Inheritances and Wealth Inequality: A Machine Learning Approach[1]

Pedro Salas-Rojo
Universidad Complutense de Madrid (Spain), ICAE and EQUALITAS
Tel: +34 91 3942626. E-mail: pedsalas@ucm.es

Juan Gabriel Rodríguez
Universidad Complutense de Madrid (Spain), ICAE, EQUALITAS and CEDESOG
Tel: +34 91 3942516. E-mail: juangabr@ucm.es

Abstract

This paper explores the relationship between received inheritances and the distribution of wealth (financial, non-financial and total) in four developed countries: the United States, Canada, Italy and Spain. We follow the inequality of opportunity (IOp) literature and −considering inheritances as the only circumstance− we show that traditional IOp approaches can lead to non-robust and arbitrary measures of IOp depending on discretionary cut-off choices of a continuous circumstance such as inheritances. To overcome this limitation, we apply Machine Learning methods ('random forest' algorithm) to optimize the choice of cut-offs and we find that IOp explains over 60% of wealth inequality in the US and Spain (using the Gini coefficient), and more than 40% in Italy and Canada. Including parental education as an additional circumstance −available for the US and Italy− we find that inheritances are still the main contributor. Finally, using the S-Gini index with different parameters to weight different parts of the distribution, we find that the effect of inheritances is more prominent at the middle of the wealth distribution, while parental education is more important for the asset-poor.

JEL Codes: C60, D31, D63, G51.

Keywords: Wealth inequality; inheritances; Machine Learning; inequality of opportunity; parental education.

[1] We acknowledge comments from LIS Datacenter seminar participants, Juan César Palomino, Dmitry Petrov, Paolo Brunori, Amaia Palencia, Alberto Palacios, Pablo de la Vega, Jordi Domenech, Josep Espasa, Piotr Paradowski, Johannes Zahnner and participants at the 14th RGS Conference, the XXVIII Encuentro de Economía Pública, the 9th ECINEQ Meeting and the 2021 EsCOE Conference on Economic Measurement. Two anonymous referees also provided very useful suggestions. This paper has received financial support from the Ministerio de Economía y Competitividad of Spain under project PID2019-107161GB-C33) and Comunidad de Madrid (Spain) under project H2019/HUM-5793-OPINBI-CM. Salas-Rojo also received financial support from the Horizon 2020 Research and Innovation Programme under Grant Agreement 730998, the InGRID-2 project, and the Ministerio de Ciencia, Innovación y Universidades through FPU grant 17/735. The usual disclaimer applies.

1. Introduction

Wealth inequality is on the rise. This persisting trend, which has been already documented for a few decades, has progressively shifted the focus of the literature to the study of its main causes. For instance, Lusardi et al. (2017) find that those with better financial knowledge extract higher profits from investments, and hence are able to accumulate higher wealth stocks. Others, such as Alstadsaeter et al. (2017), claim that wealth evasion through tax havens rises the net return to capital of the very rich, allowing them to accumulate more wealth. Meanwhile, Zucman (2019) points at the fall in progressive taxation, which has hindered the effect of distributional policies. Surprisingly, there is still a large dissent on the role that intergenerational wealth transmission through inheritances plays on shaping the wealth distribution. This paper contributes to the matter by combining the inequality of opportunity approach with Machine Learning (ML) techniques that improve the robustness of the results, and therefore can better enlighten how received bequests affect households' opportunities to accumulate wealth.

Part of the literature has found that inheritances are associated with a relevant part of wealth inequality. For instance, Piketty (2011), Piketty and Zucman (2015) and Alvaredo et al. (2017) provide empirical evidence on the rising shares of wealth accumulated at the top of the distribution, where inheritances are greater and more prevalent, and serve as a vehicle used by the very rich to channel wealth through generations. In the same vein, Fessler and Schurz (2018) claim that intergenerational wealth transfers are among the largest contributors to net wealth inequality, with other authors, such as De Nardi and Yang (2016), Adermon et al. (2018), Palomino et al. (2021) and Nolan et al. (2020) finding a positive contribution of intergenerational transfers to wealth inequality applying different agent-based models and econometric techniques.

However, other empirical evidence finds divergent results. Wolff and Gittleman (2014) for the United States, Crawford and Hood (2016) and Karagiannaki (2017) for the United Kingdom, and Elinder et al. (2018) for Sweden coincide on finding an equalizing net effect of inheritances upon the wealth distribution. They find that, as inheritances are more equally distributed than wealth, intergenerational transmission actually produces a net decrease in overall relative inequality. In the same vein, Boserup et al. (2016) find that, despite inheritances increase absolute inequality, this effect is not reflected in relative inequality measures calculated at the top of the distribution.

Aiming to contribute to the debate, we propose a different approach based on the Inequality of Opportunity (IOp) literature (see also Palomino et al., 2021). Following Roemer (1993) and Van de Gaer (1993) any economic outcome such as wealth, income or health status is the result of the interaction between two sets of factors. On the one hand, exogenous factors beyond the individuals' control, such as sex, parental education, race or the inheritances received. They are called *circumstances* because they are out of individuals' responsibility. By considering different combinations of these factors, we can divide the population into a set of mutually exclusive and exhaustive groups, called *types*. On the other hand, the remaining factors are considered to be endogenous, as they are within the individuals' set of choices. It is the case, for instance, of the number of hours worked or the nutritional habits. They are called *efforts*.

We can then decompose overall inequality in two terms: inequality of opportunity (IOp, the component attributed to circumstances) and the remaining component, usually referred to as inequality of efforts (IE). The first component is found to be undesirable from a social justice perspective, as this "unfair" inequality cannot be palliated by individual choices (Rawls, 1971; Sen, 1980). In addition, the literature has found IOp to have a negative effect on economic growth, because hindering individual opportunities for education and work causes a misallocation of talent (Marrero and Rodríguez, 2013 and 2019; Bradbury and Triest, 2016; Carranza, 2020; Marrero et al., 2021).

In this paper we consider inheritances and gifts to be circumstances whose transmission is independent from the behavior and responsibility of the recipient.[2] As such, this variable can be used to divide the population into types, allowing us to estimate IOp as the between-type component of overall wealth inequality. The idea behind this proposal is that the IOp component measures the influence of circumstances (in our case, bequests) on the final distribution of wealth. Thus, if overall wealth inequality was independent from the inheritances and gifts received, there should be no dispersion between types. Otherwise, bequests would have a role to shape the observed distribution of wealth.

Wealth is, by definition, a stock variable accumulated through life, so it is affected by life-cycle and behavioral dynamics. Analyzing how inheritances affect consumption, the saving paths or capitalization would require, for instance, explicit behavioral and agent-based models, natural experiments or detailed panel data, which are usually hard to find (Andersen and Nielsen, 2011; De Nardi, 2015). However, in this paper, these effects are already included in the wealth measures under consideration, because wealth surveyed at a certain point of life reflects all past decisions and endogenous dynamics. Thus, once we control for the size of the household, sex and age, the estimated types allow us to approximate the impact of circumstances on the wealth distribution.[3]

[2] Wolff and Laferrere (2006) propose three mechanisms to model intergenerational family transfers. First, according to the pure altruistic model, parents' utility or welfare is directly augmented by the utility of their children. Second, the impure altruistic model proposes that parents expect their children to behave in a certain way, so they obtain utility from a particular set of elements of their child's consumption or leisure time. Finally, the authors propose a non-altruistic model in which intra-family transfers correspond to explicit reciprocity contracts, so both utilities are connected and depend on one another. The data available impedes to explore in depth the relation between bequests and their recipients, so we –as most approaches in the wealth inequality and inheritances literature- assume the postulates of the pure altruistic model.

[3] Our measures are lower bounds estimates of IOp due to the limited set of circumstances available in our data. Also note that inheritances and parental education are not orthogonal to other factors such as race or place of birth during childhood. Consequently, the effects from these uncontrolled circumstances will partially overlap with our controlled variables and, therefore, will be collected by our measures.

Bearing all this in mind, and before the estimation of the between-type inequality component is conducted, an important issue must be solved. The definition of types, essential to calculate the between-type inequality component, is straightforward and easy for categorical circumstances such as parental education or sex. However, we use the value of the inheritances received. This variable being a highly skewed continuous circumstance, the generation of consistent types is more difficult. Indeed, our first results show that traditional empirical approaches to measure IOp, such as the ex-ante parametric method (Ferreira and Guignoux, 2011), lead to arbitrary and non-robust IOp estimates when ad hoc discretizations of inheritances are applied.

To solve this limitation of the traditional IOp methods and generate more robust estimates, we propose the use of Machine Learning (ML) algorithms. These computing techniques extract statistical information from the data, such as their distribution or correlation between variables, limiting the biases introduced by the researcher. To date, ML has been used in the IOp framework to select the relevant circumstances among a wide range of candidates (Brunori et al. 2019, Brunori et al. 2020, Brunori and Neidhöfer, 2020).[4] Here, we propose two ML techniques to discretize a highly skewed continuous circumstance, like the inheritances received, and thus generate statistically meaningful types. We employ *Conditional Inference Trees* (Hothorn et al. 2006), in which the partition of inheritances is based on the relation between the dependent (wealth) and the independent variables (the inheritances received), and also *Conditional Inference Random Forests* (Strobl et al. 2007) as the bootstrapped version of the tree algorithm. By means of these two methods, we analyze how the inheritances received are able to measure the opportunities to accumulate total, financial and non-financial net wealth.

We thus measure wealth IOp in the United States, Canada, Italy and Spain, data comes from the Luxembourg Wealth Survey (LWS), one of the most comprehensive databases containing information on wealth and inheritances. Rather than focusing on just one country, analyzing a set of developed countries may help us to check if our results are robust to significantly different welfare and fiscal systems. On the one hand, it could be expected that countries with a stronger public provision of services and social security schemes equalize more the households' opportunities to accumulate wealth, hence smoothing the impact of inheritances. On the other hand, these schemes can disincentive wealth accumulation since future consumption is secured by the State. In addition, higher taxes on wealth and inheritances may hinder intergenerational wealth transfers and its later accumulation, potentially affecting both economic growth and inequality. With these ideas in mind, we briefly overview the main features of these four welfare and fiscal systems.

[4] A precedent of the learning methods in the IOp literature is found in Li Donni et al. (2015) where social types were based on the statistical properties of all circumstances by a latent classes approach.

Both south European economies, Italy and Spain, have large welfare state systems, with a strong public provision of services and social security schemes. However, they differ on the fiscal treatment of inheritances and wealth. While Italy has a national inheritances tax (*Imposta sulle Successioni e Donazioni*) ranging between 4% and 8% of the total amount inherited, the application of the inheritances tax in Spain (*Impuesto de Sucesiones y Donaciones*) is heterogeneous across regions. Moreover, Italy applies a wealth tax exclusively over assets held out of the country (*Imposta sul valore degli immobile situati all'estero*, and *Imposta sul valore delle Attivita Finanziarie detenute all'estero*), while Spain taxes both, assets held inside and outside the country (*Impuesto sobre el Patrimonio*), with its precise application being assigned to the regions, who define minimum exemptions and tax rates.

North American economies also present remarkable differences. Regarding the welfare state system, Canada is similar to the European countries while, in the US, the public sector does not guarantee free access either to sanitation nor tertiary education, and the private initiative is prevalent in pension schemes. Regarding the fiscal system, both countries have property taxes levied at the state/province and municipal levels, which compensate the absence of other wealth taxes. Regarding intergenerational transfers, Canada has no specific inheritance tax, while the US mixes the Italian and the Spanish models with a national tax modified through conditional deductions implemented at the State level (*Estate Tax*).

Our main unit of analysis is the household, for whom we take the correspondent wealth and the inheritances received. Nonetheless, the demographic controls –age and gender– are taken from the household head. Moreover, parental education –a circumstance highly related to wealth inequality (Adermon et al., 2018; Palomino et al., 2021)– is also obtained from the household head. Despite data limitations restrict our analysis for parental education to Italy and the US, the differences across these two countries make the analysis of particular interest. While the educational system, including tertiary education, is free in Italy, accessing the university in the US is often expensive and even unfordable for many individuals. Consequently, achieving higher levels of education might probably be strongly dependent on the opportunities defined by previous generations.

Using inheritances as the only circumstance, our estimates of IOp explain more than 60% of overall wealth inequality (measured by the Gini coefficient) in the US and Spain, with ratios reaching 71% (US) or 68% (Spain) of financial wealth and 71% (US) or 72% (Spain) of non-financial wealth. Lower rates are found in Canada and Italy, where around 45% of overall wealth inequality is explained by our IOp estimates, rising to 61% (Canada) and 53% (Italy) in financial wealth, and descending to 51% (Canada) or 43% (Italy) in non-financial wealth. Thus, IOp is found to explain the disparate opportunities that households bear to accumulate wealth. In particular, we highlight that financial wealth inequality is more affected by IOp than non-financial wealth, maybe because inheritances act as security nets that allow households to face riskier investments.

Including the parental education in the analysis provides a deeper insight. In the US, the share of financial and non-financial wealth inequality explained by IOp then rises to 76% and 75%, respectively, while in Italy the ratios rise to 50% for total and non-financial, and 63% for financial wealth. As inheritances and parental education can be highly related (Adermon et al. 2018), we analyze the potential overlapped effects with a Shapley value decomposition (Sastre and Trannoy, 2002; Rodríguez, 2004; Shorrocks, 2013). In Italy, around one third of wealth IOp is attributed to parental education while, for the US, the ratio is less than one-fifth for any wealth IOp measure. Although both circumstances overlap part of the effects, inheritances channel a higher share of the households' uneven opportunities.

Finally, we explore whether the effects captured by the inheritances received are homogeneous along the wealth distribution. To do so, we modify the relative weight given to the observations at different parts of that distribution with the Single Parameter Gini index (Donaldson and Weymark, 1980; Yitzhaki, 1983) for several parameters of inequality aversion. We find that, in the four countries analyzed, the inheritances received are more important for the middle and upper tail households, hence fostering their opportunities to accumulate higher levels of wealth. The asset poor receive small inheritances, so there are other factors determining inequalities across them. Thus, we find that parental education for the US and Italy is highly important for the opportunities of people at the bottom of the wealth distribution.

The article is structured as follows. Section 2 introduces the wealth IOp theoretical framework and the ML algorithms employed in the paper. Section 3 describes the LWS database and the adjustments applied to the data. In Section 4 we present the main results for inheritances and parental education, while in Section 5 we explore the effect of these two circumstances along the wealth distribution. Finally, Section 6 highlights our main conclusions.

2. IOp measurement and Machine Learnings techniques

In this section we first present the IOp framework to measure the influence of inheritances on wealth inequality, and then justify and explain the implementation of Machine Learning techniques to overcome some limitations of traditional IOp approaches.

2.1. The inequality of opportunity approach

Consider a population of discrete individuals indexed by $i \in \{1, \dots, N\}$ and a variable w characterizing our economic outcome of interest, wealth, whose distribution is a function of the set of circumstances faced by the individual, C_i, and the amount of effort exerted, e_i, such that

$w_i = f(C_i, e_i)$. Circumstances are defined as a finite discrete vector of J elements and are assumed to be exogenous because they cannot be affected by individual's choices. Simultaneously, we consider effort to be a continuous variable that depends on both, personal decisions and circumstances, such that individual wealth can be rewritten as $w_i = f[C_i, e_i(C_i)]$.

Then, the population is divided into M exhaustive and mutually exclusive groups, called types, $\prod = \{T_1, \dots, T_M\}$, such that all individuals belonging to the same type T_m share the same circumstances: $T_1 \cup T_2 \cup \dots \cup T_M = \{1, \dots, N\}$, $T_r \cap T_s = \emptyset, \forall r, s,$ and $C_i = C_j, \forall i, j | i, j \in T_m, \forall m$.

Because wealth is a continuous variable, a simple way to assess the effect of circumstances (in our case, the inheritances received) on overall wealth inequality could be to compare the density function of *w* across types. Then, circumstances would have no role on the final distribution of wealth if:

$$\int w |\, T_m d_{T_m} = \int w |\, T_s d_{T_s}, \qquad \forall m, s | T_m \in \Pi, T_s \in \Pi, \tag{1}$$

where subscripts *m* and *s* indicate two different types.

Individual circumstances are irrelevant to explain wealth inequality if the distribution of wealth across types is the same. In this case the individual's outcome is independent of her circumstances. Otherwise, individual non-responsibility factors are relevant and contribute to shape the final distribution of wealth. The problem with this approach is that, in general, comparisons between distributions do not fulfill the stochastic dominance property (Atkinson, 1970), because they can be significantly different and yet cross each other. To avoid this problem and break potential ties, a common alternative is to focus on a particular moment, typically the mean. Following the ex-ante approach in Van de Gaer (1993) circumstances have no role on the final distribution of wealth if:

$$\overline{w}_s = \overline{w}_m, \quad \forall m, s | T_m, T_s \in \Pi, \tag{2}$$

where the mean wealth in a type, $\overline{w}$, is the expected wealth value that an individual can get conditioned on her type.[5] The ex-ante method requires the calculation of a counterfactual smoothed distribution $\widehat{w}$ assigning to each observation the mean wealth of her type. Applying an inequality measure $I(\cdot)$ to this counterfactual distribution gives us the part of overall inequality attributed to the set of observed circumstances (absolute IOp). Dividing this measure over total inequality reveals the share of overall inequality explained by the set of circumstances (relative IOp):

[4] A different method, the ex-post approach, compares the average outcome across types for individuals who have exerted the same *degree of effort*, i.e., who belong to the same quantile of the type distribution. We have not followed this approach because it requires strong assumptions on the definition of effort.

$$Absolute\ IOp = I(\{\widehat{w}\}) \tag{3}$$

$$Relative\ IOp = \frac{I(\{\widehat{w}\})}{I(\{w\})} \tag{4}$$

For our analysis we use the Gini coefficient but, for robustness, we also employ the Mean Logarithmic Deviation (MLD). The IOp literature has generally leaned towards the MLD, mainly because it is the only additively and path independent decomposable inequality index (Foster and Shneyerov, 2000). However, we focus our analysis on the Gini coefficient because it is the most extended measure on wealth inequality (Cowell and Van Kerm, 2015), so our results can easily be compared with the related literature. Moreover, the Gini index has recently been applied in the IOp framework by Lefranc et al. (2008), Brunori et al. (2019) and Cabrera et al. (2021) among others, highlighting some of their advantages. First, IOp indices measure inequality based on the distribution of means across types (equation (2)), so extreme values are by construction smoothed. However, the MLD is more sensitive to extreme values than the Gini coefficient, so the transformation of the original distribution into the smoothed distribution has a larger effect on the former than in the latter inequality index. Consequently, estimates based on the MLD are likely to under-estimate wealth IOp. Second, as we want to calculate the size of IOp caused by inheritances, we are not particularly interested in decomposing overall inequality into IOp and IE, so the Gini index being not additively decomposable is not an issue.[6]

The main problem we face to analyze the effect of inheritances on the wealth distribution −through the IOp framework− lies on the construction of the smoothed distribution $\widehat{w}$. The literature provides several ex-ante methods, but for comparability reasons we lean towards the standard method, so we consider the parametric approach proposed by Ferreira and Guignoux (2011) (see also Bourguignon et al., 2007).[7] This method exploits the reduced form of an OLS regression in which the outcome variable is regressed against the vector of available circumstances C_i:

$$\ln(w_i) = \alpha + \varphi C_i + \varepsilon_i. \tag{5}$$

Assuming that the estimates of φ are satisfactory, we obtain the smoothed vector $\widehat{w}$ by fitting the parameters of equation (5):

$$\overline{w}_i = \exp(\hat{\alpha} + \hat{\varphi} C_i) \tag{6}$$

[6] The MLD index has another limitation: it cannot deal with negative or zero values. As described in the data section, in this paper we are focused on net wealth which, by definition, can also be negative and zero. Consequently, to obtain IOp measures for the MLD index we need to exclude negative and zero wealth values which reduces the samples by around 5% in Canada, Italy and Spain, and 8% in the U.S.. Aimed to provide more evidence on the robustness of our estimates, Table OA3 in the Online Appendix replicates the main results obtained with the Gini index for these restricted samples.

[7] We have also replicated the whole analysis for the non-parametric approach proposed by Checchi and Peragine (2010), and the main conclusions remain similar. These results can be found in the Online Appendix (Table OA1).

Notice that when the set of circumstances is defined with discrete or categorical variables the creation of types is straightforward. But when one circumstance is continuous and highly skewed, such as the value of the inheritances received (see below), we face a problem. If we include in equation (5) a continuous circumstance without any previous discretization, when predicting vector $\widehat{w}$ we generate one type for each possible value of the circumstance. This would lead to an over fitted number of types, producing upward-biased IOp estimates (see Appendix A in Brunori et al., 2019). Furthermore, discretizing that variable under the ad-hoc criterion of the researcher might lead to non-robust results. For instance, taking the quintiles of the distribution of inheritances as cutting points can generate rather different smoothed vectors $\widehat{w}$, as non-linearities at the top tail might affect differently the imputed mean values. This can lead to arbitrary between-type inequality measures, impeding accurate comparisons across countries or precise policy assessments. To palliate this problem, we propose the use of Machine Learning techniques that base the discretization of the continuous circumstance on the statistical properties of the wealth distribution.

2.2. Machine Learning and Inequality of Opportunity

We call Machine or Statistical Learning to the stream of computing techniques that take information from the data, identify patterns and make statistical decisions with the smallest possible human intervention. The main idea of these algorithms lies on the motto “let the data talk”, as they avoid potential biases introduced by researchers such as the selection of variables, the level of significance to obtain statistical inference or the way in which we discretize continuous variables. Indeed, the discretization problem previously described is not exclusive of the IOp literature, as it is highlighted by the use of ML techniques in many different fields, such as biomedicine (Lutsgarten et al., 2011), genetics (Gallo et al., 2016), the stock market dynamics (Lalithendra and Prasad, 2018) or the price of gold (Banerjee et al., 2019).

In our context, we employ supervised Machine Learning algorithms that explore the relation between the distribution of the dependent variable (in our case, wealth) and the independent variables (the inheritances received and parental education). In particular, we lean towards conditional inference trees and forests (Hothorn et al., 2006), which include a splitting algorithm that subsequently divide the continuous variables into discrete categories. Although trees and forests are similar, they present a difference that justifies their separate implementation. Trees are usually run just once over the data, and forests are their bootstrapped version, which correct for potential inconsistencies attributed to the particular data structure. As the underlying mechanism of forests cannot be explained without understanding that of trees, the remainder of the section is devoted to present both.

Trees

Tree algorithms divide a dataset on exhaustive and mutually exclusive groups of observations, based on sequential and hierarchical decisions given some statistical criteria. Once all partitions are performed, the algorithm imputes to each observation the average value of the objective dependent variable considered, conditioned to the group it belongs. Its application to the IOp framework is straightforward, as these methods directly generate types based on the distribution of the dependent variable across the observed circumstances and then, assign to each individual the mean value of the dependent variable conditioned on her type.

From the family of tree-based methods we select those based on conditional inference algorithms, which have the advantage of not being biased towards splitting only continuous variables, as other tree-based techniques usually do (Hothorn et al. 2006). Furthermore, they have already been used in the IOp framework to select the relevant set of circumstances (Brunori et al. 2019; Brunori et al. 2020; Brunori and Neidhöfer, 2020) that define household opportunities.

The functioning of the algorithm can be explained in three consecutive steps:

1. First, it performs a t-test on the global null hypothesis of independence for each circumstance considered, at some alpha-value of significance.

$$H_0^C = D(w|C) = D(w) \tag{7}$$

For each circumstance, the algorithm provides a p-statistic, which needs to be adjusted to avoid type I errors. In this paper we apply the Bonferroni correction, quite common for multiple hypothesis testing:[8]

$$p_{adj} = 1 - (1 - p)^p \tag{8}$$

The algorithm selects the circumstance with the lowest p-adjusted value, i.e., the one with the strongest association to the dependent variable w. If $p_{adj} > \alpha$, the algorithm stops. Otherwise, it continues by setting the selected circumstance as a splitting variable.

2. Once we know that a circumstance is a splitting variable, conditional inference trees decide the cutting points. Note that when the variable is binary this step is trivial, as there is only one way to be split. But, when the variable is continuous, the algorithm needs to check all potential partitions. Thus, let us consider:

$$w_z = \{w_i : C_i < x\} \tag{9}$$

[8] Apart from the Bonferroni correction, for robustness, we also followed Genz (1992) and checked our results with a Montecarlo adjustment with 10,000 iterations. The results did not change and they can be found in Table OA2 in the Online Appendix.

$$w_{-z} = \{w_i : C_i \geq x\}, \tag{10}$$

where x defines each possible value in the continuous variable, and z the resulting subsamples. For every x, the tree algorithms test the discrepancy between both subsamples, applying a difference-in-means t-test and obtaining an associated p-value. Finally, the algorithm selects the splitting point delivered by the smallest p-value.

3. The whole process is repeated for each subsample or node, until the null hypothesis of independence cannot be rejected.[9] In that moment, the algorithm stops, and the tree is drawn. Finally, the algorithm attributes to each observation its expected wealth, conditioned on the type she belongs, allowing us to estimate IOp as defined in equation (2).

The main shortcoming of tree-based algorithms is that their results are highly dependent on several factors. One of them is the selected alpha that rejects or accepts the null hypothesis defined in equation (7). To reduce this problem and follow the spirit of ML techniques, we use an endogenously tuned alpha, obtained by the application of K-fold Cross-Validation (see the Technical Appendix). This endogenous alpha eliminates the external judgement of the researcher on setting a particular level of critical significance. Nevertheless, and following the canonical stream in econometrics, we test the robustness of our results by also presenting those for 0.05 and 0.01 alphas.

Another relevant factor for these algorithms is the data structure, i.e., the number of variables included, their correlation or their distribution (Friedman et al., 2009). Imagine the case of two highly correlated circumstances, where one delivers a slightly lower p-value than the other. In that case, the other circumstance might disappear from the splitting process, despite being almost as important as the selected one. A similar situation could be found when deciding the splitting point in step 2, in which the distribution leads to two similar cutting points. As a result, predictions inferred directly from trees are fairly sensitive to alterations in the data structure. Overall, tree methods usually perform well in-sample, but using them for out-of-sample inferences bear reasonable doubts, so we need to include a more complete technique in the analysis to reinforce the robustness of our results.

[9] Indeed, every time we test the null hypothesis of independence in the first step, we are actually testing equality of opportunity, following the same ideas expressed in equations (1) and (2). Rejecting independence implies that the distribution of an outcome variable is significantly conditioned by a certain circumstance, hence rejecting equality of opportunity (Brunori et al., 2020).

Forests

Conditional inference forests implement bootstrapping ideas into the ML framework.[10] The algorithm generates a certain number of conditional inference trees, averaging their results. The repeated extraction of subsamples guarantees the independence of each tree, so each one provides different estimates. In the end, the law of large numbers is applied to smooth the discrepancies between the constructed trees, providing a distribution consistent with the out-of-sample reality (see Schlosser et al., 2019, for a discussion).

Each tree inside the random forest grows following the same three-step structure previously explained. Nevertheless, this algorithm bears some particularities. As explained in Brunori et al. (2020), three factors determine how these forests grow. First, to control for the exclusion of highly correlated independent variables, each tree is generally grown after a random selection of circumstances. However, in our case this is not a problem, since we only have the value of the inheritances received and, when available, parental education –see the data section below-. Second, we need to consider the number of trees grown in each forest. For robustness, we have checked the results for 100, 200 and 500 trees. Finally, same as for conditional inference trees, we apply the method not only for the endogenously tuned alpha, but also for values of alpha equal to 0.05 and 0.01.

3. Database and adjustments

The data comes from the Luxembourg Wealth Study (LWS) Database, provided by the Luxembourg Income Survey (LIS) cross-national data center.[11] From the available set of countries, we present results for Canada (Survey of Financial Securities 2016), Italy (Survey of Household Income and Wealth 2014), Spain (Survey of Household Finances 2014), and the US (Survey of Consumer Finances 2016), as they are those with the most extensive data on inheritances received.[12] Another advantage of this selection of countries, as said in the Introduction, is that they present quite different welfare and fiscal systems.

[10] Despite random forests and bootstrapping are similar, they differ on a relevant aspect: the sampling process in random forests is performed without replacement because it can lead to biased results, as described in Strobl et al. (2007) and Strobl et al. (2009).
[11] LIS is a non-profit organization whose main mission consists on acquiring datasets and harmonizing them, easing cross-country comparisons. All the relevant information of the institution and the data can be found at https://www.lisdatacenter.org/.
[12] We also tried to include data from other countries, such as Austria (Household Finance and Consumption Survey, 2014), Norway (Household Wealth Statistics, 2016) and the United Kingdom (Wealth and Assets Survey, 2015). However, the limited number of valid observations with inheritances –always less than 10% of the total sample- led us to inaccurate results.

Our analysis is based on three different dependent variables. First, *non-financial wealth is* defined as the combined market value of the principal and all secondary residences, business equities, durable consumer goods and other non-financial assets owned by household members minus the sum of all real estate liabilities. Second, *financial wealth* collects the market value of all deposit accounts, cash, bonds and debt securities, stocks and other equities, investment funds and other non-pension liquid assets net of all financial liabilities. Pension assets are excluded from financial wealth. As said, we take advantage of the data harmonization performed by the LWS, which permits us to perform straightforward cross-country comparisons. However, this data homogenization also implies some shortcomings, such as the exclusion just mentioned, which potentially depresses the aggregate shares of financial wealth over total wealth. Unfortunately, we cannot be certain about the size or direction of the effect generated by this exclusion. On the one hand, private pension schemes are usually more concentrated on the wealthy, so their exclusion might downward bias our inequality estimates. On the other hand, social security pension schemes are usually more homogeneously distributed than other assets, which could equalize wealth (Wolff, 2015).

The third dependent variable is *total wealth,* which is the sum of the two previous wealth concepts. Distinguishing between the first two wealth variables provides an insightful view on the unequal opportunities that people face to invest and accumulate certain assets, while analyzing total wealth leads to a more comprehensive global inequality analysis. Because gross wealth ignores the role of intergenerational transfers on homeownership and investment behavior, we have considered net worth as our wealth definition (Boserup et al., 2016; Karagiannaki et al., 2018; Adermon et al., 2018; Nolan et al., 2020).[13] Nonetheless, the results for gross wealth are shown in Table OA4 to OA7 in the Online Appendix. Overall, we find that the dichotomy between gross and net wealth does not significantly modify the IOp estimates.[14] All monetary variables are expressed in dollars of 2017, as we use the PPP conversion rates provided by the LWS.

Throughout the whole paper the unit of analysis is the household, as this is the level at which wealth and inheritances are usually reported. Nevertheless, for the household demographics, such as age and gender, and for parental education, we consider the values reported by the household head, who is defined as the person designated by the household for the purposes of replying to

---

[13] For instance, Maroto and Severson (2019) report the increasing difficulties faced by Canadian young adults to emancipate, highlighting how the favorable parental background favored the households' homeownership opportunities. In the same vein, Mulder et al. (2015) explore the role of intergenerational transmissions on homeownership in Europe. They find that, especially in southern Europe, parental wealth is used to access mortgage credits and, in many cases, as a direct transfer or gift to defray part of the housing costs. Similar results were found for U.S. in Listokin et al. (2001), where parental assets are also used to pay high university fees, hence easing their descendants' human capital accumulation.

[14] We think that the dichotomy between net wealth and gross wealth is not an important cause for the divergent results found in the literature about the relationship between wealth and inheritances because our main results for both concepts are aligned to the same conclusion (inheritances are not equalizing).

the survey.[15] To make all countries comparable in terms of age, we restrict our final sample to households aged between 35 and 80 years old.

Analyzing stock variables such as wealth requires some previous adjustments. First, despite that the use of equivalence scales to work with households of different size is unclear (Cowell and Van Kerm, 2015), we adopt a commonly used scale of equivalence, the square root of the number of household members (Buhmann et al. 1988). Second, due to the high skewness that characterize wealth distributions, it is advisable to implement a non-linear transformation. Although the logarithm is the most common adjustment tool used in the income IOp literature, it cannot be used to smooth household net wealth, as this variable can take negative values if liabilities are higher than gross assets. Hence, following Ravallion (2017), we adjust wealth with the inverse hyperbolic sine transformation, which is also defined for non-positive values.[16]

Finally, being a woman and having a certain age are also relevant circumstances. Both factors are, by definition, beyond the respondents' control and strongly associated to the wealth distribution. For instance, the gender wealth gap in Europe varies from a lower bound set around 27% in Slovakia to the upper bound set around 48% in Greece, finding countries such as France (44%), Austria (45%), and Germany (47%) in between (Sierminska et al., 2010; Schneebaum et al., 2018). In addition, wealth is by nature strongly related to life cycle dynamics. Therefore, to compare households whose heads differ in their gender and age we must adjust our dependent variable.

Following Palomino et al. (2021) and Salas-Rojo and Rodríguez (2021), the adjustment is three-fold. First, we control by the gender of the household head. Second, we center wealth at the age of 65 which is, on average, the moment in which people retire and start de-investing. Small changes on that centering point, ranging 63-67 years old, have also been applied without finding remarkable differences on the resulting wealth distributions.[17] Third, we consider the possible

---

[15] The use of the head's attributes to characterize the household is common in the IOp literature (Ferreira and Guignoux, 2011; Brunori et al., 2019) as the partners' circumstances are not usually surveyed. To the best of our knowledge, the effects of this prevalent data limitation have not been empirically assessed. In our case, the impact is probably small. It can be seen in Table 1 (below) that both genders are evenly represented in our data, i.e., our samples seem not being gender-biased towards men or women household heads. Moreover, couples usually have similar ages, and it is reasonable to expect that age deviations between partners are, on aggregate, cancelled out. In addition, homogamy and assortative mating may resemble the parental education across couples.

[16] The inverse hyperbolic sine transformation turns a variable w_i into $\widehat{w}_i = \frac{\log\left(\theta\, w_i + \sqrt{\theta^2 w_i^2 + 1}\right)}{\theta}$, being w any wealth definition and θ a scale parameter. Moreover, as in most economic applications, we set θ equal to 1 (Bellégo et al., 2021). When the scale parameter is set below 1, to values like 0.1, 0.01 or 0.001, the transformation smooths the highest wealth observations, which reduces the level of inequality and IOp estimates. On the other hand, setting the scale parameter above 1 does not meaningfully alter the results.

[17] In 2018, the normal retirement ages in these countries were: 65 years in Canada, 67 years in Italy, 65 in Spain and 66 in the U.S. (OECD, 2019). Using 67 and 66 to adjust wealth in Italy and the U.S. have no significant impact on our results. However, because it is the most common threshold in the literature, we maintain the adjustment to 65 for all countries.

interaction between both factors, age and gender. To this end, the following regression is proposed:

$$w_i = \alpha + \beta F_i + \sum_{n=1}^{4} \gamma_n (A_i - 65)^n + \sum_{n=1}^{4} \delta_n F_i (A_i - 65)^n + \varepsilon_i, \tag{11}$$

where the dummy variable F_i is 1 when the household heads is a woman and A_i expresses the age of the household head. The forth-degree specification represents the life-cycle non-linear dynamics on wealth, as suggested in Solon (1992) (see also Palomino et al., 2021).

Table 1 deploys the summary statistics for age and gender. Overall, the mean age ranges between 50 and 60, with standard deviations surrounding the 16-17 years. Moreover, we observe that our samples are evenly distributed by gender.

Table 1. Summary statistics for the discrete variables.

CANADA (N=3,627)		
Variable	Mean	Sd
Age	50.93	19.61
Gender	0.51	0.50
ITALY (N=5,638)		
Variable	Mean	Sd
Age	61.74	17.26
Gender	0.56	0.39
SPAIN (N=4,792)		
Variable	Mean	Sd
Age	54.68	16.43
Gender	0.52	0.33
US (N=3,657)		
Variable	Mean	Sd
Age	54.73	18.56
Gender	0.52	0.39

Note: the dummy variable gender is 1 for women and Sd represents the standard deviation. Data from LWS.

Adjusted wealth ($w_{adj,i}$) is obtained after extracting the estimated coefficients from equation (11) as follows:

$$w_{adj,i} = w_i - \hat{\beta} F_i - \sum_{n=1}^{4} \hat{\gamma}_n (A_i - 65)^n - \sum_{n=1}^{4} \hat{\delta}_n F_i (A_i - 65)^n. \tag{12}$$

Table 2 presents the summary statistics for the three wealth dependent variables after the adjustment, our main circumstance of analysis (inheritances) and parental education. US households are, on average, the wealthiest, but the high value of the standard deviation highlights large inequality levels of wealth. Indeed, we find the US to be the most unequal country for the three wealth variables under consideration, reaching 94.5, 97.2 and 80.8 Gini points for total, financial and non-financial wealth, respectively. Italy and Spain have Gini coefficients for total and non-financial wealth at around 57-66 points, respectively, while Canada lies between the European and the US economies. In our data, 27.27% of observations in Canada reported to have received some positive inheritance, being this proportion 30.97% in Italy, 39.55% in Spain and 27.18% in the U.S..

Table 2. Summary statistics of the continuous variables (after adjustments).

CANADA (N=3,627)			
Variable	**Mean**	**Gini**	**MLD**
Total	413,732	82.78	2.26
Financial	84,918	92.49	3.53
Non-financial	328,813	73.08	1.86
Inheritances	52,326	89.21	1.26
ITALY (N=5,638)			
Variable	**Mean**	**Gini**	**MLD**
Total	246,357	56.75	0.85
Financial	35,034	80.40	3.37
Non-financial	211,322	56.52	0.89
Inheritances	16,931	93.02	0.97
SPAIN (N=4,792)			
Variable	**Mean**	**Gini**	**MLD**
Total	280,323	65.60	1.41
Financial	47,599	89.18	3.03
Non-financial	232,720	65.79	1.73
Inheritances	33,375	88.55	1.25
US (N=3,657)			
Variable	**Mean**	**Gini**	**MLD**
Total	548,918	94.48	3.85
Financial	203,727	97.17	4.34
Non-financial	345,191	80.74	2.41
Inheritances	81,370	94.80	1.47

Note: All monetary values are expressed in dollars of 2017, after using the LIS PPP adjustment.

Bequests are not orthogonal to other circumstances. In particular, parental education has been shown to be a good proxy of ascendants' social status which reflects other aspects of wealth accumulation, such as human and social capital intergenerational transmission (Nordblom and

Ohlsson, 2010, Adermon et al., 2018; Palomino et al., 2021). To perform the most comprehensive possible analysis and disentangle the potential overlapping effects between both circumstances, inheritances and parental education, we also consider the latter into the analysis. Unfortunately, parental education is only available for the US and Italy. For these two countries, we first use only the inheritances received as a circumstance, so the results can be compared with those obtained for Spain and Canada. Then, we also include the highest parental education level achieved by any parent in five categories: Illiterates, basic studies, basic secondary, upper secondary and university. Table 3 shows that, on average, U.S. citizens have more educated parents than their Italian counterparts. A Shapley value decomposition permits us to estimate the contribution of each circumstance to inequality of opportunity and size the potential overlapping term.[18]

Table 3. Summary statistics of parental education.

	Illiterate	Basic Studies	Basic Secondary	Upper Secondary	University
Italy (N=5,638)	3.42%	17.59%	41.89%	17.64%	19.46%
U.S. (N=3,657)	0.56%	12.66%	43.99%	19.76%	23.03%

Note: Data from LWS.

4. Main results

This section employs the IOp framework to explore how inheritances affect wealth inequality in the US, Canada, Italy and Spain. To begin with, we estimate IOp following the Ferreira and Guignoux (2011) methodology after applying several ad-hoc discretizations over the value of inheritances received. These results are shown to be arbitrary because each partition of the continuous circumstance (inheritances) provides quite different estimates of wealth IOp. This empirical fact confirms the convenience of implementing ML techniques to generate non-arbitrary types.

Table 4 deploys absolute and relative IOp measures with several ad-hoc discretizing points over the value of the inheritances received. First, we split the variable at 0$, generating a type with households who have not inherited anything, and another with those who have inherited any positive bequest. Second, we generate three types: one for those who have not inherited at all and, for those who inherit, we divide the subsample by the median value of the bequests received; third, we generate four types, dividing those who inherit by terciles.[19] Finally, we consider the

[18] The Shapley value decomposition is the only decomposition method that solves the tension between marginality and additivity (See Sastre and Trannoy (2002), Rodríguez, 2004, and Shorrocks, 2013).

[19] Terciles divide the population of households who receive bequests into three groups, each one containing one-third of the total –weighted- population.

case for which it is only relevant to inherit a big amount of wealth. We split the variable in two: those above the 75th percentile of inheritances, and the rest. Many other different discretional cuts have also been checked, only to confirm their sizable differences and arbitrariness.

Table 4. IOp measures with ad-hoc discretizations (Gini).

CANADA	**Absolute IOp**			**Relative IOp**		
Partition of inheritances	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
> 0$	22,22	29,93	19,90	26,84%	32,36%	27,23%
Median	30,05	47,89	25,07	36,30%	51,78%	34,30%
Terciles	39,02	54,91	35,81	47,14%	59,37%	49,00%
P75	33,76	52,01	31,95	40,78%	56,23%	43,72%
ITALY	**Absolute IOp**			**Relative IOp**		
Partition of inheritances	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
> 0$	18,77	23,56	19,75	33,07%	29,30%	35,13%
Median	18,78	29,14	19,36	33,09%	36,24%	34,44%
Terciles	23,52	36,46	24,35	41,44%	45,35%	43,31%
P75	14,68	30,09	15,19	25,87%	37,43%	27,02%
SPAIN	**Absolute IOp**			**Relative IOp**		
Partition of inheritances	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
> 0$	33,61	34,33	38,81	51,23%	38,50%	58,99%
Median	34,08	34,29	40,44	51,95%	38,45%	61,47%
Terciles	42,1	42,92	47,99	64,18%	48,13%	72,94%
P75	28,26	26,29	33,48	43,08%	29,48%	50,89%
US	**Absolute IOp**			**Relative IOp**		
Partition of inheritances	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
> 0$	29,07	45,24	34,25	30,77%	46,56%	42,42%
Median	36,02	63,57	38,41	38,12%	65,42%	47,57%
Terciles	42,8	75,93	44,71	45,30%	78,14%	55,38%
P75	33,65	72,13	31,77	35,62%	74,23%	39,35%

Note: IOp measures calculated with the ex-ante parametric Ferreira and Guignoux (2011) approach, with different ad-hoc discretizations over the continuous circumstance (inheritances). Data from LWS.

Results in Table 4 show that IOp estimates are quite sensible to the researcher's decision on how the different types should be generated. For instance, relative non-financial wealth IOp in Spain range between 50.9% and 72.9% of overall inequality, with remarkable differences prevailing in the other countries and wealth definitions. It is observed that applying subjective ad-hoc criteria leads to arbitrary wealth IOp estimates, hindering accurate cross-country comparisons. For instance, when analyzing absolute non-financial wealth IOp, taking the median cutting point sets Italy as the country with more equality of opportunities (19.4 Gini points), followed by Canada (25.1), the US (38.4) and, finally, Spain (40.4). However, when we set the 75th percentile as the discretizing point, the US and Canada overlap their results at around 32 Gini points, and the absolute difference between Spain and Canada descends from 15 (40.4-25.1) to just 2 (33.5-32) Gini points. Table A1 in the Appendix replicates this table for the MLD index, confirming the arbitrariness of the IOp estimates with ad-hoc cuts in the continuous circumstance.

These results also highlight an intrinsic characteristic of the ex-ante approach. By construction, including more types leads to higher IOp estimates, as it artificially generates more inequality (Ramos and Van de Gaer, 2016). Considering continuous and highly skewed circumstances widens this problem, as every additional cut increases the number of types. This explains why the discretization based on terciles is higher than that based on the median value. The implications are particularly worrying for policy assessment, as these estimates can easily provide downward or upward biased wealth IOp measures directly generated by the arbitrary criteria of the researcher. Indeed, the problem lies, precisely, on finding the point in which the cutting process should stop.

Both limitations call for a more objective method to estimate the impact of circumstances. Overcoming these methodological problems, Table 5 presents and compares wealth IOp estimated with supervised (conditional inference trees and forests) ML techniques.

Table 5. IOp measures with ML techniques (Gini).

CANADA	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree	39.31	54.75	35.98	47.49%	59.20%	49.23%
Forest	39.56	56.68	37.28	47.79%	61.28%	51.01%
ITALY	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree	24.11	40.44	24.76	42.48%	50.30%	44.04%
Forest	24.08	43.03	24.41	42.43%	53.52%	43.43%

SPAIN	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree	43.39	61.03	48.24	66.14%	68.43%	73.32%
Forest	43.47	60.27	47.59	66.27%	67.58%	72.34%
US	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree	58.34	72.64	57.57	61.75%	74.76%	71.30%
Forest	57.25	68.86	57.57	60.59%	70.87%	71.30%

Note: Trees and Forests have endogenously tuned their respective alphas following the K-fold Cross Validation technique explained in the Technical Appendix. Forests are calculated with 500 replications. Data from LWS.

Table 5 shows the IOp measures obtained with the endogenous trees and random forests algorithms.[20] Although the differences between these two methods never reach more than 2 Gini points, we consider the random forests to be our preferred results, as they control for the data dependence of trees. In line with Piketty (2011), Piketty and Zucman (2015), Alvaredo et al. (2017), Adermon et al. (2018) and Nolan et al. (2020), our results point towards inheritances as a relevant vehicle for the intergenerational transmission of wealth disparities. Indeed, our (lower-bound) relative IOp measures explain more than 60% to total wealth inequality in both, the US and Spain, descending the ratios to 47.8% and 42.4% in Canada and Italy, respectively. When looking at financial wealth inequality, the contribution ascends to 61.3% in Canada, Italy 53.5% in Italy, 67.6% in Spain and 70.9% in the US. Finally, IOp explains around 51% of non-financial inequality in Canada, 43.4% in Italy, 72.3% in Spain and 71.3% in the US. Table A2 in the Appendix shows the results for the Gini index with alphas of 0.01 and 0.05, while Table A3 replicates the complete analysis for the MLD index.[21]

Unsurprisingly, financial wealth is highly affected by bequests. The accumulation of financial assets requires either stronger investment skills or a favorable economic support (Lusardi et al., 2017) so, in this manner, inheritances may act as insurances or safety nets who reduce the investors' risk aversion, fostering their participation in the financial markets. For instance,

[20] The tree-plots and precise endogenous alphas tuned with Cross-Validation are available in the Online Appendix. As said in the methods section, the particular partitions and endogenous alphas are sensible to data variations.

[21] As expected, the MLD index provides much lower estimates of wealth IOp. In this respect, Palomino et al. (2021) recently studied the contribution of intergenerational transfers and socioeconomic background to wealth inequality in France, Spain, the UK and the US, ranging their estimates of the gross inheritance contribution measured with the MLD between 32.8% in the UK and 39.3% in France.

Andersen and Nielsen (2011) study the bequeathing spillovers on the stock market participation in Denmark, finding that receiving around one million DKR increases the probability of acquiring listed and unlisted shares by 21%.

Although financial assets are important drivers of the rising wealth inequality levels, housing remains as the main asset in most household portfolios (Jordá et al., 2019). With the exception of Spain, non-financial IOp levels are below financial, reflecting that households may participate with more or less intensity in financial markets, but most of them end up owning their main residence. Nonetheless, the intergenerational transfers still play a major role to explain the different household opportunities to accumulate non-financial wealth. Overall, the effect of intergenerational transfers is stronger where the rising housing prices make homeownership less affordable (Mulder et al., 2015), which is the case of southern European countries. Thus, Poggio (2013) alerts that the weak of public intervention programs and the insecure labor markets have delayed the transition to adulthood in Italy, hence impeding poor young households to purchase houses and making them depend on renting. These findings are extendable to Spain, which also suffered the magnified effects of the housing boom during the 2000s (Anghel et al., 2018).

Similarly, homeownership ratios in Canada have also decreased in the last decades, particularly among young households who, due to life-cycle reasons, are less-likely to have been bequeathed. However, the not-so-high IOp ratios in this country might also be attributed to the government-managed mortgage industry, the tax-sheltered options for down payments for first housings and the scarce public support for rental housing (Maroto and Severson, 2019). After the Great Recession, homeownership rates have also declined substantially in the U.S., being race the main demographic driver of the uneven housing purchases (Goodman and Mayer, 2018).

Overall, the cross-country fiscal and welfare system disparities explained in the introduction do not seem to account for the cross-country IOp differences. For instance, the higher inheritance taxes in Italy might explain why in this country the IOp levels are usually below Spain and the US, although the argument cannot be used for Canada, where no inheritances tax exists. Spain has a strong welfare system similar to Italy, but has wealth IOp levels close to the US, where the private initiative is more preeminent. With all this, the factors explaining cross-country wealth IOp differences require further investigation.[22]

Another important circumstance that is not orthogonal to the inheritances received is parental education. As shown by Palomino et al. (2021), more educated parents have, in general, higher income and saving levels that, in the long term, improve their bequeathing capabilities. In addition, the literature has repeatedly found that household's income is highly correlated to

[22] In this regard, Salas-Rojo and Rodríguez (2021) employ a counterfactual decomposition method to show how the covariates distribution –education, labor status, and income- channel the unequal opportunities to accumulate wealth and explain a remarkable part of wealth IOp disparities between the US and Spain.

parental education of the head (Cabrera et al., 2021), hence improving the possibilities to accumulate wealth. With these ideas in mind, and provided that the LWS data includes information on parental education for Italy and the US, we incorporate this circumstance in our study.

Table 6 presents wealth IOp estimates for Italy and the US using two circumstances: the inheritances and the parental education.[23] Tables A4 and A5 in the Appendix present the rest of IOp indexes confirming the robustness of the proposed ML techniques and with the IOp estimates based on the MLD lying always below those obtained with the Gini coefficient.

Table 6. IOp measures with inheritances and parental education (Gini).

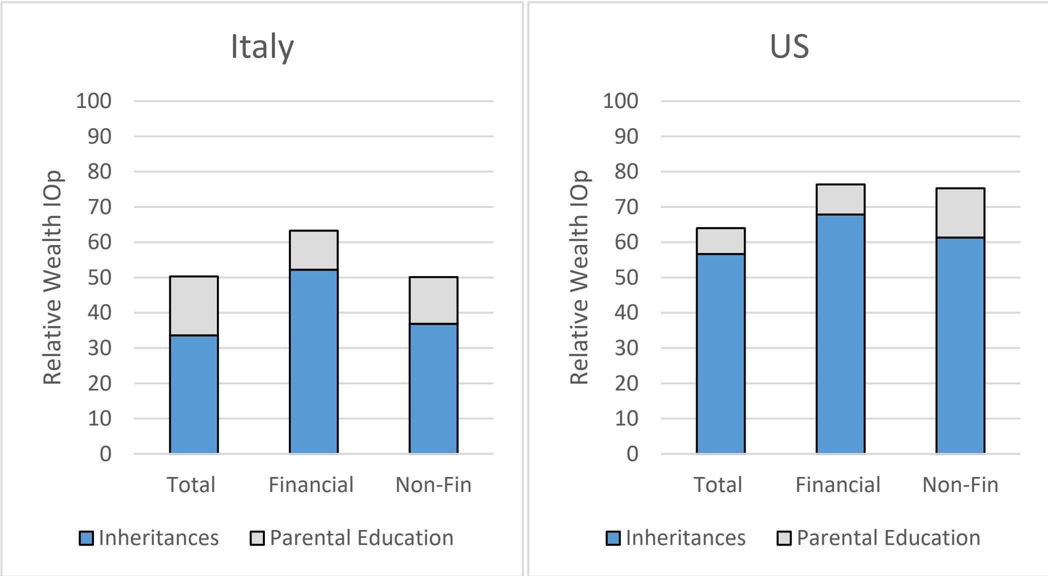
Italy	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree	28.92	50.73	28.43	50.96%	63.10%	50.57%
Forest	28.54	50.87	28.17	50.29%	63.27%	50.11%
US	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree	59.86	73.68	62.43	63.36%	75.83%	74.32%
Forest	60.48	74.21	63.22	64.01%	76.37%	75.30%

Note: Trees and Forests have endogenously tuned their respective alphas following the K-fold Cross Validation technique explained in the Technical Appendix. Forests are calculated with 500 replications. Data from LWS.

In Italy, including parental education as circumstance increases absolute and relative IOp measures. Using the Gini index, 50.3% of total inequality is explained by our IOp measures with both circumstances, reaching up to 63.3% in financial and 50.1% in non-financial wealth inequality. Similarly, US estimates ascend to 64% in total, 76.4% in financial and 75.3% in non-financial wealth. These results show that parental education conducts the unequal opportunities through a different channel than inheritances despite that may exist some overlapping between both circumstances since the inheritances received are, potentially, a comprehensive variable that collects many different aspects of individual’s background. Considering this possibility, Figure 1 plots the relative contribution of each circumstance to IOp, calculated with the Shapley value decomposition over the random forest indexes calculated in Table 6.

[23] Wealth IOp estimates with the parental education alone are shown in Table A7.

Figure 1. Shapley Value Decomposition of wealth IOp.



Note: Total stands for total net wealth, Financial for financial net wealth, and Non-Fin for non-financial net wealth. Shapley value decomposition applied over the relative wealth IOp measures obtained in Table 6. Data from LWS.

In Italy, around two-thirds of total wealth IOp can be explained by the distribution of bequests, being the remaining attributed to parental education. The inheritances received can explain up to four-fifths (82.5%) of financial and three-quarters (76.5%) of non-financial wealth IOp, being the complementary shares up to 100% attributed to parental education. In the US, the relative importance of inheritances is higher, always explaining more than 80% of wealth IOp. These results highlight the importance of inheritances on defining the opportunities to household wealth accumulation, although they do not neglect the importance of other factors, such as the human capital transmission, that may overlap part of their effects.

So far, our analysis has been focused on analyzing the aggregate household opportunities to accumulate wealth. Aimed to be more precise, in the next section we study whether the effect of inheritances and parental education is homogeneous across the wealth distribution or, rather, depends on the part of the wealth distribution under consideration.

5. Circumstances and Wealth Distribution

The distributions of wealth and inheritances are highly skewed, so it is pertinent to check whether their relationship is homogeneous along the wealth distribution or, rather, more accentuated at the tails or the middle (see Rodríguez, 2008). For this task, we expand our analysis by using the Single-Parameter Gini (S-Gini) proposed in Donaldson and Weymark (1980) and Yitzhaki (1983).

S-Gini indexes are born from the idea that the canonical Gini index weights equally all parts of the wealth distribution. Thus, they introduce a vector of weights, defined by a parameter v of inequality aversion which modifies the weight given to different percentile positions, q. Thus, as v increases, the delivered weights give more (less) importance to the lower (upper) part of the distribution. Formally, the S-Gini index is:

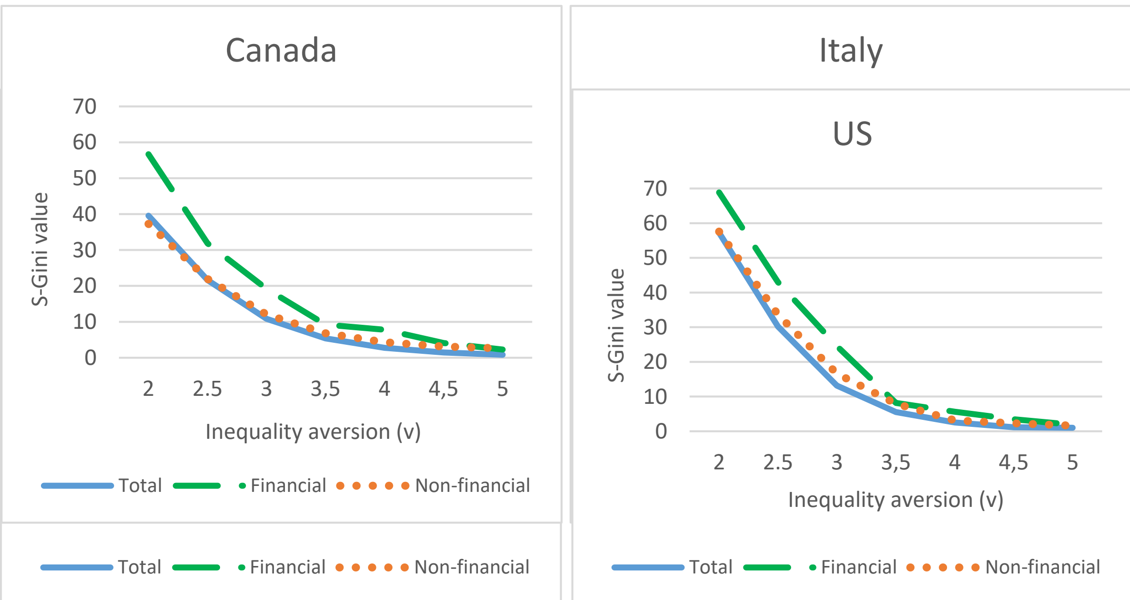
$$I_{S-Gini}(F;v) = 1 - v[v-1]\int_0^1 [1-q]^{v-2} L(F;q)dq \tag{13}$$

where L is the Lorenz curve, and $v > 1$ is an inequality aversion parameter. For $v = 2$, equation (13) provides the traditional Gini index.

Now, by applying this index to the counterfactual smoothed distribution $\hat{w}$, we obtain an absolute measure of IOp for an inequality aversion parameter v, $IOp(v)$. We might find three possible cases when the value of v is increased. First, if there were no significant changes in the estimates, the circumstances under consideration would have a homogeneous effect along the wealth distribution. Second, if absolute IOp increased with the inequality aversion parameter v, we could safely say that the effect of circumstances is higher at the lower tail of the wealth distribution. Finally, if absolute IOp decreased when the inequality aversion parameter rose, the effect of circumstances would be more intense at the upper tail of the wealth distribution.

Before continuing, we remark that we do not include relative IOp in these results, because its interpretation can be misleading: changes provoked in these relative measures could not only be caused by absolute IOp, but also by overall inequality. As a result, the calculated variation in the relative IOp measures would not exclusively be explained by the heterogeneous effect of the circumstances at different parts of the wealth distribution but, rather, by their interaction with overall inequality.

Figure 2. S-Gini wealth IOp with inheritances.



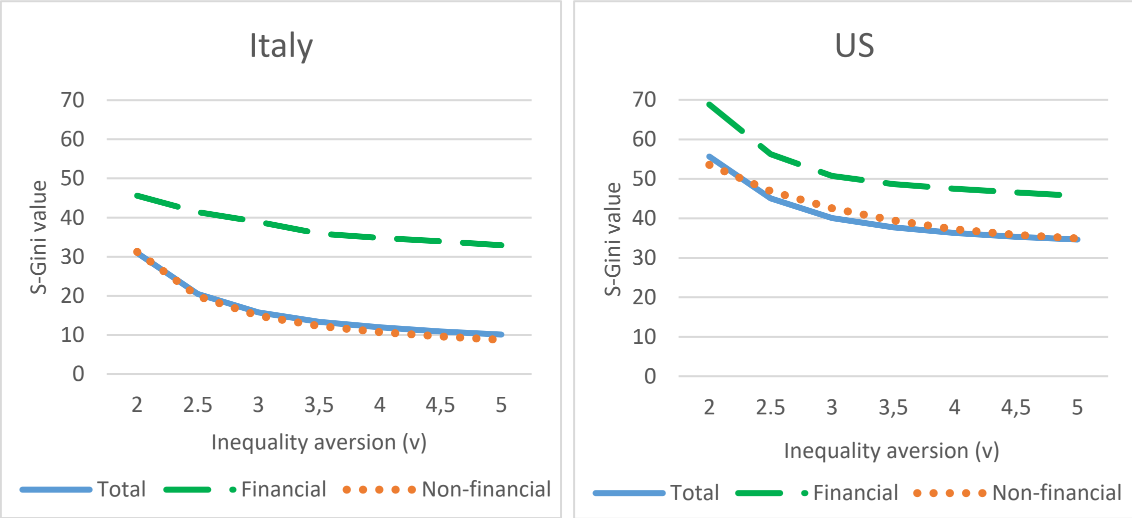
Note: Absolute IOp measures for different inequality aversion parameters calculated with random forests and endogenously tuned alphas. Data from LWS.

We estimate the effect of circumstances along the wealth distribution for $v \in [2, 5]$. The precise values are included in Tables A6 and A7 in the Appendix, while Figure 2 plots the evolution of $IOp(v)$ for the inheritances received. It is clearly observed that the higher the weight on the lower tail of the distribution, the smaller the value of $IOp(v)$, with this index being close to zero for $v = 5$ in all wealth measures for Canada, Italy and the US. In these countries, the inheritances received lose importance when we focus at the lower tail of the wealth distribution, i.e., the significant contribution to wealth inequality shown in Tables 5 and 6 is mainly explained by their effect on the opportunities of the middle class and, particularly, wealthy households. However, for Spain, the $IOp(v)$ values are reduced at a smoother rate. As explained in OECD (2021), the inheritances received in this country still play a role in the wealth of poor households, although we find them to be much more important to define the opportunities of the rich.

The same analysis is replicated in Figure 3 for inheritances and parental education in Italy and the US. In the first country, $IOp(v)$declines again, although at a lower rate than in Figure 2. Thus, $IOp(v)$ in Italy descends from 28.5 to 9.5 in total wealth, from 50.9 to 24.9 in financial wealth and from 28.2 to 8.1 in non-financial wealth. Similarly, $IOp(v)$ smoothly descends in the US for total wealth (from 60.5 to 35.3), financial wealth (from 74.2 to 41.7), and non-financial wealth (from 63.2 to 33.8). This result suggests that for those located at the lower part of the wealth distribution parental education gains importance to determine the household opportunities to accumulate wealth, particularly in the US.[24]

[24] In line with Lusardi et al. (2017), parental education in the US seems to be a relevant factor explaining financial wealth inequality, particularly at the bottom of the wealth distribution. A successful participation in the financial markets usually requires some skills highly affected by human capital. See also Cabrera et al. (2021).

Figure 3. S-Gini wealth IOp with inheritances and parental education.



Note: Absolute IOp measures for different inequality aversion parameters calculated with random forests and endogenously tuned alphas. Data from LWS.

By employing S-Gini indexes we highlight the heterogeneous transmission of opportunities at different parts of the distribution, which otherwise are hidden by other traditional approaches. On the one hand, as inheritances are typically concentrated at the upper tail of the wealth distribution (Nolan et al, 2020; OECD, 2021), diminishing the weight on this part of the distribution makes the effect of inheritances on inequality to disappear. This fact together with the large non-linearities of both, wealth and inheritances, might explain why inheritances are so relevant to explain wealth inequality, as they would intensively foster the opportunities of the wealthy. On the other hand, as inheritances at the lower tail of the wealth distribution are rather small, the opportunities of those at this part of the distribution are conditioned by other factors such as parental education.

6. Conclusions

In this paper, we estimate wealth IOp associated with inheritances received for four developed western economies: Canada, Italy, Spain and the US. We estimate this by the between-types component of total, financial and non-financial wealth inequality. After controlling by household size, age and sex, and computing the type distributions based on the inheritances received, there should not exist any dispersion between types. Otherwise, overall wealth inequality would be conditioned by the bequest distribution.

The traditional definition of types in the literature (Ferreira and Guignoux, 2011) for a continuous circumstance like inheritances requires a discretization process that, when left to the researchers' criterion, leads to arbitrary results. We show here that the choice of bequests cut-off points to define the types greatly influences the contribution of bequests to overall inequality, which can prevent a correct cross-country comparison or policy advising. To overcome this limitation, we apply Machine Learning techniques that objectivize the discretization process, statistically justifying the cuts on the inheritances variable. Among all available methods, we select the *endogenously tuned random forests* as the approach that provides non-discretionary estimates of the impact of inheritances on wealth inequality.

Our results show that a remarkable share of overall wealth inequality is explained by our estimates of IOp when using bequests as the only circumstance. Particularly, this measure explains up to 54% of financial wealth inequality in Italy, 61% in Canada, 68% in Spain and 71% in the US. In the case of non-financial wealth, inheritances explain 51% overall inequality in Canada, 43% in Italy, 71% in the US and a significant 72% in Spain. On aggregate, almost 48% (Canada), 42% (Italy), 66% (Spain) and 61% (the US) of total net wealth inequality can be assigned to wealth IOp.

When we include parental education as an additional circumstance in the countries that have that variable available in the dataset, IOp barely increases, being 63% in financial wealth, 50% in non-financial and total wealth in Italy. For the US, it explains 64% of total wealth, the ratios being 76% and 75% for financial and non-financial wealth, respectively. The literature has already shown that inheritances and parental education are not orthogonal (Adermon, 2018; Palomino et al., 2021), so we check for the existence of overlapping. For this task, we perform a Shapley value decomposition and find that, in the US, more than five-sixths of wealth IOp can be attributed to inheritances, while the rest is explained by the parental education. In Italy, around two-thirds of wealth IOp is attributed to inheritances, being the remainder explained by the parental education.

Finally, due to the high skewness of the wealth and inheritance distributions, we apply the S-Gini index for several parameters of inequality aversion: the higher the parameter values, the larger the weight on the lower tail of the wealth distribution for the index calculation. We find that, the more we focus on the left tail of the wealth distributions, the less important inheritances received are to explain their opportunities to accumulate wealth, while the opposite happens with parental education. That is, inheritances and parental education seem to have a heterogeneous effect along the wealth distribution, and influence different tails of the distribution. Therefore, policies having a bearing on inheritances and on parental education are complementary (not substitutive) for the reduction of wealth inequality.

Conflict of interests

The authors deny the existence of any conflict of interest.

Data:

We do not generate new data in this paper. The data we use is available at the Luxembourg Income Survey repository: https://www.lisdatacenter.org/data-access/

References

Adermon, A., Lindahl, M. and Waldenström, D.: Intergenerational wealth mobility and the role of inheritance: evidence from multiple generations. The Eco Jou. 128, 482 -513. (2018)

Alstadsaeter, A., Johannesen and N., Zucman, G.: Tax evasion and inequality. NBER Working Paper, no 23772. (2017)

Alvaredo, F., Chancel, L., Piketty, T., Saez, E. and Zucman., G.: Global inequality dynamics: New findings from WID.world. Ame. Eco. Rev. 107, 404-409. (2017)

Andersen, S. and Nielsen, K. M.: Participation constraints in the stock market: evidence from unexpected inheritance due to sudden death. The Rev. of Fin. Stu, 24, 1667-1697. (2011)

Anghel, B., Basso, H. Bover, O., Casado, J. M., Hospido, L., Izquierdo, M., Kataryniuk, I. A., Lacuesta, A., Montero, J. M. and Vozmediano, E.: Income, consumption and wealth inequality in Spain. Ser. 9, 351-387. (2018)

Atkinson, A.: On the Measurement of Inequality. Jou. of Eco. The., 2, 244-263. (1970)

Banerjee, D., Ghosal, A. and Mukherjee, I.: Prediction of gold price movement using discretization procedure. Comp. Int. in Dat. Min. 1, 345-356. (2019)

Bellégo, C., Benatia, D. and Pape, L.: Dealing with the log of zero in regression models. SSRN Ele. Jou. https://doi.org/10.2139/ssrn.3444996. (2021)

Boserup, S. H., Kopczuk, W. and Kreiner, C. T.: The role of bequests in shaping wealth inequality: Evidence from Danish wealth records. Ame. Eco. Rev. 106, 656-61. (2016)

Bourguignon, F., Ferreira, F. H. and Menéndez, M.: Inequality of opportunity in Brazil. Rev. of Inc. and Wea. 53, 585-618. (2007)

Bradbury, K. and Triest, R. K.: Inequality of opportunity and aggregate economic performance. The Rus. Sag. Fou. Jou. of the Soc. Sci. 2, 178-201. (2016)

Brunori, P., Hufe, P. and Mahler, D. G.: The roots of inequality: Estimating inequality of opportunity from regression trees. World Bank Group Policy Research Working Paper no 8349. (2020)

Brunori, P. and Neidhöfer, G.: The evolution of inequality of opportunity in Germany: A machine learning approach. ZEW Discussion Papers, no 20-013. (2020)

Brunori, P., Peragine, V. and Serlenga, L.: Upward and downward bias when measuring inequality of opportunity. Soc. Cho. and Wel. 52, 635-661. (2019)

Buhmann, B., Rainwater, L., Schmaus, G. and Smeeding, T. M.: Equivalence scales, well-being, inequality and poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study Database. Rev. of Inc. and Wea. 34, 115-142. (1988)

Cabrera, L., Marrero, G. A., Rodríguez, J. G. and Salas-Rojo, P.: Inequality of opportunity in Spain: New insights from new data. Rev. of Pub. Eco. 237, 153-185. (2021)

Carranza, R.: Inequality of outcomes, inequality of opportunity, and economic growth. ECINEQ Working Paper no 534. (2020)

Checchi, D. and Peragine, V.: Inequality of opportunity in Italy. The Jou. of Eco. Ine. 8, 429-450. (2010)

Crawford, R. and Hood, A.: Lifetime receipt of inheritances and the distribution of wealth in England. Fis. Stu. 37, 55-75. (2016)

Cowell, F. and Van Kerm, P.: Wealth inequality: A survey. Jou. of Eco. Sur. 29, 671-710. (2015)

De Nardi, M.: Quantitative models of wealth inequality: A survey. NBER Working Paper no 21106. (2015)

De Nardi, M. and Yang, F.: Wealth inequality, family background and estate taxation. Jou. of Mon. Eco. 77, 130-145. (2016)

Donaldson, D. and Weymark, J. A.: A single parameter generalization of the Gini indices of inequality. Jou. of Eco. The. 22, 67-86. (1980)

Elinder, M., Erixson, O. and Waldenström, D.: Inheritance and wealth inequality: Evidence from population registers. Jou. of Pub. Eco. 165, 17-30. (2018)

Ferreira, F. H. G. and Guignoux, J.: The measurement of inequality of opportunity: Theory and application to Latin America. Rev. of Inc. and Wea. 57, 622-657. (2011)

Fessler, P. and Schurz, M.: Private wealth across European countries: The role of income, inheritance and the welfare state. Jou. of Hum. Dev. and Cap. 19, 1-29. (2018)

Fushiki, T.: Estimation of prediction error by using K-Fold Cross Validation. Sta. and Com. 21, 137-146. (2011)

Friedman, F., Hastie, T. and Tibshirani, R.: The elements of Statistical Learning. Springer, Berlin. (2009)

Foster, J. E. and Shneyorov, A.: Path independent inequality measures. Jou. of Eco. The. 91, 199-222. (2000)

Gallo, C. A., Cecchirini, R. L., Carballido, J. A., Micheletto, S. and Ponzoni, I.: Discretization of gene expression data revised. Bri. Bio. 17, 758-770. (2016)

Genz, A.: Numerical computation of multivariate normal probabilities. Jou. of Com. and Gra. Sta. 1, 141-149. (1992)

Goodman, L. S. and Mayer, C.: Homeownership and the American dream. Jou. of Eco. Per. 32, 31-58. (2018).

Hothorn, T., Hornik, K. and Zeileis, A.: Unbiased recursive partitioning: A conditional inference framework. Jou. of Com. and Gra. Sta. 15, 651-674. (2006)

Jordá, O., Knoll, K., Kuvshinov, D., Schularick and M., Taylor, A. M.: The rate of return on everything, 1870-2015. The Qua. Jou. of Eco. 134, 1225-1298. (2019)

Karagiannaki, E.: The impact of inheritance on the distribution of wealth: Evidence from Great Britain. Rev. of Inc. and Wea. 63, 394-408. (2017)

Lalithendra, N. V. and Prasad, G. S.: Support vector machine in the anticipation of currency markets. Int. Jou. of Eng. and Tec. (2018)

Lefranc, A., Pistolesi, N. and Trannoy A.: Inequality of opportunities vs. Inequality of outcomes: Are Western societies all alike? Rev. of Inc. and Wea. 54, 513 -546. (2008)

Li Donni, P., Rodríguez, J. G. and Rosa Días, P.: Empirical definition of social types in the analysis of inequality of opportunity: A latent classes approach. Soc. Cho. and Wel. 44, 673-701. (2015)

Listokin, D., Wyly, E. K., Schmitt, B. and Voicu, I.: The potential and limitations of mortgage innovation in fostering homeownership in the United States. Hou. Pol. Deb. 12, 465-513. (2001)

Lusardi, A., Michaud, P. C. and Mitchell, O. S.: Optimal financial knowledge and wealth inequality. Jou. of Pol. Eco. 125, 431-477. (2017)

Lutsgarten, J. L., Visweswaran, S. and Gopalakrishnan, V.: Application of an efficient bayesian discretization method to biomedical data. Bio. 12, 1-15. (2011)

Luxembourg Wealth Study (LWS) Database, (Italy, Canada, U.S., Spain; access 03/03/2020 – 30/05/2021). Luxembourg: LIS.

Maroto, M. and Severson, M.: Owning, renting or living with parents? Changing housing situations among Canadian young adults, 2001 to 2011. Hou. Stu. Doi: https://doi.org/10.1080/02673037.2019.1630559. (2019).

Marrero, G. A. and Rodríguez, J. G.: Inequality of Opportunity and Growth. Jou. of Dev. Eco. 104, 107-122. (2013)

Marrero, G. A. and Rodríguez, J. G.: Inequality and growth: The cholesterol hypothesis. ECINEQ Working Paper no 501. (2019)

Marrero, G. A., Rodríguez, J. G. and van der Weide, R.: Race and gender inequality in the U.S.: Effects of the 'double gap' on growth along the income ladder". Mimeo (2021)

Mulder, C. H., Dewilde, C., Duijn, M. and Smits, A.: The association between parents' and adult children's homeownership: a comparative analysis. Eur. Jou. Pop. 31, 495-527. (2015)

Nolan, B., Palomino, J.C., Van Kerm, P. V. and Morelli, S.: The Intergenerational Transmission of Wealth in Rich Countries. Centre for Studies in Economics and Finances Working Papers, no. 578. (2020)

Nordblom K. and Ohlsson, H.: Bequests, gifts and education: Links between intergenerational transfers. Emp. Eco. 40, 343-358. (2010)

OECD.: Pensions at a Glance 2019: OECD and G20 Indicators. OECD Publishing, Paris. Doi: https://doi.org/10.1787/b6d3dcfc-en. (2019)

OECD.: Inheritance Taxation in OECD Countries. OECD Tax Policy Studies, OECD Publishing, Paris. https://doi.org/10.1787/e2879a7d-en. (2021)

Palomino, J. G., Marrero, G. A., Nolan, B. and Rodríguez, J.G.: Inheritances, family background and wealth inequality. Oxf. Eco. Pap., https://doi.org/10.1093/oep/gpab052 (2021)

Piketty, T.: On the long-run evolution of inheritance: France 1820-2050. The Qua. Jou. of Eco. 126, 1041-1131. (2011)

Piketty, T. and Zucman, G.: Wealth and inheritance in the long run. Han. of Inc. Dis. 2, 1303-1368. (2015)

Poggio, T.: The first steps into the Italian housing system. Inequality between generational gaps and family intergenerational transfers. In R. Forrest and N. M. Yip (Eds.), Young people and housing transitions, trajectories and generational fractures. London: Rougledge. (2013)

Ramos, X. and Van de Gaer, D.: Approaches to inequality of opportunity: Principles, measures and evidence. Jou. of Eco. Sur. 30, 855-883. (2016)

Ravallion, M.: A concave log-like transformation allowing non-positive values. Eco. Let. 161, 130-132. (2017).

Rawls, J.: A theory of justice. Harvard University Press, Cambridge. (1971)

Rodríguez, J. G.: Descomposición Factorial de la Desigualdad de la Renta. Rev. de Eco. Apl. 36, 25-46. (2004)

Rodríguez, J. G.: Partial equality of opportunity orderings. Soc. Cho. and Wel. 31, 435-456. (2008)

Rodríguez, J. D., Pérez, A. and Lozano, J. A.: Sensitivity analysis of K-Fold Cross Validation in prediction error estimation. Tran. on Pat. Ana. and Mac. Int. 32, 569-575. (2009)

Roemer, J. E.: A Pragmatic approach to responsibility for the egalitarian planner. Phi. and Pub. Aff. 20, 146-166. (1993)

Salas-Rojo, P. and Rodríguez, J. G.: The distribution of wealth in Spain and the U.S.: The role of socioeconomic factors. Ser. 12, 389–421. (2021)

Sastre, M. and Trannoy, A.: Shapley Inequality Decomposition by Factor Components: Some Methodological Issues. Jou. of Eco. 77, 51-89. (2002)

Schlosser, L., Hothorn, T., Stauffer and R., Zeileis, A.: Distributional regression forests for probabilistic precipitation forecasting in complex terrain. The Ann. of App. Sta. 13, 1564-1589. (2019)

Schneebaum, A., Rehm. M., Mader and K., Hollan, K.: The gender wealth gap across European countries. Rev. of Inc. and Wea. 64, 295-331. (2018)

Sen, A.: Equality of what? The Tan. Lec. on Hum. Val., University of Utah Press, Salt Lake City. (1980)

Shorrocks, A. F.: Decomposition procedures for distributional analysis: A unified framework based on the Shapley value. The Jou. of Eco. Ine. 11, 99-126. (2013)

Sierminska, E. M., Frick, J. R. and Grabka, M., M.: Examining the gender wealth gap. Oxf. Eco. Pap. 62, 669-690. (2010)

Solon, G.: Intergenerational income mobility in the United States. Ame. Eco. Rev. 82, 393-408. (1992)

Strobl, C., Boulesteix, A. L., Zeileis, A. and Hothorn, T.: Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. Bio. 8, 25-46. (2007)

Strobl C. Malley, J. and Tutz, G.: An Introduction to Recursive Partitioning: Rationale, Application and Characteristics of Classification and Regression Trees, Bagging and Random Forests. Psy. Met. 14, 323-348. (2009)

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

Wolff, E. N. and Gittleman, M.: Inheritances and the Distribution of Wealth or Whatever Happened to the Great Inheritance Boom? The Jou. of Eco. Ine. 12, 439-468. (2014)

Wolff, E. N.: U.S. Pensions in the 2000s: The lost decade. Rev. of Inc. and Wea. 61, 599-629. (2015)

Wolff, F. C. and Laferrere, A.: Microeconomic models of family transfers. Han. On the Eco. Of Giv, Rec, and Alt. In S. Kolm and M. Ythier (ed.), 1, 13, 889-969. Elsevier. (2006)

Yitzhaki, S.: On an Extension of the Gini Inequality Index. Int. Eco. Rev. 24, 617-628. (1983)

Zucman, G.: The Triumph of Injustice: How the Rich Dodge Taxes and How to Make Them Pay. Norton and Company, New York. (2019)

Appendix

Table A1: IOp with ad-hoc discretizations (MLD).

CANADA	**Absolute IOp**			**Relative IOp**		
Partition of inheritances	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
> 0$	0.11	0.18	0.07	4.87%	5.10%	3.76%
Median	0.19	0.47	0.14	8.41%	13.31%	7.53%
Terciles	0.29	0.59	0.25	12.83%	16.71%	13.44%
P75	0.27	0.59	0.25	11.95%	16.71%	13.44%
ITALY	**Absolute IOp**			**Relative IOp**		
Partition of inheritances	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
> 0$	0.08	0.12	0.09	9.41%	3.56%	10.11%
Median	0.11	0.21	0.11	12.94%	6.23%	12.36%
Terciles	0.12	0.29	0.13	14.12%	8.61%	14.61%
P75	0.09	0.27	0.10	10.59%	8.01%	11.24%
SPAIN	**Absolute IOp**			**Relative IOp**		
Partition of inheritances	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
> 0$	0.26	0.27	0.36	18.44%	8.91%	20.81%
Median	0.24	0.25	0.34	17.02%	8.25%	19.65%
Terciles	0.32	0.35	0.44	22.70%	11.55%	25.43%
P75	0.20	0.17	0.27	14.18%	5.61%	15.61%
US	**Absolute IOp**			**Relative IOp**		
Partition of inheritances	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
> 0$	0.17	0.44	0.24	4.42%	10.14%	9.96%
Median	0.29	0.91	0.33	7.53%	20.97%	13.69%
Terciles	0.36	1.34	0.38	9.35%	30.88%	15.77%
P75	0.30	1.22	0.28	7.79%	28.11%	11.62%

Note: IOp measures calculated with the ex-ante parametric Ferreira and Guignoux (2011) approach, with different ad-hoc discretizations over the continuous circumstance (inheritances). Data from LWS.

Table A2: IOp with Trees and Forests, alphas at 1% and 5% (Gini).

CANADA	Absolute IOp			Relative IOp		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree 1%	39.40	54.65	36.23	47.60%	59.09%	49.58%
Tree 5%	38.79	54.75	36.01	46.86%	59.20%	49.27%
Forest 1%	39.36	56.89	37.24	47.55%	61.51%	50.96%
Forest 5%	38.81	56.75	37.03	46.88%	61.36%	50.67%
ITALY	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree 1%	24.16	40.43	24.83	42.57%	50.29%	44.17%
Tree 5%	22.63	40.44	25.64	39.88%	50.30%	45.61%
Forest 1%	24.32	43.43	24.38	42.85%	54.02%	43.37%
Forest 5%	24.30	42.99	24.69	42.82%	53.47%	43.92%
SPAIN	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree 1%	43.69	61.05	48.24	66.60%	68.46%	73.32%
Tree 5%	43.65	60.98	48.24	66.54%	68.38%	73.32%
Forest 1%	43.62	60.32	47.60	66.49%	67.64%	72.35%
Forest 5%	43.12	60.17	45.63	65.73%	67.47%	69.36%
US	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree 1%	58.64	72.86	57.86	62.07%	74.98%	71.66%
Tree 5%	58.40	72.96	57.57	61.81%	75.08%	71.30%
Forest 1%	57.30	67.86	57.86	60.65%	69.84%	71.66%
Forest 5%	56.99	68.75	57.57	60.32%	70.75%	71.30%

Note: IOp measures using trees and random forests. Alphas are exogenously set to 1% and 5%. Forests are calculated with 500 replications. Data from LWS.

Table A3: IOp with Trees and Forests, alphas endogenous at 1% and 5% (MLD).

CANADA	Absolute IOp			Relative IOp		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree 1%	0.28	1.14	0.23	12.39%	32.29%	12.37%
Tree 5%	0.29	1.16	0.24	12.83%	32.86%	12.90%
Tree end	0.29	1.16	0.24	12.83%	32.86%	12.90%
Forest 1%	0.28	1.15	0.24	12.39%	32.58%	12.90%
Forest 5%	0.29	1.20	0.24	12.83%	33.99%	12.90%
Forest end.	0.28	1.21	0.25	12.39%	34.28%	13.44%
ITALY	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree 1%	0.12	0.32	0.25	14.12%	9.50%	28.09%
Tree 5%	0.13	0.34	0.26	15.29%	10.09%	29.21%
Tree end	0.13	0.33	0.24	15.29%	9.79%	26.97%
Forest 1%	0.13	0.36	0.15	15.29%	10.68%	16.85%
Forest 5%	0.13	0.36	0.16	15.29%	10.68%	17.98%
Forest end.	0.13	0.36	0.14	15.29%	10.68%	15.73%
SPAIN	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree 1%	0.33	0.68	0.40	23.40%	22.44%	23.12%
Tree 5%	0.32	0.69	0.41	22.70%	22.77%	23.70%
Tree end	0.33	0.68	0.40	23.40%	22.44%	23.12%
Forest 1%	0.32	0.65	0.40	22.70%	21.45%	23.12%
Forest 5%	0.32	0.66	0.40	22.70%	21.78%	23.12%
Forest end.	0.32	0.65	0.39	22.70%	21.45%	22.54%
US	**Absolute IOp**			**Relative IOp**		
ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Tree 1%	0.71	2.35	0.66	18.44%	54.15%	27.39%
Tree 5%	0.72	2.40	0.65	18.70%	55.30%	26.97%
Tree end	0.72	2.34	0.65	18.70%	53.92%	26.97%
Forest 1%	0.66	2.23	0.64	17.14%	51.38%	26.56%
Forest 5%	0.65	2.24	0.63	16.88%	51.61%	26.14%
Forest end.	0.65	2.09	0.63	16.88%	48.16%	26.14%

Note: IOp measures calculated with several Machine Learning techniques. Trees and Forests have endogenously tuned their respective alphas following the K-fold Cross Validation technique explained in the Technical Appendix. Other alphas are exogenously set to 1% and 5%. Forests are calculated with 500 replications. Data from LWS.

Table A4: IOp estimates with the inheritances and parental education (Italy).

Italy	Absolute IOp			Relative IOp		
Partition and ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Gini						
> 0$	31.38	55.99	30.57	55.30%	69.64%	54.38%
Median	31.98	57.32	30.99	56.35%	71.29%	55.12%
Terciles	35.29	60.52	34.54	62.19%	75.27%	61.44%
P75	30.99	61.11	29.56	54.61%	76.01%	52.58%
Tree 1%	28.92	50.86	28.43	50.96%	63.26%	50.57%
Tree 5%	28.76	50.81	28.43	50.68%	63.20%	50.57%
Forest 1%	28.52	50.86	28.15	50.26%	63.26%	50.07%
Forest 5%	28.52	50.87	28.17	50.26%	63.27%	50.11%
MLD						
> 0$	0.16	0.56	0.15	18.82%	16.62%	16.85%
Median	0.18	0.59	0.17	21.18%	17.51%	19.10%
Terciles	0.21	0.65	0.21	24.71%	19.29%	23.60%
P75	0.18	0.68	0.17	21.18%	20.18%	19.10%
Tree 1%	0.15	0.49	0.16	17.65%	14.54%	17.98%
Tree 5%	0.15	0.49	0.16	17.65%	14.54%	17.98%
Tree end	0.15	0.49	0.16	17.65%	14.54%	17.98%
Forest 1%	0.15	0.49	0.16	17.65%	14.54%	17.98%
Forest 5%	0.14	0.49	0.15	16.47%	14.54%	16.85%
Forest end	0.15	0.49	0.15	17.65%	14.54%	16.85%

Note: IOp measures calculated with several Machine Learning techniques. Trees and Forests have endogenously tuned their respective alphas following the K-fold Cross Validation technique explained in the Technical Appendix. Other alphas are exogenously set to 1% and 5%. Forests are calculated with 500 replications. Data from LWS.

Table A5: IOp estimates with the inheritances and parental education (the US).

US	Absolute IOp			Relative IOp		
Partition and ML method	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
Gini						
> 0$	45.78	59.16	49.94	48.45%	60.88%	61.85%
Median	51.05	65.41	53.9	54.03%	67.32%	66.76%
Terciles	53.05	68.67	55.79	56.15%	70.67%	69.10%
P75	50.01	66.49	50.29	52.93%	68.43%	62.29%
Tree 1%	59.42	73.94	62.44	62.89%	76.09%	77.33%
Tree 5%	59.68	73.22	61.95	63.17%	75.35%	76.73%
Forest 1%	60.45	73.99	63.34	63.98%	76.14%	75.45%
Forest 5%	60.94	74.17	64.16	64.50%	76.33%	75.46%
MLD						
> 0$	0.36	0.68	0.44	9.35%	15.67%	18.26%
Median	0.45	0.84	0.5	11.69%	19.35%	20.75%
Terciles	0.48	0.92	0.54	12.47%	21.20%	22.41%
P75	0.43	0.86	0.44	11.17%	19.82%	18.26%
Tree 1%	0.88	1.56	0.85	22.86%	35.94%	35.27%
Tree 5%	0.88	1.54	0.85	22.86%	35.48%	35.27%
Tree end	0.87	1.55	0.85	22.60%	35.71%	35.27%
Forest 1%	0.84	0.76	0.62	21.82%	17.51%	25.73%
Forest 5%	0.83	0.75	0.63	21.56%	17.28%	26.14%
Forest end.	0.83	0.76	0.63	21.56%	17.51%	26.14%

Note: IOp measures calculated with several Machine Learning techniques. Trees and Forests have endogenously tuned their respective alphas following the K-fold Cross Validation technique explained in the Technical Appendix. Other alphas are exogenously set to 1% and 5%. Forests are calculated with 500 replications. Data from LWS.

Table A6: Single Parameter Gini Absolute IOp with inheritances.

Canada	Total Wealth	Financial Wealth	Non-Financial Wealth
v=2	39.56	56.68	37.28
v=2.5	21.64	31.84	21.96
v=3	10.84	19.16	12.03
v=3.5	5.39	9.21	6.82
v=4	2.75	7.78	4.31
v=4.5	1.47	4.08	3.13
v=5	0.83	2.28	2.61
Italy	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
v=2	24.08	43.03	24.41
v=2.5	9.82	21.99	9.88
v=3	4.15	9.93	4.17
v=3.5	1.74	4.29	1.76
v=4	0.77	1.86	0.75
v=4.5	0.34	0.87	0.37
v=5	0.13	0.49	0.11
Spain	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
v=2	43.47	60.27	47.59
v=2.5	33.37	49.04	38.48
v=3	24.71	37.19	29.86
v=3.5	17.82	26.81	22.26
v=4	12.61	18.77	16.17
v=4.5	8.82	13.43	11.53
v=5	6.13	9.05	8.11
US	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
v=2	57.25	68.86	57.57
v=2.5	30.18	42.97	33.68
v=3	13.15	24.56	16.64
v=3.5	5.54	8.17	7.95
v=4	2.56	5.65	3.12
v=4.5	1.14	3.47	2.33
v=5	1.01	1.89	1.62

Note: IOp measures calculated with random forests with 500 replications and an alpha endogenously tuned following the K-fold Cross Validation technique explained in the Technical Appendix. Data from LWS.

Table A7: Single Parameter Gini Absolute IOp with inheritances and parental education.

Italy	Parental Education			Parental Education and Inheritances		
	Total Wealth	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
v=2	21.65	41.65	17.65	28.54	50.87	28.17
v=2.5	15.17	35.14	13.52	16.62	39.65	15.56
v=3	12.92	33.79	11.25	11.78	31.35	10.53
v=3.5	12.11	34.32	10.37	9.99	27.02	8.65
v=4	11.85	35.95	9.37	9.63	25.63	8.16
v=4.5	11.14	35.26	8.68	9.53	24.88	8.13
v=5	10.55	32.12	8.82	9.46	24.87	8.08
US	**Parental Education**			**Parental Education and Inheritances**		
	Total Wealth	**Financial Wealth**	**Non-Financial Wealth**	**Total Wealth**	**Financial Wealth**	**Non-Financial Wealth**
v=2	52.34	66.82	53.21	60.48	74.21	63.22
v=2.5	49.61	61.72	47.32	51.78	63.83	51.73
v=3	46.83	59.11	45.99	42.99	62.31	44.23
v=3.5	43.51	55.35	43.65	38.09	49.41	39.45
v=4	39.21	49.43	41.65	36.94	46.53	37.94
v=4.5	38.62	48.76	39.82	35.59	44.05	34.64
v=5	38.18	47.92	38.62	35.29	41.65	33.78

Note: IOp measures calculated with random forests with 500 replications and an alpha endogenously tuned following the K-fold Cross Validation technique explained in the Technical Appendix. Data from LWS.

Technical Appendix

The K-fold Cross Validation Method

Imposing a certain alpha level, such as 0.01 or 0.05, can bias the results, as they are exogenous econometric conventions that might impede us reflect all information gathered in the data. To solve this issue and test the accuracy of our results we implement a *K*-fold Cross Validation (CV) method, one of the most popular ML techniques (Rodríguez et al., 2009; Fushiki, 2011).

The CV divides the whole sample into K subsamples, also called folds. The optimal number of possible folds depends on the dataset. To make sure that we always have enough observations in every fold, we set $K = 5$. For robustness, we have also tested $K = 6$ and $K = 7$, without significant changes in our results.

The conditional inference tree algorithm is run on the union of $K - 1$ folds (training sample, *m*) for varying levels of alpha, while taking out the last k^{th} fold (validation sample, *k*). After that, we use the mean squared error (MSE) to evaluate the difference of the prediction in the training sample with respect to the validation sample:

$$MSE_k(\alpha) = \sum \frac{N^m}{N^k} \sum \frac{1}{N^m} \left(w_i^k - \mu^m(\alpha) \right)^2, \qquad (T.A.1)$$

where w_i^k is the output vector of the validation fold, and $\mu^m(\alpha)$ collects the predictions emanated from the training sample for a certain alpha level. Note that N^m and N^k are the sample sizes of the training and validation folds, respectively. This exercise is repeated four times more for the same alpha level, sequentially leaving out one *K*-fold at a time.

Finally, we construct the average mean squared error of the cross-validation process (MSE_{CV}) as the average of the five MSEs:

$$MSE_{CV}(\alpha) = \frac{1}{5} \sum_{k=1}^{5} MSE_k(\alpha) \qquad (T.A.2)$$

After running the algorithm for a set of possible alpha levels, we select the one providing the smallest MSE_{CV}.