	Dummy =1 if With great difficulty
Dfin_sit143	Ability to make ends meet	Dummy =1 if With difficulty
Dfin_sit144	Ability to make ends meet	Dummy =1 if With some difficulty
Dfin_sit145	Ability to make ends meet	Dummy =1 if Fairly easily
Dfin_sit146	Ability to make ends meet	Dummy =1 if Easily
Dfin_sit147	Ability to make ends meet	Dummy =1 if Very easily
DprecenceP1	Presence of parents	Dummy =1 if Lived with both parents (or persons considered as parents)
DprecenceP2	Presence of parents	Dummy =1 if Lived with father only (or person considered as a father)
DprecenceP3	Presence of parents	Dummy =1 if Lived with mother only (or person considered as a mother)
DprecenceP4	Presence of parents	Dummy =1 if Lived in a private household without any parent
Dten_status1	Tenancy status of the house in which the respondent was living	Dummy =1 if don't know
Dten_status2	Tenancy status of the house in which the respondent was living	Dummy =1 if Accommodation was provided free
Dten_status3	Tenancy status of the house in which the respondent was living	Dummy =1 if Tenant
Dten_status4	Tenancy status of the house in which the respondent was living	Dummy =1 if Owner

Tab A.4: Father characteristics - Heatmap based on posterior inclusion probability (g-BRIC)

Country	GDP_P	DCitiz_F1	DCitiz_F2	DCitiz_F3	DCitiz_F4	DCountry_F2	DCountry_F3	DCountry_F4	DEdu_F2	DEdu_F3	DEdu_F4	DEdu_F5	DMOcc_FF	DMOcc_FR3	DMOcc_FR4	DOcc_F6	DOcc_F7
Luxembourg	108010,7		0,596	0,001	0,002	0,192	0,001	0,005	0,140	0,257	1,000	1,000	0,993	0,009	1,000	0,973	
Switzerland	66342,01		0,012	0,040	0,167	0,051	0,083	0,429	0,012	0,089	1,000	1,000	0,000	1,000	0,004	0,000	0,000
Austria	53102,95		0,863	0,545		0,009	0,173		0,002	0,007	0,002	0,998	0,610	0,274	0,765	0,020	0,022
Netherlands	52594,23		0,012	0,017	0,186	0,009	0,008	0,236	0,006	0,846	0,918	0,913	0,092	0,474	0,603	0,055	0,047
Denmark	51293,21		0,249	0,060		0,653	0,043	0,019		0,012	0,012			0,682	0,370	0,022	
Germany	49684,61		0,007			0,003			0,000	0,287	0,791	0,998	1,000	0,809	1,000	0,136	0,478
Iceland	48186,17		0,224	0,010	0,331	0,359	0,013	0,266	0,005	0,064	0,007	0,098	0,003	0,081	0,149	0,051	0,053
Belgium	48153,94		0,001	0,138	0,176	0,001	0,397	0,001	0,001	0,032	0,068	0,184	0,856	0,178	0,001		0,000
Italy	42892,31		0,086	0,090	0,371	0,001	0,004	0,018	0,584	1,000	1,000		0,721	0,994	0,220	0,706	
France	42864,02		0,009	0,002	0,982	0,049	0,002	0,030	0,093	0,171	0,176	1,000	0,126	0,771	0,552	0,067	0,945
United Kingdom	42294,12		0,126	0,029	0,026	0,992	0,043	0,581	0,004	0,745	0,961	0,999	0,042	0,675	0,358	0,023	
Cyprus	37562,45		0,321	0,052	0,002	0,679	0,897	0,001	0,142	0,008	0,001	0,125	0,097	0,271	0,996	0,002	0,002
Spain	36884,26		0,347	0,027	0,110	0,593	0,005	0,001	0,719	0,443	0,572	1,000	0,001	1,000	0,015	0,001	
Czech Republic	34002,19		0,007	0,004	0,003	0,009	0,003	0,004	0,026	0,130	0,856	0,982	0,024	0,108	0,894	0,730	
Malta	32881,83		0,171			0,048			0,999	0,993	0,013	0,018	0,020	0,740	0,998	0,007	0,011
Portugal	31304,82		0,782	0,992		0,038	0,006		0,058	0,894	1,000	1,000	0,790	0,216	0,845	0,004	0,006
Greece	30372,76		0,010	0,721	0,023	0,007	0,004	0,022	0,280	0,728	0,986	1,000	0,876	0,128	0,886	0,010	0,008
Estonia	28159,58		0,016			0,970				0,114	0,894	0,975	0,007	0,007	1,000	0,010	0,051
Slovakia	26107,83		0,009	0,022		0,007	0,030	0,002	0,011	0,047	0,142	0,993	0,022	0,500	0,599	0,003	0,008
Lithuania	25968,58		0,010	0,008	0,003	0,005	0,024	0,005	0,004	0,064	0,051	1,000	0,002	0,005	0,007	0,612	0,004
Poland	25124,28		0,012	0,002	0,000	0,422	0,015	0,002		0,133	0,878	1,000	0,215	0,287	0,999	0,000	
Hungary	24971,48		0,141	0,004	0,000	0,004	0,017	0,000	0,000	0,000	1,000	1,000	0,425	0,009	1,000	0,000	0,004
Croatia	24316,25		0,027	0,007	0,007	0,036	0,006	0,007	0,027	0,014	0,994	1,000	0,624	0,632	0,384	0,007	0,006
Latvija	22795,2		0,168				0,284		0,066	0,974	0,032	0,999	0,011	0,254	0,400	0,006	0,007
Romania	20792,17		0,004	0,027	0,004	0,005	0,011	0,005	0,005	0,030	0,994	1,000	0,043	0,012	0,162	0,497	0,512
Bulgaria	17966,16	0,041							0,007	0,198	0,301	1,000	0,027	1,000	0,119	0,008	0,009
Finland	10787,66		0,007	0,023	0,143	0,010	0,014	0,195	0,006	0,009	0,010	0,725	0,029	0,021	0,524		0,009

Table A.5: Mother characteristics - Heatmap based on posterior inclusion probability (g-BRIC)

Country	GDP_P	DCitiz_M1	DCitiz_M2	DCitiz_M3	DCitiz_M4	DCountry_M2	DCountry_M3	DCountry_M4	DEdu_M2	DEdu_M3	DEdu_M4	DEdu_M5	DOcc_M2	DOcc_M3	DOcc_M4	DOcc_M5	DOcc_M6	DOcc_M7
Luxembourg	108010,7		0,217	0,991	0,001	0,059	0,010	1,000	0,950	0,712	0,300	0,357	0,002	0,042	0,001	0,001	0,614	0,056
Switzerland	66342,01		0,419	0,163	0,150	0,526	0,674	0,257	0,207	0,003	0,044	0,758	0,009	0,261	0,000	0,001	0,056	0,001
Austria	53102,95		0,129	0,466		0,001	0,023		0,003	0,647	0,267	0,007	0,004	0,001	0,001	0,001	0,005	0,024
Netherlands	52594,23		0,003	0,016	0,060	0,006	0,006	0,024	0,007	0,044	0,152	0,013	0,004	0,050	0,003	0,006	0,645	0,110
Denmark	51293,21		0,034	0,155	0,014	0,026	0,118	0,040		0,015	0,014		0,019	0,013	0,034	0,015	0,995	
Germany	49684,61		0,690			0,084			0,001	0,269	0,270	0,278	0,000	0,990	0,066	0,000	0,011	0,064
Iceland	48186,17		0,375	0,006	0,373	0,020	0,005	0,025	0,009	0,004	0,004	0,004	0,004	0,855	0,005	0,022	0,023	
Belgium	48153,94		0,001	0,158	0,241	0,002	0,308	0,588	0,013	0,158	0,123	0,086		0,000	0,000	0,000	0,000	0,000
Italy	42892,31		0,819	0,830	0,601	0,097	0,079	0,013	0,614	1,000	0,964		0,005	0,003	0,001	0,000	0,262	
France	42864,02		0,002	0,001	0,020	0,003	0,003	0,008	0,012	0,010	1,000	0,120	0,012	0,005	0,007	0,003	0,007	0,001
United Kingdom	42294,12		0,045	0,046	0,018	0,003	0,078	0,241	0,199	0,007	0,008	0,036		0,002	0,050	0,003	0,002	0,016
Cyprus	37562,45		0,000	0,015	0,057	0,000	0,036	0,940	0,675	0,620	0,536	0,639	0,001	0,000	0,001	0,000	0,996	0,000
Spain	36884,26		0,001	0,319	0,899	0,064	0,656	0,467	0,937	0,329	0,360	1,000	0,035	0,008	0,003	0,001	0,006	0,001
Czech Republic	34002,19		0,004	0,004	0,003	0,039	0,004	0,003	0,185	0,220	0,825	0,845	0,093	0,005	0,003	0,278	0,010	
Malta	32881,83		0,785			0,061			0,047	0,089	0,907	1,000	0,005	0,150	0,016	0,009	0,268	0,014
Portugal	31304,82		0,183	0,007		0,007	0,002		0,363	0,638	1,000	1,000	0,008	0,905	0,002	0,011	0,015	0,076
Greece	30372,76		0,007	0,265	0,038	0,008	0,053	0,995	0,012	0,009	0,039	0,214	0,006	0,008	0,004	0,004	0,003	0,015
Estonia	28159,58		0,009			0,985			0,013	0,856	0,151	0,951	0,010	0,027	0,009	0,009	0,017	0,018
Slovakia	26107,83		0,006	0,163		0,013	0,228	0,004	0,006	0,570	0,423	0,639	0,007	0,078	0,002	0,004	0,854	0,008
Lithuania	25968,58		0,004	0,072	0,003		0,007	0,004	1,000	0,004	0,003	0,003	0,006	0,011	0,003	0,003	0,003	0,003
Poland	25124,28		0,007	0,000	0,000	0,001	0,497	0,002		0,208	0,546	0,997		0,000	0,000	0,000	0,002	1,000
Hungary	24971,48		0,625	0,002	0,682	0,077	0,003	0,339	0,091	0,902	1,000	1,000	0,000	0,287	0,007	0,005	0,634	0,065
Croatia	24316,25		0,006	0,006	0,011	0,030	0,007	0,507	0,969	0,012	1,000	1,000	0,007	0,011	0,007	0,012	0,016	0,044
Latvija	22795,2		0,363				0,178		0,014	0,974	0,031	1,000	0,034	0,005	0,065	0,017	0,006	0,028
Romania	20792,17		0,005	0,005	0,016	0,011	0,005	0,007	1,000	0,090	0,980	0,997	0,702	0,539	0,008	0,004	0,982	0,487
Bulgaria	17966,16	0,029							1,000	0,997	0,748	0,254	0,012	0,016	0,015	0,006	0,007	0,007
Finland	10787,66		0,007	0,019	0,188	0,006	0,019	0,172	0,007	0,011	0,020	0,013	0,008	0,596	0,136	0,008	0,167	0,104

Table A.6: Households characteristics - Heatmap based on posterior inclusion probability (g-BRIC)

Country	GDP_P	Dfin_sit_a_142	Dfin_sit_a_143	Dfin_sit_a_144	Dfin_sit_a_145	Dfin_sit_a_146	Dfin_sit_a_147	Dfin_sit142	Dfin_sit143	Dfin_sit144	Dfin_sit145	Dfin_sit146	Dfin_sit147	DprecenceP2	DprecenceP3	DprecenceP4	DTen_status2	DTen_status3	DTen_status4	dum_duls	dum_nchilds	dum_nwork
Luxembourg	108010,7	0,083	0,001	0,000	0,002	0,003	0,006	0,141	0,023	0,002	0,009	0,011	0,426	0,001	0,016	0,004	0,001	0,073	0,218	0,002	0,987	0,952
Switzerland	66342,01	0,024	0,001	0,000	0,008	0,000	0,003	0,000	0,000	0,000	0,010	0,008	0,000	0,000	0,665	0,334	0,000	0,000	0,000	0,000	0,001	0,088
Austria	53102,95	0,014	0,001	0,002	0,470	0,003	0,939	0,010	0,002	0,395	0,035	0,002	0,042	0,005	0,111	0,001	0,002	0,014	0,043	0,001	0,097	0,916
Netherlands	52594,23	0,003	0,028	0,004	0,003	0,008	0,004	0,014	0,004	0,011	0,020	0,023	0,003	0,022	0,007		0,003	0,004	0,003	0,004	0,007	0,019
Denmark	51293,21	0,011	0,011	0,023	0,012	0,013	0,026		0,057	0,012	0,035	0,011	0,019				0,016	0,016		0,023	0,011	0,901
Germany	49684,61	0,563	0,654	0,047	0,281	0,111	0,004	0,003	0,002	0,251	0,021	0,011	0,023	0,006	0,041	0,001	0,029	0,483	0,518	0,002	0,001	0,991
Iceland	48186,17	0,004	0,021	0,004	0,003	0,004		0,012	0,005	0,004	0,006	0,225	0,009	0,005	0,005		0,011	0,080	0,160	0,004	0,004	0,004
Belgium	48153,94	0,073	0,001	0,001	0,003	0,212	0,435	0,007	0,026	0,002	0,001	0,026	0,565	0,003	0,134	0,000	0,000	0,110	0,180	0,000	0,002	0,000
Italy	42892,31	0,003	0,001	0,058	0,001	1,000		0,011	0,001	0,166	0,002	0,114					0,001	0,001		0,106	1,000	0,087
France	42864,02	0,004	0,003	0,003	0,003	0,005	0,002	0,003	0,002	0,551	0,005	0,002	0,005	0,465	0,595	0,026	0,005	0,017	0,007	0,004	0,974	0,084
United Kingdom	42294,12	0,003	0,010	0,003	0,003	0,002	0,005	0,004	0,014	0,005	0,004	0,034	0,007	0,003	0,003	0,008	0,003	0,275	0,728	0,094	0,010	0,028
Cyprus	37562,45	0,029	0,059	0,024	0,798	0,526	0,012	0,019	0,000	0,000	0,000	0,012	0,027	0,000	0,042	0,008	0,421	0,001	0,005	0,006	0,077	0,335
Spain	36884,26	0,001	0,000	0,003	0,830	0,181	0,001	0,001	0,003	0,009	1,000	0,999	0,001	0,001	0,996	0,076	0,065	0,001	0,006	0,004	0,057	0,001
Czech Republic	34002,19	0,004	0,004	0,620	0,004	0,004	0,008	0,036	0,005	0,006	0,015	0,003		0,005	0,731	0,005	0,012	0,046		0,018	1,000	0,007
Malta	32881,83	0,004	0,008	0,004	0,005	0,037	0,068	0,006	0,004	0,004	0,006	0,190	0,004	0,006	0,062	0,007	0,038	0,076	0,786	0,006	0,019	0,866
Portugal	31304,82	0,762	0,433	0,034	0,003	0,033	0,141	0,015	0,003	0,030	0,703	0,883	0,003	0,003	0,298	0,007	0,006	0,002	0,004	0,002	0,945	0,892
Greece	30372,76	0,129	0,028	0,004	0,040	0,005	0,276	0,021	0,007	0,008	0,004	0,036	0,756	0,004	0,022	0,010	0,004	0,008	0,006	0,005	0,895	0,037
Estonia	28159,58	0,007	0,013	0,022	0,129	0,816	0,031	0,009	0,007	0,013	0,017	0,241	0,031	0,022	0,018	0,010	0,028	0,726	0,280	0,074	0,023	0,012
Slovakia	26107,83	0,124	0,019	0,005	0,002	0,003	0,002	0,006	0,242	0,003	0,011	0,003	0,035	0,003	0,003	0,004	0,010	0,005	0,024	0,004	0,003	0,004
Lithuania	25968,58	1,000	0,065	0,083	0,195	0,251	0,324	0,004	0,418	0,458	0,206	0,136	0,073	0,010	0,399	0,003	0,002	0,003	0,003	0,003	0,275	0,004
Poland	25124,28		0,020	0,002	0,048	0,561	0,000		0,128	0,079	0,886	0,865	0,001	0,001	0,005	0,001		0,000	0,000	0,337	0,857	0,006
Hungary	24971,48	0,948	0,764	0,004	0,003	0,495	0,150	0,052	0,089	0,005	0,009	0,006	0,219	0,002	0,008	0,003	0,001	0,867	0,049	0,001	1,000	0,001
Croatia	24316,25	0,088	0,007	0,006	0,006	0,065	0,009	0,852	0,012	0,011	0,006	0,023	0,033	0,011	0,196	0,048	0,007	0,016	0,024	0,006	0,018	0,008
Latvija	22795,2		0,006	0,008	0,017	0,440	0,081		0,006	0,006	0,023	0,018	0,275	0,014	0,042	0,005		0,006	0,005	0,005	0,028	0,006
Romania	20792,17	0,200	0,979	0,006	0,007	0,019	0,026	0,865	0,017	0,006	0,005	0,010	0,978	0,005	0,006	0,005	0,005	1,000	0,004	0,698	0,237	0,123
Bulgaria	17966,16	0,023	0,922	0,017	0,018	0,013	0,185	0,014	0,050	0,057	0,019	0,008	0,171				0,009	0,084	0,017	0,071	1,000	0,007
Finland	10787,66	0,006	0,006	0,025	0,008	0,013	0,007	0,006	0,006	0,007	0,007	0,007	1,000	0,010	0,017	0,006		0,007	0,008	0,018	0,008	0,041