Inequality of opportunity in wealth: accounting for differences between the US and Spain

Pedro Salas Rojo

Universidad Complutense de Madrid

March 4, 2019

Outline

Portfolio composition and wealth inequality

- The literature has generally defined two types of wealth: Financial and Real Estate.
- Financial wealth is the most important contributor to increases in wealth inequality (Bilias et al, 2005; Lusardi et al, 2017)
- A high diversity in the portfolio composition is found through the wealth distribution.
    - Richer and more educated households tend to invest more in finances (Gennaioli et al, 2014; Badarinza et al, 2016)

Cross-country comparisons

- In the last years, the literature has been centered on analyzing the origin of differences among countries.
- Bover (2010), Christelis et al (2013), Sierminska and Doorley (2017) and Cowell et al (2018) find that household's covariates are relatively homogeneous among countries. Hence, institutions would explain the lion's share of the differences.

Inequality of opportunity (IO)

- Inequality should not be related to circumstances beyond our own control (Rawls, 1971; Sen, 1980)
- In reality, these circumstances actually affect the individual's outcome (income, wealth, health...)
- Then, total inequality can be undestood as a combination of inequality of opportunity and inequality of effort (Roemer, 1993; Van de Gaer, 1993)
- We should minimize the former, so that outcomes are only conditioned by the latter.
- IO framework has not yet expanded to wealth analysis.

Contribution and motivation

- We analyze wealth general inequality and, more interestingly, wealth IO for US and Spain using post-crisis data.
- We check whether the results between both dimensions coincide.
- To do so, we use the decomposition method proposed by DiNardo, Fortin and Lemieux (1995)
- If they don't coincide, general wealth inequality analyses might be biased, as they could be mixing IO and IE effects.

Outline

The measurement of IO

In order to study IO, first we need to define types:

- Types are mutually exclusive groups of individuals that share certain circumstances (exogenous characteristics of individuals: Roemer, 1993).
- The literature considers that differences *between* types, once all circumstances have been taken into account, are explained by inequality in opportunities.
- Differences *within* each group would be explained by effort.

We should be worried about the former, and not about the latter.

The measurement of IO

- In order to measure IO, we follow the ex-ante parametric Ferreira and Gignoux (2011) approach.

$$ln(w_t) = \alpha C + \beta E + u$$
$$E = \delta C + v$$
$$ln(w_t) = (\alpha + \beta\delta)C + \beta v + u$$

- That can be estimated by using a simple OLS, and predicted as follows:

$$ln(w_t) = \phi C + \epsilon$$
$$\widetilde{w_t} = exp[\widehat{\phi C} + \widehat{\epsilon}]$$

- The new vector $\widetilde{w_t}$ is the type-correspondent wealth of each observation. Then, we apply MLD to this distribution. Note: we are only getting lower bound IO measures!

The measurement of IO

We use the MLD because it belongs to the family of Generalize Entropy Indexes, that share an interesting characteristic: they are additively decomposable, and its decomposition is consistent (Foster and Shenereyov, 2000). Then:

$$MLD^T = MLD^B + \sum \alpha MLD^W$$

MLD^B can be understood as an absolute IO measure:

$$MLD^B = \frac{1}{n} \sum ln \frac{\widetilde{w}_i}{\overline{\widetilde{w}}}$$

The ratio of MLD^B over MLD^T would be the relative IO measure.

The DFL decomposition

- The rest of our analysis is based on a set of counterfactuals, generated with the DiNardo-Fortin-Lemieux (DFL) method.
- This method redefines the sample weights of the treated country (Spain), in order to reflect the controled covariates of the refference country (US).

The DFL decomposition: the reweighting factor

- Take countries -i and -j, and consider w to be our objective variable. Also define a vector z of covariates. We define the wealth distribution as:

$$\int F(w|z,i)dF(z|i)$$

- Now impose the characteristics of -j

$$\int F(w|z,i)dF(z|{\color{red}j})$$

$$\int F(w|z,i)\Psi dF(z|i)$$

- Where, by means of the Bayes Rule we get to:

$$\Psi = \frac{dF(z|j)}{dF(z|i)} = \frac{P(j|z) * P(i)}{P(i|z) * P(j)}$$

The DFL decomposition: covariate effects

- We could also study the separate effect of each covariate:

$$F^i = \int F(w|a,e,y,l,h,i)dF(a|e,y,l,h,i)dF(e|y,l,h,i)$$
$$dF(y|l,h,i)dF(l|h,i)dF(h|i)$$

- Consider

$$F^1 = \int F(w|a,e,y,l,h,i)dF(a|e,y,l,h,j)dF(e|y,l,h,i)$$
$$dF(y|l,h,i)dF(l|h,i)dF(h|i)$$
$$F^2 = \int F(w|a,e,y,l,h,i)dF(a|e,y,l,h,j)dF(e|y,l,h,j)$$
$$dF(y|l,h,i)dF(l|h,i)dF(h|i)$$

The DFL decomposition: covariate effects

- The difference between F^i and F^1 would explain the effect of imposing age of -j into wealth distribution of -i.
- The difference between F^1 and F^2 would explain the effect of imposing education of -j into wealth distribution of -i, once we have already imposed the age distribution.
- Then, we could have:

$$F^i - F^j = [F^i - F^1] + [F^1 - F^2] + [F^2 - F^3] + [F^3 - F^4] + [F^4 - F^5] + [F^5 - F^j]$$

The DFL decomposition: covariate effects

- The problem: we don't know the correct order of the conditional expectations:

$$F^i = \int F(w|y, e, a, h, l, i)dF(y|e, a, h, l, i)dF(e|a, h, l, i)$$
$$dF(a|h, l, i)dF(h|l, i)dF(l|i)$$

- We overcome this problem by applying a Shapley decomposition on the covariates.

The DFL decomposition: Shapley decomposition

- The Shapley value decomposition solution is a concept that comes from the cooperative games theory.
- Our covariates can create up to 120 (5!) different coalitions. As we do not know which one is the correct, we just calculate them all and assume that they all have the same probability to happen.
- Thus, our reweighting factors are the result of the average of all possible combinations of covariates.

General idea

To recap:

- We take a vector z of US characteristics and we impose them on Spanish wealth distribution.
- Then, we generate a counterfactual that mixes US characteristics with Spanish institutions (and other unobservable variables)

This will provide several results:

- Actual difference = Spanish inequality - US inequality
- Compositional effect = Spanish inequality - Counterfactual inequality
- Residual = US inequality - Counterfactual inequality

Outline

The US and Spain databases

For our analysis we are using the SCF (2016) and the EFF (2014)
We selected these two countries because of:

- The complete wealth recording and the availability of circumstance data to create types, necessary for the IO analysis.
- The literature has found consistent differences among them:
  - Disparities on the welfare systems condition investment/consumption decisions, strongly affecting wealth.
  - Bover (2010) found that wealth inequality differences are particularly strong on the tails of the distribution.
  - Azpitarte (2012) analyzes the asset vulnerability.

Both studies use pre-crisis data (2001-2002) and are encapsulated in the traditional general inequality approach.

Types

Data to construct types is scarce, particularly in wealth surveys. In our database we have:

- Gender (men/women)
- Highest parental qualification (high/medium/low)
- Bequests (have received/not)

Thus, we are left with 12 different types.
Note that our methodology only makes sense if we run our counterfactuals by types.

Summary statistics: types

Variable	US	Spain
Gender (women)	26,25	45,4
Parents high qual.	25,96	22,62
Parents interm qual.	45,25	34,3
Parents low qual.	28,79	43,08
Have inherited	21.14	42.82

Table: Summary statistics of the type variables

Wealth definitions

We will use three gross wealth measures:

- Financial wealth: deposits, listed and unlisted shares, stocks, bonds...
- Real Estate wealth: real estate properties (houses, garages...)
- Total wealth: Financial wealth + Real Estate wealth

And, also, three debt measures:

- Financial debt: personal loans, negative credit balances...
- Real Estate debt: mortgages and real-estate related debts
- Total debt: Financial debt + Real Estate debt

Control variables

Our five control variables (vector z) are quite standard in the wealth inequality literature:

- Age: following Pfeffer and Killewald (2016) we defined ranks, from 25 to 74.
- Education attainment: illiterates and primary, secondary and tertiary.
- Income: divided in deciles.
- Labor status: three categories: workers, unemployed and others.
- Household structure: married(yes/no) and children(yes/no)

Outline

Wealth results

Factor	Income	Total	Financial	Real Estate
Spain GINI	0.34	0,603	0,834	0,586
US GINI	0.39	0,829	0,925	0,812
Counterfactual GINI	-	0,613	0,843	0,596
Actual difference	-0.05	-0,226	-0,091	-0,226
Compositional effect	-	-0,010	-0,009	-0,010
Compositional effect	-	4.40%	9.89%	4.44%
Residual	-	0,216	0,082	0,216
Residual	-	95.60%	90.11%	95.56%

Table: General wealth inequality results

Strong differences on wealth inequality between US and Spain.
Quite small compositional effect (around 0,01)
Match with the literature: actual differences are mainly explained by the residual.

Wealth results

Factor	Total	Financial	Real Estate
Contribution of variables	-0,010	-0,009	-0,010
Age	-0.030	-0.024	-0.024
Education	0.006	0.005	0.009
Income	-0.003	-0.015	-0.007
Labor	0.015	0.012	0.017
HH Structure	$7.41^{-0.4}$	0.013	-0.005

Table: General wealth inequality: Shapley decomposition

Positive sign: decreases inequality in the counterfactual.
Negative sign: increases inequality in the counterfactual.
This is the result of adding variables. But, what about the relative magnitude?

Wealth results

Factor	Total	Financial	Real Estate
Contribution of variables	-0,010	-0,009	-0,010
Age	300,58%	267,10%	249,49%
Education	-69,92%	-55,68%	-94,47%
Income	35,89%	176,03%	74,11%
Labor	-159,14%	-135,94%	-171,76%
HH Structure	-7,41%	-151,51%	42,63%

Table: General wealth inequality: Shapley decomposition

Age and income increased inequality in the counterfactual.
Strong and compensated effects among covariates.
No causal interpretation should be inferred from these results.

Wealth results

Factor	Total	Financial	Real Estate
Spain IO ratio	22.44%	19.48 %	12.06%
US IO ratio	10.44%	13.76%	5.26%
Counterfactual IO ratio	17.00%	11.12%	11.68%
Actual difference (pc)	12.00	5.72	6.80
Compositional effect (pc)	5.44	8.36	0.38
Compositional effect	45.41%	146.15%	5.58%
Residual (pc)	-6.56	2.64	-6.42
Residual	54.59%	(-)46.15%	94.42%

Table: Wealth IO results

Strong differences in IO between US and Spain.
Remarkable compositional effect: up to 40% in Financial wealth.
Differences are now explained by residual and covariates.

Wealth results

Factor	Total	Financial	Real Estate
Contribution of variables	5.448	8.355	0.380
Age	-140.61%	-44.25%	-1828.77%
Education	208.90%	91.63%	2180.62%
Income	25.43%	35.4%	251.85%
Labor	60.80%	20.57%	268.41%
HH Structure	-54.52%	-3.35%	-772.11%

Table: Wealth IO: Shapley decomposition

Positive sign: decreases inequality in the counterfactual.
Negative sign: increases inequality in the counterfactual.
Strong and compensated effects among covariates.
Age, education and labor have the same effect in both frameworks

Debt results

Factor	Total	Financial	Real Estate
Spain GINI	0,759	0,898	0,776
US GINI	0,730	0,781	0,766
Counterfactual GINI	0,750	0,887	0,770
Actual difference	0,029	0,117	0,010
Compositional effect	0,009	0,011	0,006
Residual	-0,020	-0,106	-0,004

Table: General debt inequality results

Small differences on debt inequality between US and Spain, except for financial debt

Quite small compositional effect (smaller than 0,01)

Residuals still explain the lion's share (except in Real Estate, where there is no difference)

Debt results

Factor	Total	Financial	Real Estate
Contribution of variables	0,008	0,011	0,006
Age	-721,97%	-92,25%	-950,57%
Education	560,16%	62,35%	789,44%
Income	-395,15%	-70,41%	-519,20%
Labor	186,07%	62,93%	227,99%
HH Structure	470,89%	137,38%	552,25%

Table: General debt inequality Shapley decomposition

Positive sign: decreases inequality in the counterfactual.
Negative sign: increases inequality in the counterfactual.
Strong and compensated effects among covariates.
Exactly the same as wealth general inequality.

Debt results

Factor	Total	Financial	Real Estate
Spain IO ratio	5.55%	4.26%	4.84%
US IO ratio	4.63%	3.17%	2,84%
Counterfactual IO ratio	5.39%	3.96%	4.77%
Actual effect (pc)	0,92	1,09	2.00
Compositional effect (pc)	0,16	0,30	0.07
Residual (pc)	-0.76	-0.79	-1.93

Table: Debt IO results

Quite small IO ratios for all three types of wealth.
The counterfactual makes a small difference.
Lion's share of differences are attributed to institutions.

Debt results

Factor	Total	Financial	Real Estate
Contribution of variables	0.163	0.295	0,074
Age	-1026,35%	-703,94%	-2242.32%
Education	1042,05%	1134,22%	1998.85%
Income	301,12%	-285,21%	494.40%
Labor	255,04%	839,06%	600.34%
HH Structure	-471,86%	-884,13%	-751.29%

Table: Debt IO Shapley decomposition

Strong and opposed effects.

Same signs as wealth IO. Results seem to be robust even for negative wealth.

Age and education have the strongest effects, mutually compensating each other.

Outline

General inequality analysis

- There are remarkable differences between Spain and US wealth distributions.
- Imposing US covariates into Spain do not seem to change inequality in the receptor country (Spain).
- Then, those differences can be attributed to institutions.
- Age and income distributions in the US contributed to increase inequality in the counterfactual.
- Education, labor and HH Structure decreased inequality in the counterfactual.

IO analysis

- We also find remarkable differences: US has greater equality of opportunities.
- Imposing US covariates into Spain generates a quite different counterfactual in financial and total wealth.
- Then, differences can be attributed to covariates and institutions.
- Education, income and labor increased decreased inequality in the counterfactual.
- Age and HH structure increase inequality in the counterfactual.

Take-home ideas

- IO and general inequality approaches do not provide the same results in the counterfactuals.
- The Shapley decomposition provides mixed results:
  - Age always increases inequality.
  - Education and labor always decrease inequality.
  - Income and Household Structure have different effects.
- Thus, some covariates seem to behave differently, and others the same. Anyway, they seem to be relevant to explain IO.

Outline

Summary statistics: wealth measures

Variable	Mean	Sd	p10	p50	p90
Total Assets	310,892	1188,364	3,296	174,055	599,574
Financial	59,614	522,142	0,11	8,349	115,339
Real Estate	251,08	946,505	0,002	153,786	510,292
Total debt	41,259	77,848	0,001	1,209	137,308
Fin. Debt	2,569	11,236	0,001	0,001	7,69
R. E. Debt	38,69	76,042	0,001	0,001	131,816

Table: Summary statistics of wealth and debt measures. All values are expressed in thousand USD of 2016

Summary statistics: covariates

Variable	US	Spain
Age (mean)	51,19	52,65
High educated	35,26	24,37
Intermediate education	62,16	41,09
Low education	2,58	34,55
Worker	62,09	47,64
Unemployed	3,58	17,28
Other	34,33	35,04
Single	41,12	40,73
With Kids	43,73	31,5

Table: Summary statistics of the covariates

MLD results

Factor	Total	Financial	Real Estate
Spain	1.235	2.526	1.890
US	2.551	3.423	4.168
Counterfactual	1.311	2.204	1.987

Table: General wealth inequality results measured with MLD

Factor	Total	Financial	Real Estate
Spain	5.116	5.687	6,113
US	3.435	3.787	5.863
Counterfactual	5.896	5.761	6.864

Table: General debt inequality results measured with MLD

Application of the Bayes Rule to get the Ψ

The Bayes rule is defined as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Then,

$$\Psi = \frac{P(z|j)}{P(z|i)} = \frac{\frac{P(j|z)P(z)}{P(j)}}{\frac{P(i|z)P(z)}{P(i)}} = \frac{P(j|z)P(i)}{P(i|z)P(B)}$$