Evidence and Drivers of Income Polarization in Italy

A Relative Distribution Analysis Using Recentered Influence Function Regressions

F. Clementi M. Fabiani

University of Macerata, Italy

Conference "Income and Wealth Inequality: Drivers and Consequences",
Gdańsk University of Technology, Poland

Outline

Motivation

- Income and wealth inequality have been long-standing issues in Italy.
- Despite being one of the largest economies in Europe, economic inequality has been on the rise in the country since the 1980s, with the wealthiest individuals accumulating more wealth while the lower-income groups continue to struggle.
- The issue of inequality in Italy is a complex one, with factors such as regional disparities, high unemployment rates, and the presence of a large informal economy contributing to the problem.
- One aspect that the distributional studies focusing on Italy have only marginally considered is that of income polarization, which considers not only how far the income distribution strays from its center [1, 2] but also how some earnings groups (poles) form around local means [3].
- The polarization phenomenon is frequently viewed as dangerous because it denotes a decline in social cohesion, which can result in social conflict [4].
- In Italy, the analysis of income polarization has helped to explain the discrepancy between empirical evidence of stability in distributional indices and the deterioration of confidence and expectations among Italian households in the 2000s [5].
- The analysis of income polarization has also emerged as crucial in assessing the effects of economic crises like the Great Recession of 2007–2009, which led to widespread deprivation and widened gaps between social groups in the Italian income distribution [6].

Aim of This Work

- The contribution of this paper is twofold.
- First, it uses a very intuitive method, the "relative distribution", to analyze the distributional changes occurred in Italy between 1991 and 2020; the strength of this method rests in providing a non-parametric framework for taking into account all the distributional differences that could arise in the comparison of distributions over time and that would not be detected easily from a comparison of standard measures of inequality.
- Second, and most importantly, the paper develops within the relative distribution framework a novel methodology to identify the covariates of distributional changes; the main value added being it enables a very specific and useful analysis of the factors that cause income polarization.

Polarization: Conceptualization and Measurement

- While related to inequality, polarization is a distinct concept; in general, it is concerned with the «[. . .] appearance (or disappearance) of groups in a distribution» [7, p. 105], whereas inequality provides an indication of the overall dispersion of a distribution.
- We find two families of polarization metrics in the literature [8]:
  - → the first family defines polarization as the process by which a distribution becomes "bi-polarized", i.e. a process by which incomes increasingly cluster at the distribution's extremities and the middle class disappears; Foster and Wolfson [1, 2, 9, 10] were the first to propose income polarization measures that fit this perspective;
  - → indices in the second family of measures regard polarization as the "clustering" of a population around two or more poles of the income distribution—regardless of where they fall on the income scale— which might give rise to social conflicts and tensions; indicators based on the idea of income polarization as a conflict between various groups have been primarily conceived by Esteban, Ray, and coauthors [3, 11, 12].
- Another strand of research combines kernel density estimates and mixture models to explain how polarization patterns evolve over time in both personal incomes [13] and the cross-country distribution of per-capita income [14, 15].

The Relative Distribution Approach

- To investigate Italian income polarization, we use the "relative distribution" method [16, 17].
- The relative distribution can be obtained as the ratio of the density of the "comparison" population to the density of the "reference" population and describes where households at various quantiles in the comparison distribution are concentrated in terms of the quantiles of the reference distribution.
- This method offers the possibility to decompose the relative distribution into changes in "location" and changes in "shape":
  - → the "location effect" represents the relative distribution if there had been no change in distributional shape but only a location shift of the density;
  - → the "shape effect" represents the relative density net of the location effect and isolates differences in distributional shape between the two populations.
- One can also use summary measures to quantify the observed pattern of changes:
  - → the "median relative polarization" (MRP) index captures the degree to which there is divergence from, or convergence toward, the center of the distribution;
  - → the "lower relative polarization" (LRP) and the "upper relative polarization" (URP) indices capture the contributions to distributional change made by the segments of the distribution above and below the median.

Recentered Influence Function (RIF) Regression

- One novelty of this study is the use of "recentered influence function" (RIF) regression to analyze the drivers of income polarization in Italy.
- RIF regression is a statistical tool for the analysis of partial effects of explanatory variables on selected distributional statistics [18–21].
- The RIF is calculated by adding back the distributional statistic to the influence function.
  - → The "influence function" is an analytical tool that has long been used for analyzing the robustness of distributional statistics [22–24].
- The importance of this transformation lies in the fact that the expectation of the RIF is precisely the distributional statistic of interest.
- The RIF is used as the dependent variable in an ordinary least squares regression to capture how small changes in the distribution of the independent variables affect the distributional statistic.

RIF Regression for Relative Polarization Indices

- In practice, following [21], we first compute the RIF of the MRP index for each income i.
- The coefficients β can be estimated by OLS through the following equation:

$$RIF_i\left(\widehat{MRP}\right) = \alpha + \sum_{k=1}^{K} \beta_k x_{i,k} + \epsilon_i$$

- The estimated model parameter $\hat{\beta}_k$ can be interpreted as the effect of a small change in the distribution of X_k on $\widehat{MRP}$—when the distribution of other covariates remains unchanged—or as linear approximation of the effect of large changes of X_k on $\widehat{MRP}$ [25].
- RIF regression for the LRP and URP can be derived analogously.

The Survey on Household Income and Wealth

- Income data are drawn from the Historical Archive (version 11.3, released in March 2023) of the Survey on Household Income and Wealth (SHIW), a representative survey of the Italian resident population conducted by the Bank of Italy since the mid-1960s to gather information on income, saving, consumption expenditure, wealth, demographics, and labor force participation of Italian households.
- Specifically, the data set employed includes two independent cross sections of Italian households referring to 1991 and 2020, for a total of 14,427 observations.
- The basic definition of income is net of taxation and social security contributions; it is the sum of four main components: compensation of employees; pensions and net transfers; net income from self-employment; property income (including income from buildings and income from financial assets).
- The income variable is adjusted for differences in household size and composition using the "modified OECD" equivalence scale and weighted by the provided sampling weights; in addition, household incomes (expressed in euros) are examined at 2020 prices.

Table 1: Summary measures of Italian household disposable income, 1991 and 2020

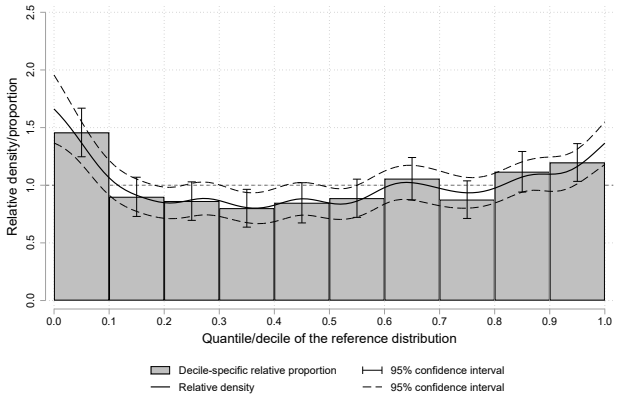
	1991	2020
1st decile	4,388.0	7,202.1
1st quartile	6,023.6	11,113.7
Median	8,671.9	17,112.7
Mean	9,811.4	19,741.7
3rd quartile	12,252.5	24,442.5
9th decile	16,314.7	33,575.9
Gini index	28.8	33.3
Low-income individuals[a]	16.7	21.4
Foster-Wolfson index	24.8	28.0

[a] Threshold 60% of the median.

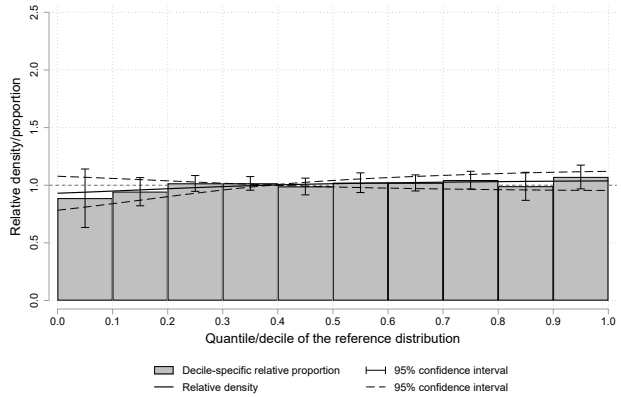
Changes in the Italian Personal Income Distribution

- The relative distribution plots reveal that, despite the higher median, the greater dispersion of household income in the 2020 edition of the SHIW resulted in relatively more low-income households than in 1991, and this effect was mainly concentrated in the bottom decile.
- In contrast, at the top end the higher spread increased the share of households in the last decile of the 2020 income distribution by almost 17%.
- The comparison income distribution, therefore, is more polarized than the reference income distribution: there are more comparison observations at the top and bottom of the income scale, and fewer in the middle.
- The size and sign of the estimated relative polarization indices confirm the impression left by the graphical display.

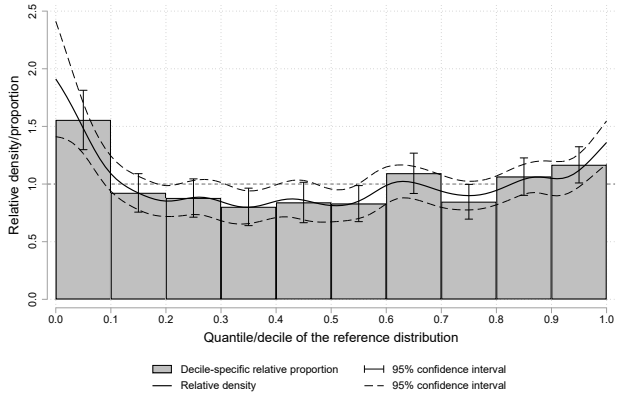
Figure 1: Relative distribution plots



(a) Overall relative density



(b) Location effect



(c) Shape effect

Table 2: Relative polarization indices

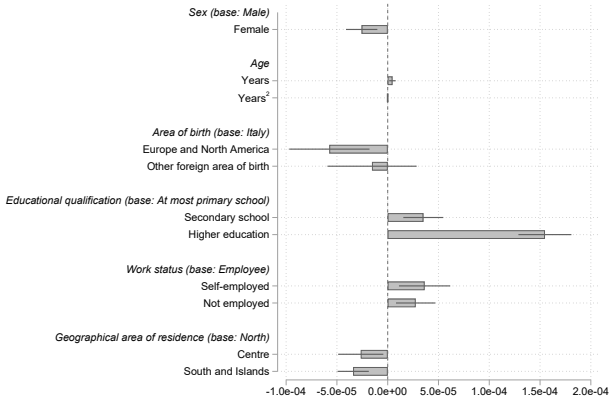
Index[a]	Value	LB[b]	UB[c]	*p*-value[d]
MRP	0.098	0.059	0.137	0.000
LRP	0.137	0.064	0.211	0.000
URP	0.058	0.010	0.106	0.017

[a] MRP = median relative polarization index, LRP = lower relative polarization index, URP = upper relative polarization index.
[b] Lower bound of the 95% confidence interval.
[c] Upper bound of the 95% confidence interval.
[d] Refers to the null hypothesis of no change with respect to the reference distribution, i.e. that the index equals 0.

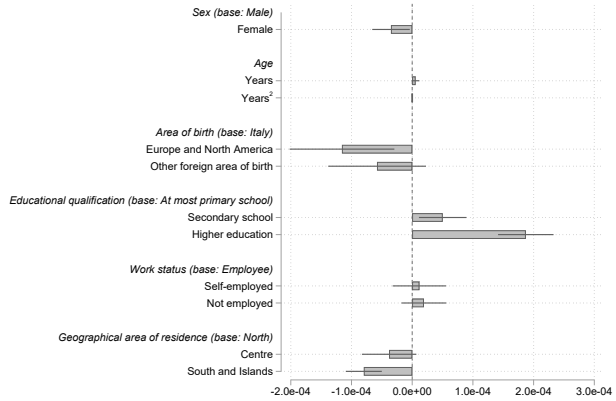
The Effects of Covariates on Relative Polarization Indices

- In estimation of RIF regressions for relative polarization indices we consider a set of SHIW variables accounting for the role of some potential determinants of household income and its polarization.
- Specifically, we control for:
  - → demographic characteristics—sex, age, and foreign area of birth (Europe and North America or other foreign country) of the household's head;
  - → educational attainment—a dummy representing three educational levels (at most primary school, secondary school, and higher education);
  - → types of employment status—employee, self-employed, and not employed);
  - → geography—a dummy variable allowing the division of Italy into three geographical areas (North, Centre, and South and Islands).
- Statistical estimates of coefficients from RIF-regression models can be displayed in form of graphs.

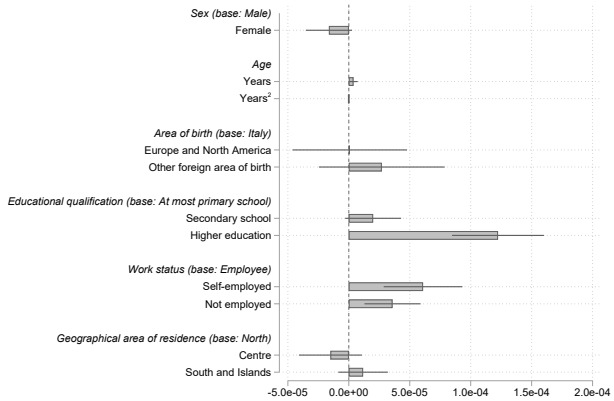
Figure 2: RIF regression for relative polarization indices



(a) MRP



(b) LRP



(c) URP

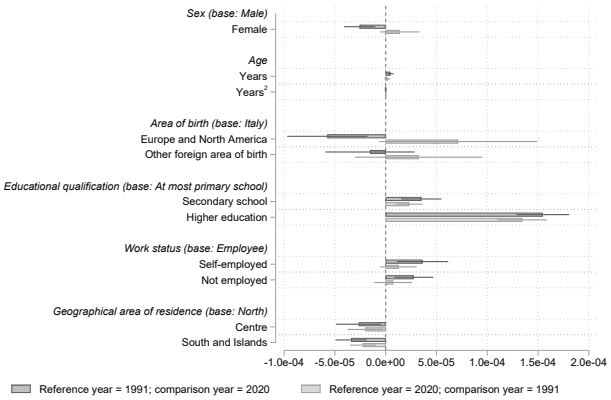
Main Results

- *Demographic characteristics*
  - → Sex: female-headed households reduce polarization, mostly in the lower tail.
  - → Age: polarization first increases with age (positive sign of the regression coefficient) and then begins to decrease (negative sign of the age-squared coefficient).
  - → Area of birth: compared with native Italians, households headed by foreign-born (mainly in Europe and North America) have a counter-polarization effect.
- *Educational attainment*
  - → Higher levels of education lead to greater polarization, a result that is significant especially in the upper tail of the distribution.
- *Work status*
  - → Being self-employed or not employed, compared to being employed, negatively impacts (increases) polarization.
- *Geographical area of residence*
  - → Living in the Centre or South/Islands reduces the impact on polarization.

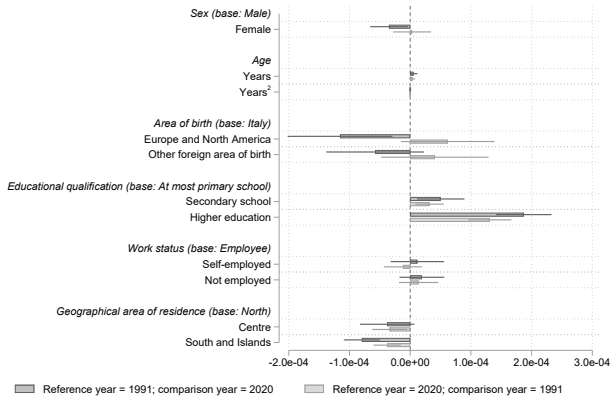
Swapping the Reference and Comparison Distributions

- We run a couple of checks to verify the robustness of the RIF-regression results.
- First, we swap the reference and comparison distributions, assuming that the reference distribution is for 2020 and the comparison distribution is for 1991.
  - → As the relative polarization indices are "anti-symmetric"—meaning that swapping the comparison and the reference will yield indices with the same magnitude but opposite sign [17, 26]—it is expected that the sign of the estimated regression coefficients will not change.
  - → In fact, if an explanatory variable has a pro-polarization (counter-polarization) effect when the comparison distribution of 2020 is more polarized than the reference distribution of 1991, then the same variable must keep the sign of its coefficient unaltered when the two distributions are swapped, hence indicating that its effect is of counter-polarization (pro-polarization) when we consider the 1991 distribution as a comparison and the 2020 distribution as a reference—with the former being less polarized.
- Results from this check are similar to those obtained in the leading case, and show that for the bulk of the explanatory variables in the model the coefficient estimates keep the sign unchanged.
- In those few cases where the sign of the regressor changes, the coefficient is not statistically different from zero.

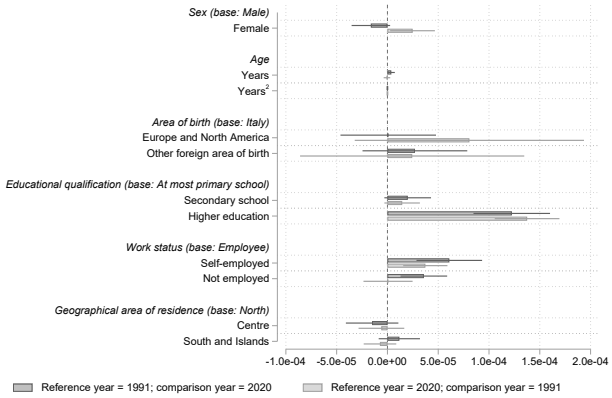
Figure 3: RIF regression with swapped comparison and reference distributions



(a) MRP



(b) LRP

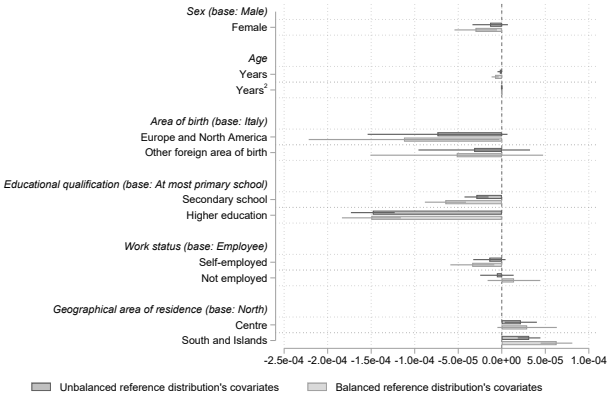


(c) URP

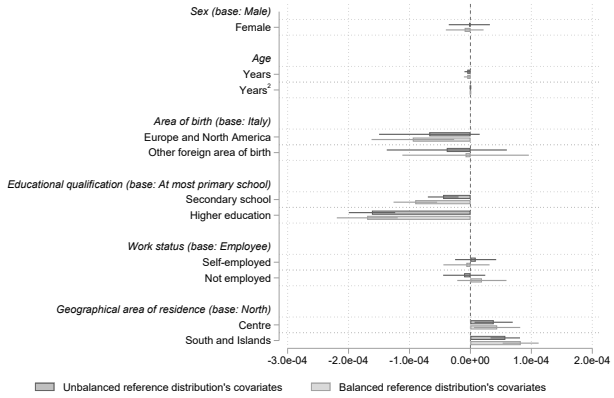
Adjusting the Distributions of RIF-regression Covariates

- A further validation check of the main findings was undertaken by adjusting the distributions of RIF-regression covariates in the reference and comparison years.
- The adjustment was achieved by reweighting the covariate distributions from one year such that their first, second, and possibly higher moments were equal to the sample moments from their distributions in the other year.
- This check mainly helps to ensure that the findings produced are not affected by the year selected for the covariates, which in the leading case is 2020.
- By comparing our leading case estimates with estimated RIF-regression coefficients after reweighting of the reference year's covariate distributions, it is observed that for many of the estimated coefficients the sign coincides in the two cases considered—the occurrences in which the sign of the coefficients changes are infrequent and non-significant—an indication that it is irrelevant whether one uses the original covariates from the comparison year 2020 or those from 1991 after they have been reweighted to have the same distributional characteristics as those from 2020.
- The same conclusion can be drawn by comparing leading case estimates with estimated RIF-regression coefficients after comparison year's covariate distributions are reweighted, which means that the use of the original 2020 comparison year's covariates or the reweighted covariates to have the same distributional characteristics as those in 1991 is largely irrelevant.

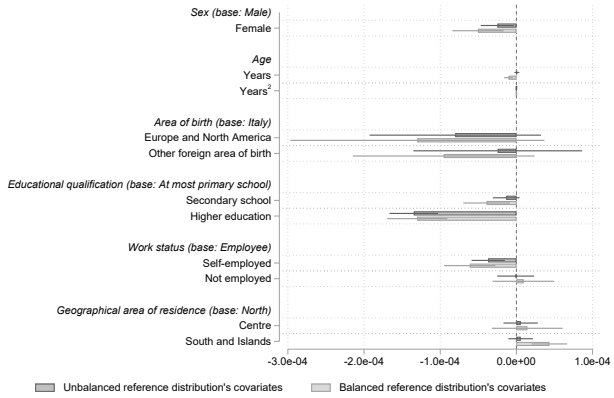
Figure 4: RIF regression with reweighted reference distribution's covariates



(a) MRP

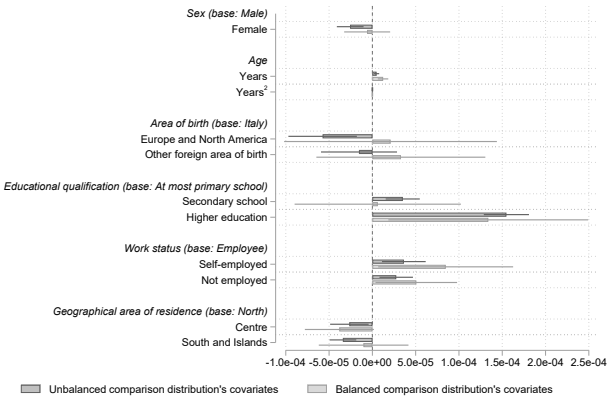


(b) LRP

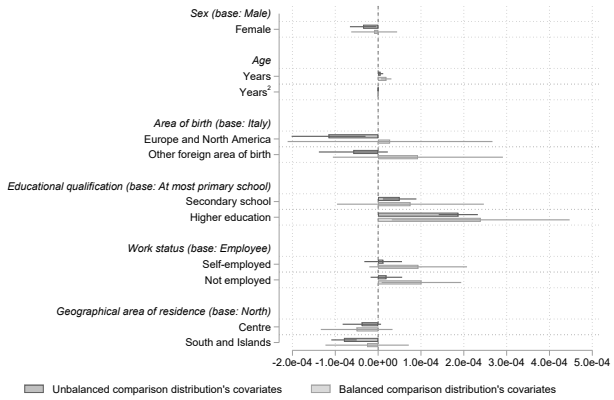


(c) URP

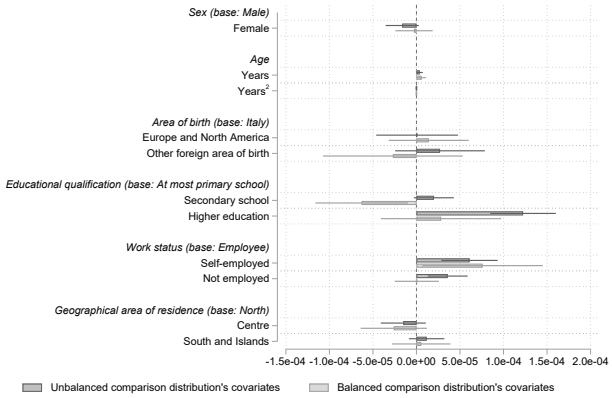
Figure 5: RIF regression with reweighted comparison distribution's covariates



(a) MRP



(b) LRP



(c) URP

Summary of Findings

- The topic of the growing gap between the rich and the poor is gaining momentum thanks, in particular, to the large attention that has been obtained in recent research on world inequalities [27–32]; the general idea that can be derived from this investigation is that throughout the most recent 30 to 40 years, both developing and affluent countries went through substantial distributional changes that aggravated inequalities.
- In Italy, researchers and policymakers have focused on distributional concerns due to "household impoverishment", particularly in the middle class; however, most of the analyses that have dealt with the vulnerability of the middle class, the fall in households' purchasing power, and their difficulty in making ends meet, are based on summary statistics of household incomes, which may not capture the whole income distribution, possibly leaving some of the most significant changes that have occurred unexplored.
- In this work, we have used the relative distribution method to analyze changes in the Italian personal income distribution between 1991 and 2020; in contrast to methods that rely on summary statistics, this nonparametric approach uses all information about the shape of the distributions.
- The relative distribution analysis was able to document a clear tendency to polarization in household incomes: in fact, the results suggest that the distributional changes hollowed out the middle of the Italian personal income distribution and increased the concentration of households around the highest and lowest deciles, thus generating a U-shaped distribution relative to the baseline.
- The other main contribution of this paper was proposing a tool that quantifies the drivers of the observed polarization process by blending two different frameworks of distributional analysis: the relative distribution and the unconditional quantile regression; the advantage over other methodologies is that it allows to directly relate the impact of changes in the expected values of the covariates on the relative polarization indices.
- Results indicate that characteristics of the household's head such as not being elderly, having attained the highest levels of education, and with a working condition other than employed, all tend to increase the size of the upper and lower tails of the income distribution over time and, consequently, the degree of polarization; a counter-polarization effect is observed, on the other hand, when the household's head is female, was born in a foreign country, or lives in the Centre or South/Islands of the country, due mainly to the lower concentration of income in these population subgroups compared to the reference category.

[1] J. E. Foster and M. C. Wolfson. *Polarization and the Decline of the Middle Class: Canada and the US*. Working Paper 31. Available at https://ophi.org.uk/working-paper-number-31/. Oxford: Oxford Poverty & Human Development Initiative, 1992.

[2] J. E. Foster and M. C. Wolfson. "Polarization and the Decline of the Middle Class: Canada and the U.S.". In: *The Journal of Economic Inequality* 8 (2010), pp. 247–273.

[3] J.-M. Esteban and D. Ray. "On the Measurement of Polarization". In: *Econometrica* 62 (1994), pp. 819–851.

[4] J.-M. Esteban and D. Ray. "Linking Conflict to Inequality and Polarization". In: *The American Economic Review* 101 (2011), pp. 1345–1374.

[5] R. Massari, M. G. Pittau, and R. Zelli. "A Dwindling Middle Class? Italian Evidence in the 2000s". In: *Journal of Economic Inequality* 7 (2009), pp. 333–350.

[6] C. A. Ricci and S. Scicchitano. "Decomposing Changes in Income Polarization by Population Group: What Happened During the Crisis?" In: *Economia Politica* 38.1 (2021), pp. 235–259.

[7] S. R. Chakravarty. *Inequality, Polarization and Poverty: Advances in Distributional Analysis*. New York, NY: Springer-Verlag Inc., 2009.

[8] J.-Y. Duclos and A.-M. Taptué. "Polarization". In: *Handbook of Income Distribution, Volume 2A*. Ed. by A. B. Atkinson and F. Bourguignon. Amsterdam: North-Holland, 2015, pp. 301–358.

[9] M. C. Wolfson. "When Inequalities Diverge". In: *The American Economic Review* 84 (1994), pp. 353–358.

[10] M. C. Wolfson. "Divergent Inequalities: Theory and Empirical Results". In: *Review of Income and Wealth* 43 (1997), pp. 401–421.

[11] J.-Y. Duclos, J.-M. Esteban, and D. Ray. "Polarization: Concepts, Measurement, Estimation". In: *Econometrica* 72 (2004), pp. 1737–1772.

[12] J. Esteban, C. Gradín, and D. Ray. "An Extension of a Measure of Polarization, with an Application to the Income Distribution of Five OECD Countries". In: *The Journal of Economic Inequality* 5 (2007), pp. 1–19.

[13] S. P. Jenkins. "Did the Middle Class Shrink During the 1980s? UK Evidence From Kernel Density Estimates". In: *Economic Letters* 49 (1995), pp. 407–413.

[14] G. Anderson, O. Linton, and T. W. Leo. "A Polarization-Cohesion Perspective on Cross-Country Convergence". In: *Journal of Economic Growth* 17 (2012), pp. 49–69.

[15] D. T. Quah. "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics". In: *The Economic Journal* 106 (1996), pp. 1045–1055.

[16] M. S. Handcock and M. Morris. "Relative Distribution Methods". In: *Sociological Methodology* 28 (1998), pp. 53–97.

[17] M. S. Handcock and M. Morris. *Relative Distribution Methods in the Social Sciences*. New York, NY: Springer-Verlag Inc., 1999.

[18] S. Firpo, N. M. Fortin, and T. Lemieux. "Unconditional Quantile Regressions". In: *Econometrica* 77 (2009), pp. 953–973.

[19] B. Essama-Nssah and P. J. Lambert. "Influence Functions for Policy Impact Analysis". In: *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber*. Ed. by J. A. Bishop and R. Salas. Vol. 20. Bingley, UK: Emerald Group Publishing Limited, 2012, pp. 135–159.

[20] F. Rios-Avila. "Recentered Influence Functions (RIFs) in Stata: RIF Regression and RIF Decomposition". In: *The Stata Journal* 20 (2020), pp. 51–94.

[21] B. Jann. "Relative Distribution Analysis in Stata". In: *The Stata Journal* 21 (2021), pp. 885–951.

[22] F. A. Cowell and M.-P. Victoria-Feser. "Robustness Properties of Inequality Measures". In: *Econometrica* 64 (1996), pp. 77–101.

[23] F. A. Cowell and E. Flachaire. "Income Distribution and Inequality Measurement: The Problem of Extreme Values". In: *Journal of Econometrics* 141 (2007), pp. 1044–1072.

[24] F. A. Cowell and E. Flachaire. "Statistical Methods for Distributional Analysis". In: *Handbook of Income Distribution*. Ed. by A. B. Atkinson and F. Bourguignon. Vol. 2A. Amsterdam: North-Holland, 2015, pp. 359–465.

[25] S. Firpo, N. Fortin, and T. Lemieux. "Decomposing Wage Distributions Using Recentered Influence Function Regressions". In: *Econometrics* 6 (2018), p. 28.

[26] L. Hao and D. Q. Naiman. *Assessing Inequality*. Thousand Oaks, CA: SAGE Publications, Inc., 2010.

[27] B. Milanovic. *The Haves and the Have-Nots: A Brief and Idiosyncratic History of Global Inequality*. New York, NY: Basic Books, 2010, p. 280.

[28] J. E. Stiglitz. *The Price of Inequality: How Today's Divided Society Endangers Our Future*. New York, NY: W. W. Norton & Company, 2012.

[29] T. Piketty. *Capital in the Twenty-First Century*. Cambridge, MA: Belknap Press of Harvard University Press, 2014.

[30] A. B. Atkinson. *Inequality: What Can Be Done?* Cambridge, MA: Harvard University Press, 2015, p. 384.

[31] B. Milanovic. *Global Inequality: A New Approach for the Age of Globalization*. New York, NY: Belknap Press of Harvard University Press, 2016.

[32] J. E. Stiglitz. *The Great Divide: Unequal Societies and What We Can Do about Them*. New York, NY: W. W. Norton & Company, 2016.

Thank you all!