Received: 6 November 2017 | Revised: 19 December 2018 | Accepted: 7 January 2019

DOI: 10.1111/rsp3.12181

ORIGINAL ARTICLE

Inequality of opportunity, a matter of space?

Jesus Perez-Mayo

Department of Economics, University of Extremadura, Elvas Av. 06006, Badajoz, Spain

Correspondence
Jesus Perez-Mayo, Department of Economics, University of Extremadura, Elvas Av. 06006 Badajoz, Spain.
Email: jperez@unex.es

JEL Classification: D63; D6; R00

Abstract

Inequality has become a very important issue in developed countries in the last decades. Among the different definitions of inequality, inequality of opportunity stands out as one of the main topics. At what extent are personal circumstances out of control the root of unequal outcomes? Space, defined as the place where people live, can constrain the set of opportunities that people face or, even, people with the same personal circumstances, excepting region, and the same set of opportunities can obtain very different outcomes depending on the region. This issue is measured in Spain by using the 2011 EU-SILC microdata.

KEYWORDS

Equity, Justice, Inequality, and Other Normative Criteria and Measurement, Welfare Economics, Urban, Rural, and Regional Economics

1 | INTRODUCTION

Equality of opportunity is one of the main issues in modern societies. It can be defined as the independence of current (future) living conditions from past (current) ones. Gender, family, education, or social class are usually found in the literature among these background characteristics that can constrain well-being, thus provoking the opposite phenomenon, inequality of opportunity.

Nevertheless, place of birth or residence is one of the personal characteristics that deeply affect future personal life. The influence of place in inequality replicates the discussion between "place poverty" and "people poverty" of the territorial factors compared with the personal ones. For instance, being unemployed is the main factor or there are some regional peculiarities—regional labour markets—that make this factor more serious.

The same as growing in a poor family makes more likely being a poor adult, growing or living in an impoverished area can cause a higher risk of poverty in adulthood. Education levels and quality, employment chances, economic

TABLE 1 Regional inequality indices 2011 (restricted sample)

Region	index	Population share	Absolute contribution	Relative contribution
1: Andalucia	0.2659	0.1751	0.0466	0.2108
2: Aragon	0.1933	0.0279	0.0054	0.0244
3: Asturias	0.1669	0.0233	0.0039	0.0176
4: Balearic Is.	0.2058	0.0240	0.0049	0.0223
5: Canary Is.	0.3174	0.0468	0.0149	0.0673
6: Cantabria	0.1943	0.0129	0.0025	0.0114
7: Castilla-Leon	0.1914	0.0522	0.0100	0.0452
8: Castilla-La Mancha	0.2706	0.0445	0.0120	0.0545
9: Catalonia	0.1883	0.1637	0.0308	0.1395
10: Valencia	0.1693	0.1053	0.0178	0.0807
11: Extremadura	0.2478	0.0229	0.0057	0.0257
12: Galicia	0.2098	0.0576	0.0121	0.0547
13: Madrid	0.1975	0.1439	0.0284	0.1286
14: Murcia	0.1343	0.0315	0.0042	0.0192
15: Navarre	0.1495	0.0141	0.0021	0.0095
16: Basque Country	0.1661	0.0472	0.0078	0.0355
17: Rioja (La)	0.2409	0.0069	0.0017	0.0075
Within	---	---	0.2109	0.9545
Between	---	---	0.0112	0.0509
Population	0.2209	1.0000	0.2209	1.0000

Source: Author's elaboration from Stata DASP module.

dynamism differ between regions or counties. Therefore, it is critical to consider the space in the analysis of the inequality of opportunities.

This paper follows this research line and tries to figure out to what extent place limits and affects people's life developments. To perform this analysis, a decomposition of inequality measures based on counterfactuals will be used. Thus, it will be possible to distinguish at what extent subjective circumstances constrain people opportunities to achieve life goals. By following this approach, spatial inequality of opportunities will be estimated when the same person would occupy a different position in the income distribution or would have different poverty risk if he/she lived or grown up in a different place. If the same people with different treatment will obtain different life outcomes, inequality of opportunities will be found.

Regional inequalities are widely studied in the literature. Theil or entropy inequality indices are the most used due to the distinctive property: additive decomposition.[1] Thus, if a population can be divided into some exhaustive and mutually exclusive groups, the inequality index for the whole population can be calculated as a linear combination of the group indices, where the relevance of each index is the population share of each group. When the grouping variable is spatial—countries, regions, provinces or districts—this decomposition procedure can show the share in inequality of every spatial level. Moreover, global inequality can be split into two components, between groups and within groups, as Table 1 will report in this paper. Even, some authors as Akita (2003) proposed the decomposition in more than two spatial levels and other (Márquez, Lasarte-Navamuel, & Lufin, 2016; Márquez, Lasarte-Navamuel, & Lufin, 2017) use this procedure to include some "neighbour" effects in the decomposition.

[1]Some authors use Gini indices by using the decomposition methodology proposed by Silber (1989) instead of Theil.

In the analysis of regional disparities, one can find, on one hand, papers where regions are the units of analysis and, therefore, the interregional inequality is their goal. On the other side, some papers study the interpersonal inequalities by using the spatial dimension as the grouping variable (for instance, Ayala, Jurado, & Pedraja, 2009). However, this paper goes beyond. Instead of analysing income inequality, its aim is studying inequality of opportunities by considering the region as a limiting factor. This kind of analysis is still rare to be found in the literature of inequality of opportunities. Only some papers try to measure this issue in French (Carpantier & Sapata, 2013) or Italian (Checchi & Peragine, 2010) regions.

In this case, the measurement of regional inequalities of opportunity will be explored for the Spanish regions. Spain shows two features that make very interesting to examine this phenomenon. First, it is one of the most decentralized countries in the world where some public policies as education, health or job search services are managed and funded by Regional Governments. Therefore, the same person may have to face different life risks or can find dissimilar education systems depending on her place of residence. Besides, poverty and inequality regional disparities have been observed in Spain from the last decades of the last century, even despite the economic and social development experienced by the whole country. Maps in Appendix Figure A1 show that, although some regions have changed their relative ranks and there is some level of convergence from the 1970s, it is very likely to find that the same regions in the extremes of the distribution in each decade are found. Therefore, inquiring about the role of space or regions in the inequality of opportunities gains importance.

These differences are not simply related to income poverty. As the Table A1 in the Appendix shows, regional disparities clearly appear in issues such as GDP, incomes, and wages. In addition, labour markets in Spain do not depict a common picture with different unemployment rates, mainly when youth unemployment is analysed, or labour participation or activity rates for women. Lastly, a very important issue related to the capacity of being competitive and flexible in the global markets can be found in the data for school dropout rates or the number of firms that work in the hi-tech sectors. All these topics usually overlap so that opportunities set can be assumed to be constrained in some regions.

In addition, this study is possible because of the increasing data availability. EU-SILC microdata for Spain will be used due to their regional (NUTS 2) statistical significance, their special design to collect data about income, social exclusion, and living conditions as well as annual modules specially designed to measure the intergenerational transmission of poverty.

The paper is organised as follows. Section 2 presents the methodology and data used. Then, empirical results after applying the methodology to 2011 special module dataset are reported and discussed in Section 3 and, finally in Section 4, some concluding remarks are enumerated.

2 | METHODOLOGY AND DATA

The definition of inequality of opportunity lies on the following idea: not all the inequalities are equally bad. Some part of the unequal outcomes that any person can observe may be due to different levels of personal effort, another part, perhaps, stems from features out of personal control, such as, as parents' education or financial difficulties in childhood and, finally, luck can be the factor that explains why people who grew up in the same background end up with different life outcomes, even after making the same efforts. Roemer (1998) states that mainly the first inequalities should be removed because not all the inequalities are equally bad, while others do not share the same level of priority. Additionally, according to Checchi and Peragine (2010), the second kind of inequality could be described as non-offensive in ethical terms. Since effort is expected to lead to higher levels of outcome, inequality due to effort could be acceptable or even desirable. Thus, ethically offensive inequality will be the part due to circumstances out of control of people. These circumstances are all elements that individuals cannot change with their effort and that impact on their outcome or advantages, following Roemer's approach. Some examples of them are family background, gender, race or region, in this case. Therefore, if a perfect equality of opportunities exists, people's outcome will be unaffected by people's circumstances. Which then, is the objective of this analysis? The goal is to decompose total inequality into two parts: ethically offensive and non-offensive.

There are many approaches to empirically apply the above definitions. One of them, proposed by Bourguignon, Ferreira, and Menéndez (2007), can be called "regression approach." In this case, the outcome is estimated as a function of circumstances and efforts by using a linear regression model, which is used to simulate a counterfactual distribution of outcomes being removed the effect of circumstances. Afterwards, the actual distribution is compared with the simulated counterfactuals so that aggregate inequality is split into a component coming from circumstances and a residual.

In contrast, Checchi and Peragine (2010) decompose total inequality into non-offensive and offensive components by utilising the standard between-group decomposition of inequality indices. Characteristics define groups—Roemer's types—and, hence, the between-group component can be interpreted as the *ex ante* measure of inequality of opportunity. If groups are built from their relative position in the effort distribution, inequality within groups will express an *ex post* measure. Another approach, by Lefranc, Pistolesi, and Trannoy (2008), is built on stochastic dominance of distribution given specific types. However, the latter is very demanding in terms of sample size.

Roemer's usual analysis of inequality of opportunity assumes a set of individuals expressed by $i \in \{1, \cdots, N\}$, where N is a large finite number. These individuals show a set of attributes $\{y_i, C_i, e_i\}$, where y stands for an outcome, C for a set of personal characteristics and e is the effort. Following Roemer again, there will be only one variable for outcomes and the effort will be a scalar. If the effort is assumed to be a continuous variable, the set C_i has J elements for each circumstance j and each element C_{ij} can take a finite number of values, x_j, the population can be partitioned into homogenous subgroups or types. This partition $\Pi = \{T_1, \cdots, T_K\}$ is exhaustive and exclusive, the joint distribution of outcomes and circumstances can be denoted as $\{y, C\}$ and the marginal distribution of outcomes is expressed by the vector $y = (y_1, y_N)$.

If $F^k(y)$ is the cumulative distribution function of outcome in type k and assuming two different types k and l, Roemer's strong criterion of equal opportunities can be defined as:

$$F^k(y) = F^l(y). \tag{1}$$

Therefore, if outcome distributions over types are identical in the case of equality of opportunity, measuring inequality of opportunity makes one check if both distributions are different and the degree of disparities, in the case they are not equal. This is the approach followed by Lefranc et al. (2008), who use a stochastic dominance method to compare conditional income distributions across types. Although this approach is very interesting, estimating those distribution functions requires a relevant number of observations in each type and the current sample sizes are not usually large enough. Using a small number of types is troubling because the results provide the blended effects of several circumstances.

Alternatively, a weaker criterion can be implemented to overcome this "sample size" problem. In this case, the equality of opportunity definition is based on the outcome means across types instead of cumulative distributions, and it can be expressed as:

$$\mu^k(y) = \mu^l(y), \tag{2}$$

where μ is the mean outcome of each type k.

Moving from Equation (1) to Equation (2) solves the problem of sample size and allows one to consider more types, which are more proper to describe reality. Besides, this change in measuring inequality of opportunity can relate to the discussion between *ex post* and *ex ante* definitions of inequality of opportunity.[2] In the former, it is

[2]More detailed discussions of both approaches can be read in Checchi and Peragine (2010), Ooghe, Schokkaert, and Van de Gaer (2007) or Carpantier and Sapata (2013); Ferreira and Gignoux (2011).

assumed that people who have made the same effort earn different outcomes whichever circumstances they have. To estimate this inequality, advantage differences should be aggregate between people in the same effort quantile across types. This procedure implies the existence of equality in all the quantiles and, therefore, in the whole distribution, as the Equation (1) shows.

In contrast, defining *ex ante* inequality requires us to see inequality between groups of people with the same characteristics. Hence, comparisons of individual efforts are not required, and only the opportunity set in each type must be assessed. This valuation, as Van de Gaer (1993) proposes, can be based on the mean value of the outcome, as Equation (2) assumes.

Based on this weaker criterion, Ferreira and Gignoux (2011) propose measuring inequality of opportunities based on a smoothed distribution of outcomes instead of the marginal distribution of advantages. This sort of smoothed distribution, denoted by $\{\mu^k\}$, was defined by Foster and Shneyerov (2000)[3] and applied to the estimation of inequality of opportunity by Checchi and Peragine (2010). The smoothed distribution $\{\mu^k\}$ is built for an outcome distribution y and a partition of type Π by replacing each individual outcome y_i^k with the type-specific mean $\mu^k(y)$.

Thus, an index of inequality of opportunity can be computed as:

$$\theta_a = I\left(\left\{\mu_i^k\right\}\right), \tag{3}$$

where I is one of the inequality indices in the literature.

From this absolute index θ_a, it is possible to build a relative version of the index:

$$\theta_r = \frac{I\left(\left\{\mu_i^k\right\}\right)}{I(y)}. \tag{4}$$

While θ_a measures the absolute level of inequality of opportunity (IOL), this level is compared to overall inequality in θ_r, so that this index is an inequality of opportunity ratio (IOR). Being told before, the index *I* can be any inequality index, but it would be desirable that the index *I* fulfilled the axiomatic properties for inequality indices (see Cowell, 1995).

As Shorrocks (1980) and Foster (1985) demonstrate, only an inequality measure that belongs to the generalized entropy (E_α) class can satisfy the four basic properties and, besides, being additively decomposable. Although the range of possible indices is lower, it is possible to obtain different inequality values depending on the parameter α for a specific smoothed distribution with a specific distribution of outcomes and a specific partition of circumstance types because each Entropy measure is sensitive to different parts of the distribution.

Nevertheless, as Ferreira and Gignoux (2011) show, a precise index can be selected by imposing the path-independent decomposability axiom, proposed by Foster and Shneyerov (2000). This axiom requires a standardized distribution, $\left\{v_i^k\right\}$, that is built by replacing each individual outcome y_i^k with $y_i^k \frac{\mu}{\mu^k}$, where μ is the overall mean of the distribution. This standardized distribution removes inequality between types as well as a smoothed distribution eliminates within-group inequality. Based on the standardized distribution, the path-independent decomposability axiom can be expressed as:

$$I\left(\left\{\mu_i^k\right\}\right) = I(y) - I\left(\left\{v_i^k\right\}\right), \tag{5}$$

Foster and Shneyerov (2000) show that the only inequality measure that satisfies this additional axiom, among those that use the arithmetic mean as the reference outcome and satisfy the Pigou–Dalton transfer axiom, is the mean

[3]This smoothed distribution lies on the inequality decomposition techniques proposed by Bourguignon et al. (2007), Cowell (1980), and Shorrocks (1980).

logarithmic deviation or the generalized entropy index when $\alpha = 0$. Therefore, Equations (3) and (4) can be respectively expressed as:

$$\theta_a = E_0\left(\{\mu_i^k\}\right), \tag{6}$$

and

$$\theta_r = \frac{E_0\left(\{\mu_i^k\}\right)}{E_0(y)}. \tag{7}$$

Therefore, from a sample that collects information about the joint distribution of outcome and circumstance variables $\{y, C\}$, it is possible to compute θ_a and θ_r by determining the between-group component in the standard decomposition of the mean logarithmic deviation by population subgroups.

Although this non-parametric technique can be applied in most of the analysis if there are few types in the partition Π, sometimes more variables—and, hence, more types—can be considered as personal circumstances that affect opportunities. The increase in types can hinder the computation of the index of inequality of opportunities. Thus, Ferreira and Gignoux (2011) propose following the regression approach by Bourguignon et al. (2007) that provides an efficient estimation. These authors state that Roemer's idea of inequality of outcomes established by circumstances, efforts and luck can be expressed by the following outcome model, $y = f(C, E, u)$. Besides, since circumstances are exogenous and efforts may be affected by circumstances, the model above can be written as:

$$y = f(C, E(C, v), u). \tag{8}$$

This model can be re-written to measure inequality of opportunities as $y = \phi(C, \varepsilon)$. If Equation (8) before is log-linearized, it can be expressed as $lny = C\varphi + \varepsilon$ and estimated by OLS where the parameters φ contain the information about the direct effects of circumstances on the outcome and their indirect effect through efforts. From the estimated coefficients a parametric estimate of the smoothed distribution can be built as:

$$\widehat{\mu}_i = \exp(C_i, \widehat{\varphi}), \tag{9}$$

where $\widehat{\varphi}$ are parameter OLS estimates and $\widehat{\mu}_i$ is a counterfactual outcome. Since equation (9) replaces individual outcomes with their predictions –and so, residuals are removed—the vector $\widehat{\mu}$ is parametrically analogue to the smoothed distribution $\{\mu^k\}$. The standardized distribution can be also parametrized by:

$$\widehat{v}_i = \exp\left(\overline{C}_i\widehat{\varphi} + \widehat{\varepsilon}_i\right), \tag{10}$$

$\widehat{\mu}$ is analogue to $\{\mu^k\}$ because it assigns the vector of average circumstances to every individual and keeps the within-type variation by means of $\widehat{\varepsilon}_i$. From these models, expressions of IOL and IOR can be re-written in parametric terms as:

$$\theta_a^p = E_0(\widetilde{\mu}), \tag{11}$$

and

$$\theta_a^p = \frac{E_0(\widetilde{\mu})}{E_0(y)}. \tag{12}$$

There are some caveats about this method of computing inequality of opportunity. Omitted circumstances represent the first problem. The vector C_i that appears in any dataset is a subset of the hypothetical vector C_i^* made of all observed and unobserved circumstances that cause people outcomes. Actual measures of inequality of opportunity require that the entire set of circumstances should be considered. Since it very unlikely to use this hypothetical vector C_i^*, the estimates of IOL and IOR should be interpreted as lower-bound estimates of inequality of opportunity.

Another caveat is related to the estimation of the partial effects of one (or a subset) circumstance, controlling for the others, through alternative counterfactual distributions. From the latter, one could compute some partial IOR or inequality shares of a specific circumstance. Nevertheless, Ferreira and Gignoux (2011) clearly remark that these shares are only significant as estimates of the (total) contribution of a given circumstance to inequality of opportunities.

2.1 | Data

The database used throughout this paper is the European Union Survey of Income and Living Conditions (hereafter, EUSILC) 2011 sample for Spain. The main goal sought by EUROSTAT when EU-SILC started was comparability and harmonization between European Union Member States. EU-SILC database comprises detailed information about each household member's incomes as well as indicators of material well-being, social exclusion, living conditions, labour and health status, educational attainment, and other social issues. It consists of four datasets: two with demographic data for households and individuals and the other two with more detailed personal and household information, respectively. Therefore, information from these datasets must be brought together to gather the required data for each analysis. Despite the wide range of available data, EU-SILC is the most fitting dataset to perform income distribution and poverty analyses in the European framework. Some special data modules are planned every year to allow one to develop some specific studies on social participation, accessibility to services, material hardship, well-being, financial exclusion, living conditions, or inter-generational transmission of poverty. The latter is the module used in this paper because it contains extended information related to people between 25 and 59 at the interview about parents' education, parents' labour status and occupation, financial situation, or difficulty to meet their ends. Despite these data availability to measure inequality of opportunities, some constraints of the database should be considered.[4]

Since the individual will be the unit of analysis of this paper and the number of adults is restricted in the special module, the samples comprise 16,427 observations for 2011, the last wave when the inter-generational transmission of poverty module was collected.[5]

When EU-SILC started, only Household Budget Surveys could be used with a regional (NUTS 2) dimension every decade and the European Community Household Panel (ECHP) was collected every year, but the spatial units were NUTS 1 or supra-regions. The geographical reference units in this paper are the Spanish NUTS 2 or regions (*Comunidades Autonomas*), the highest level of spatial disaggregation in microdata on income and living conditions in the database. There is a trade-off between NUTS 3 or provinces (smaller units), where sample sizes could be very large for statistical significance, and NUTS 1, with larger sample sizes but poorer information because of grouping regions with different characteristics and using regions (NUTS 2) can be a useful solution. Besides, NUTS 2 are appropriate for this analysis because of policy reasons. Social work policies, as well as education and health, are managed by regional governments and, therefore, the role of regions in the inequality of opportunities becomes very interesting.

[4]For instance, there is no information about the place of birth, excepting the country, so that self-selection or spatial sorting phenomena cannot be estimated. Besides, this module only appears in the cross-section database.

[5]Every EU-SILC wave includes a special module focused on a specific social issue. Inter-generational transmission of poverty has been measured only in 2005 and 2011 waves. However, since the methodology of collecting income data changed in 2013 and both modules do not consider the same variables, the analysis is constrained to the last available wave with this module.

3 | RESULTS

The outcome considered to measure the inequality of opportunities in this paper is income. Regarding this variable, the usual procedures in the literature are applied. First, net household income is computed by summing all the household members' individual incomes from every source as earnings, self-employment profits, capital incomes, pensions, and other social benefits (unemployment, illness, family, housing, etc.) and transfers from other private agents. Once the household total net disposable income is computed, differences in needs must be considered. These differences in needs come from differences in household size and type because not all households, even if they have the same size, need the same amount of money to cope with their normal lives. To compute the net equivalent income—that is, to consider differences in needs—the total household income is divided by the equivalent household size. In this case, the modified OECD equivalence scale[6] is used to compute the latter.

Regarding the circumstances vector (*C*), some variables related to the family background and the household financial stress during adults' childhood appear in the dataset. Among the family background, father's and mother's education are measured by both categorical variables running from no schooling (1) to higher education (4). Parents' labour status variables have been slightly modified from the original categorical variables and the current categories are "wage-earner," "self-employed," "unemployed" and "inactive." Besides, two variables related to father's and mother's education are included. Both variables are categorical with 1-digit ISCO-08 occupations. Finally, the household financial stress during childhood is measured by a categorical variable running from "very hard" (1) to "very good" (6). It is expected that parents' higher education, having parents who are employees or employers, good parents' occupations or having lived in a household with decent living conditions during childhood are variables that can help people to achieve a better future.

In addition to these "past" variables, another current variable is included to reflect gender to find some specific effects between men and women. All the variables in the model are usually considered in the literature about inequality of opportunity. Besides, the role of regions is included as a grouping variable to find some regional peculiarities, which can make fighting against non-ethic inequalities easier or harder. Space is not an out-of-control variable since people can move and choose their place of residence.[7] Nevertheless, it can become a factor that constrains the opportunity set of people or the extent of taking advantage of it. Considering space, measured by the region, into the variables that impact on inequality of opportunities is the main contribution of this paper in the literature of inequality of opportunities, where it has been an omitted variable.

Regional inequality differences in Spain stand out. This phenomenon is long-lasting because these disparities have been reported since the first surveys on inequality and poverty in the 1960s. The implementation of a modern public welfare system as well as the Spanish entry into the European Union, the then European Economic Community, enhanced the pace of convergence between Spanish regions in terms of well-being. However, the deep changes suffered by the Spanish economy at the end of the last century slowed down first and later stopped the convergence. Given their diverse economic specialization and labour markets, not all the regions took the same advantage of the remarkable boom observed in Spain before the economic crisis. Likewise, regional differences hampered the economic resilience of regions with lower well-being. Thus, those regions with more investment in R&D and specialized in the industrial sector have faced the economic shocks better. Besides these disparities in economic sectors, there are significant social and political differences. Spain's regional structure seems to be quasi-federal since the regions have strong political power: their regional parliaments can pass laws and some essential welfare policies like education, health, or social issues, with the exception of Social Security, are the responsibility of regional governments as well as some taxes that are totally or partially collected by regions. Therefore, regions can be a source of inequality because people's opportunity set can be constrained depending on the region where people live.

[6]This modified OECD equivalent scale assigns different weights to each household member: 1 to the first adult in the household, 0.5 to the other adults and 0.3 to children. Later, all the weights are summed to obtain the equivalent household size.

[7]Although the level of inner mobility in Spain has not been very important lately based on the official population data.

TABLE 2 Inequality of opportunities in Spain 2011 (for equivalent income)

Method	Absolute	Relative
Ferreira-Gignoux (with scale)	0.017894	0.095588
Ferreira-Gignoux (without scale	not defined	0.072116
Decomposition (Shapley method)	Variable	Value in percentage
Father's education	0.004186	23.39%
Mother's education	0.004112	22.98%
Father's occupation	0.004198	23.46%
Mother's occupation	0.002810	15.70%
Financial stress	0.002588	14.46%
Total	0.019664	100%

Source: Author's elaboration from Stata IOP package.

Table 1 reports the estimation of disparities in inequality mentioned before. Since the mean log deviation, Theil, or generalized entropy index with $\alpha = 0$, is computed, the property of additive decomposition can be applied. Thus, around 95% of the aggregate inequality observed in Spain, 0.2209, is due to intra-regional distribution and the remaining 5% comes from between regions income disparities.

Nevertheless, this decomposition does not measure the extent that regions limits the opportunity set. The indices presented in the former section, IOL and IOR, will be computed by following the regression approach proposed by Ferreira and Gignoux (2011) so that all the information contained in the dataset can be used.

The IOP Stata package (Wendelspiess Chávez Juárez & Soloaga, 2014) is implemented to estimate both indices. This package can extend the information included in them. Moving beyond the point estimated provided by the regression approach, a decomposition of inequality of opportunity can be interesting to understand why it exists. Among the decompositions available in the package, one based on the Shapley[8] value is the most adequate for continuous dependent variables or outcomes. This decomposition allows one to divide the total inequality of opportunity into several components by attributing the respective share to each circumstance. This package estimates the inequality for all possible permutations of the circumstance variables and, after that, the average marginal effect of each circumstance variable on the measure of inequality of opportunity is computed. Thus, this decomposition is order independent and, besides, there are no residuals and the different components sum up to the total value.

Table 2 reports a level of inequality of opportunity in Spain of around 10%. Namely, 10% of the income heterogeneity in Spain is due to people's circumstances that constrain personal goals. Even though people exert the same level of effort, some circumstances make more difficult to achieve the same level of income.

In the regression approach all the factors, excepting gender, are significant and have the expected signs. Thus, family background stands out as the more relevant variable to understand the inequality of opportunity. Among them, father's education explains almost a quarter of the inequality of opportunities. Mother's education also presents a relevant share, showing that mothers with higher levels of education improve children's outcome. It should be noted that the female access to higher education in Spain was lower than male access. Therefore, the usual combination of parents' education for the adults in the sample is “male education higher that female education”, so that higher education in mothers means a well-educated family and, hence, very significant opportunities that boost personal outcomes. Parents' occupations are also significant with a value in the father's occupation case very similar to parents' education level. Better parents' occupations increase income and make labour status non-significant.

[8]Sastre and Trannoy (2002) applied the Shapley value to the decomposition of income inequality. Besides, a full demonstration of applying this decomposition to any function can be found in Shorrocks (1999).

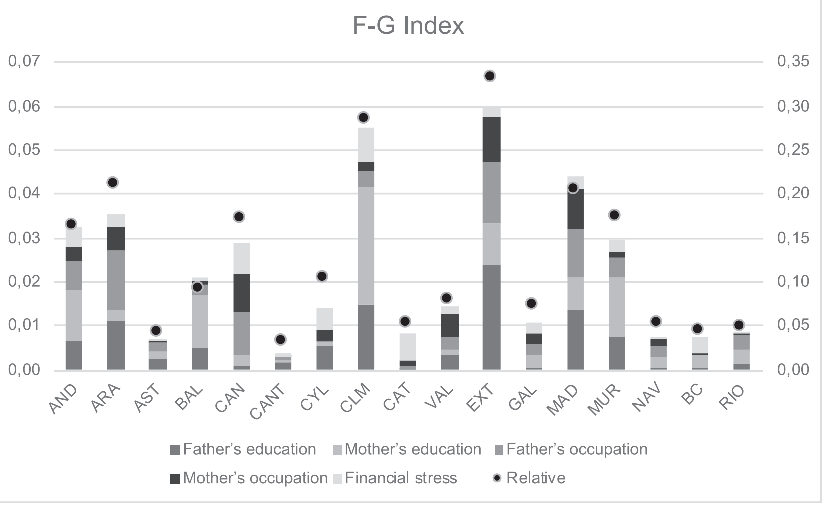

FIGURE 1 Regional distribution of inequality of opportunities and factor decompositions (for equivalent income)
Source: Author"s elaboration from Stata IOP package. Factor decompositions in bars follows the secondary (right) axis.

Table 2 also reports the role of financial stress suffered by individuals in the childhood, which is very related to income with the expected sign -better economic situation before is related to more income now, although it is the less important factor after controlling by the other.

Once inequality of opportunities is measured in Spain, it is important to include the spatial dimension by estimating it in every region as well as the relative importance of factors.[9] Figure 1 clearly shows significant disparities between regions. For instance, relative values range comes from 3% of income inequality in Cantabria to 33% in Extremadura. Besides, regions such as Catalonia, Navarre, or Basque Country, where poverty and living conditions have been historically low and better, respectively, seems to be territories with more opportunities.

Finally, there is a remarkable relationship between absolute and relative values and, therefore, one can state that those regions with higher levels of income inequality also show a higher level of inequality of opportunities. Therefore, regional characteristics appear to be more limiting in most unequal regions. Not only are current life outcomes very different, but future life outcomes also. Due to the database design, the analysis performed in this paper measures the current impact of past circumstances. However, given the current regional distribution of school dropout rates, investments in R&D, rankings in PISA surveys, unemployment rates and other variables that enhance the economic dynamism, regional convergence in income inequality between Spanish regions could be expected to slow down, and even revert to divergence.

Besides, there are no clear regional patterns when the role of factors is analysed, although parents' education is one of the most important variables in almost every region so that policies that improve the education system and enhance labour participation are required to break the vicious circle of inequality and poverty. Both factors have been pointed out in a recent survey by OECD (OECD, 2018) as the key issues to achieve and promote social mobility.

In sum, depending on the region they live, people can achieve different life outcomes in terms of income, in addition to living conditions in childhood.

Equivalent income is the output variable studied until now but considering net personal income instead could be interesting. At first sight, personal income could be regarded as the best variable. However, issues as assortative mating or wealth and bequests can make people with good environments in their childhood appear with zero or low

[9]Both are reported in Figure 1 by means of the main and the secondary axis, respectively.

TABLE 3 Inequality of opportunities in Spain 2011 (for personal net income)

Method	Absolute	Relative
Ferreira-Gignoux (with scale)	0.025308	0.274633
Ferreira-Gignoux (without scale	not defined	0.051565
Decomposition (Shapley method)	Variable	Value in percentage
Sex	0.015243	60.23%
Mother's laboral status	0.001021	4.03%
Father's occupation	0.004341	17.15%
Financial stress	0.004703	18.58%
Total	0.025308	100%

Source: Author's elaboration from Stata IOP package.

current earnings. Although equivalent income can avoid these problems, inequality in net incomes is computed to complement the analysis before (see Table 3).

Since it is estimated that more than a quarter of the personal income inequality can be explained by people's out-of-control circumstances, one can state that inequality of opportunities is outstandingly higher for net personal income than equivalent income. This result should be expected because this indicator omits significant facts as differences in household size and needs, income pooling by household members or economies of scale within the household.

In this case, gender stands out as the most important variable because it explains 60% of the inequality of opportunities. This result shows some the effects of the gender gap in wages and labour market or the level of assortative mating. Thus, women who have been brought up in a good family measured in terms of socioeconomic status keep their situation, despite lower personal earnings. Equivalent income adds and shares all the incomes in the household and, therefore, people with no income declare their actual living conditions. On the other side, issues such as father's occupation and the financial stress in childhood explain the rest of the observed inequality.

Besides, regarding the regional distribution, gender stands out as a very relevant factor related to inequality of opportunities in personal income (Figure 2). This outcome contrasts the regional profiles for inequality of opportunities in equivalent income (see Figure 1). This disparity could be due to the increasing weigh of gender when personal income is used, which overlaps the influence of the remaining variables.

Finally, an issue very much related to inequality of opportunities—even, both issues are considered equivalent—is the inter-generational transmission of poverty. That is, it is interesting to measure the extent of what poverty is inherited.

Table 4 reports that 7% of the current poverty risk could be explained by the out-of-control circumstances. After removing those variables without statistical significance, fathers' education, and occupation, as well as financial stress, in childhood arise as the explaining variables, where the latter stands out as the most important. Being poor while growing up explains a half of inequality of opportunities.

The regional distribution of inequality of opportunities measured by poverty risk gives one a very relevant picture. Regions such as Andalusia, Extremadura, Castile La Mancha, or Murcia, where poverty risk rates are usually high, stand out very markedly. Especially, almost a fifth of the current poverty risk can be attributed to the situation in the past. Besides, these regions present labour markets with very high unemployment rates (both general and youth) as well as high level of school dropouts. Therefore, it is possible to affirm that there are some regions that can be identified as actual poverty traps.

Table 5 reports regional values in its main diagonal. It shows the results of an Oaxaca-style decomposition that follow the same approach than the decomposition usually applied to wage discrimination in labour economics. When wages are estimated by a Mincerian equation—that is, a linear regression—wage differences can be expressed as a combination of those differences caused by disparities in characteristics and those that come from coefficients. In

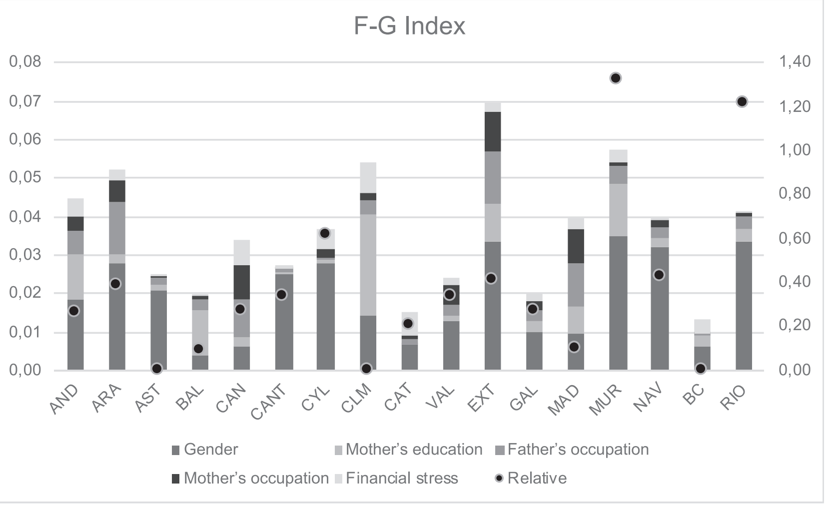

FIGURE 2 Regional distribution of inequality of opportunities and factor decompositions (for personal income)
Source: Author"s elaboration from Stata IOP package. Factor decompositions in bars follows the secondary (right) axis.

TABLE 4 Inequality of opportunities in Spain 2011 (for poverty risk)

Method	Absolute	
Dissimilarity index	0.11321	
Adapted DI	0.0759161	
Decomposition (Shapley method)	**Variable**	**Value in percentage**
Father's education	0.020968	18.52%
Father's occupation	0.036595	32.32%
Financial stress	0.055647	49.15%
Total	0.11321	100%

Source: Author's elaboration from Stata IOP package.

this case, a higher (lower) degree of inequality of opportunity can come from the presence of lower (higher) fathers' education, worse (better) fathers' occupation and worse (better) economic situations at childhood in a given region or from the regression coefficients of the model used to estimate poverty risk. The former will be inequality of opportunity caused by characteristics and the latter the one caused by coefficients. Based on this approach, Table 5 should be read according to the following rule: every cell reports the counterfactuals given by the characteristics in the row region and the coefficients in the column region. The application of this technique confirms the conclusion of the main diagonal values. Returns to characteristics in poorer regions make it difficult to overcome initial positions in out-of-control variables because almost all the regions show higher values when poorer region coefficients are used in the estimation.

The regional ranking obtained after computing inequality of opportunities and inter-generational transmission of poverty is rather in line with the ones that come from other variables like income inequality, poverty rates, GDP *per capita*, unemployment rates or even some indices of well-being or quality of life proposed in Spain respectively (Chasco, 2014; Jurado & Perez-Mayo, 2012).

TABLE 5 Regional distribution of inequality of opportunities and Oaxaca-style decompositions (for poverty risk)

		Coefficients																
		AND	ARA	AST	BAL	CAN	CANT	CYL	CLM	CAT	VAL	EXT	GAL	MAD	MUR	NAV	BC	RIO
Characteristics	AND	**0.1941**	0.0922	0.0256	0.0960	0.1033	0.1657	0.0234	0.1809	0.1240	0.1865	0.2118	0.0558	0.0763	0.1627	0.0174	0.0959	0.1736
	ARA	0.1236	**0.0459**	0.0201	0.0467	0.0951	0.1061	0.0162	0.1053	0.0665	0.1205	0.1667	0.0395	0.0563	0.1149	0.0115	0.0647	0.0745
	AST	0.1436	0.0485	**0.0206**	0.0488	0.1061	0.1126	0.0186	0.1212	0.0774	0.1427	0.1682	0.0453	0.0647	0.1335	0.0126	0.0663	0.0786
	BAL	0.1805	0.0801	0.0228	**0.0816**	0.1042	0.1432	0.0219	0.1647	0.1036	0.1665	0.2105	0.0546	0.0739	0.1571	0.0155	0.0836	0.1448
	CAN	0.1769	0.0736	0.0224	0.0747	**0.0999**	0.1414	0.0215	0.1600	0.1088	0.1699	0.1931	0.0518	0.0718	0.1514	0.0155	0.0820	0.1369
	CANT	0.1467	0.0500	0.0201	0.0522	0.0993	**0.1197**	0.0187	0.1233	0.0827	0.1430	0.1643	0.0463	0.0648	0.1291	0.0129	0.0713	0.0803
	CYL	0.1421	0.0433	0.0222	0.0434	0.0955	0.1357	**0.0182**	0.1200	0.0937	0.1449	0.1543	0.0420	0.0609	0.1254	0.0143	0.0804	0.0792
	CLM	0.1670	0.0807	0.0205	0.0815	0.1017	0.1228	0.0191	**0.1540**	0.0911	0.1557	0.1976	0.0504	0.0705	0.1631	0.0133	0.0714	0.1404
	CAT	0.1566	0.0718	0.0186	0.0719	0.1021	0.1056	0.0179	0.1428	**0.0745**	0.1415	0.1930	0.0499	0.0691	0.1527	0.0114	0.0626	0.1230
	VAL	0.1413	0.0476	0.0206	0.0484	0.0975	0.1192	0.0186	0.1195	0.0804	**0.1378**	0.1663	0.0433	0.0610	0.1239	0.0130	0.0704	0.0799
	EXT	0.1869	0.0685	0.0231	0.0685	0.0979	0.1503	0.0225	0.1671	0.1195	0.1851	**0.1825**	0.0526	0.0735	0.1580	0.0164	0.0865	0.1360
	GAL	0.1692	0.0590	0.0247	0.0598	0.0999	0.1658	0.0213	0.1522	0.1213	0.1677	0.1652	**0.0481**	0.0678	0.1353	0.0169	0.0973	0.1294
	MAD	0.1694	0.0752	0.0212	0.0758	0.1078	0.1290	0.0209	0.1532	0.0892	0.1537	0.1995	0.0533	**0.0722**	0.1517	0.0139	0.0758	0.1287
	MUR	0.1454	0.0439	0.0204	0.0416	0.1081	0.1240	0.0177	0.1235	0.0907	0.1535	0.1461	0.0441	0.0661	**0.1378**	0.0129	0.0734	0.0849
	NAV	0.1601	0.0530	0.0217	0.0542	0.1029	0.1312	0.0207	0.1377	0.0956	0.1569	0.1709	0.0491	0.0685	0.1372	**0.0145**	0.0766	0.0930
	BC	0.1577	0.0628	0.0218	0.0654	0.1117	0.1277	0.0199	0.1370	0.0869	0.1502	0.1913	0.0510	0.0710	0.1456	0.0137	**0.0749**	0.1014
	RIO	0.1512	0.0554	0.0199	0.0548	0.1030	0.1155	0.0188	0.1316	0.0777	0.1428	0.1738	0.0477	0.0667	0.1406	0.0122	0.0684	**0.0923**

4 | CONCLUDING REMARKS

Inequality of opportunities is analysed in this paper by including the spatial dimension in Spain. Given the apparent level of regional disparities in this country, where differences in poverty, unemployment or economic growth are reported throughout years, considering life chances—alternatively, different needs of personal effort depending on the starting point—can be interesting to give some insights of the future outcomes to be expected. The concentration of fewer opportunities with worse outcomes can end up in a vicious circle hard to break and this issue should be used to design the best policies aimed at achieving more social welfare and quality of life.

In this paper, the regression approach by Bourguignon et al. (2007) is combined with the methodology proposed by Ferreira and Gignoux (2011) to parametrically compute the indices of inequality of opportunity. These measures satisfy Roemer's definition as well as the usual axiomatic properties in the literature for inequality indices.

Based on the 2011 special module on intergenerational transmission of poverty of EU-SILC dataset, IOL and IOR indices are estimated and report that around 10% of the whole Spanish income inequality is related to differences in opportunities. Some regions stand out as places with more disadvantaged people, while others seem to be better environments to overcome initial bad endowments. This result is even more striking when inter-generational transmission of poverty is analysed. Poorer regions, besides of presenting worse levels of out-of-control characteristics, have worse returns to those characteristics, so that these regions are poverty traps where the risk of poverty inheritance is considerably higher than the average value in Spain. Thus, a strategy that seeks improving people's life chances in the future can tackle the regional effects that hinder personal development.

ORCID

Jesus Perez-Mayo https://orcid.org/0000-0001-6757-7857

REFERENCES

Akita, T. (2003). Decomposing regional income inequality in China and Indonesia using two-stage nested Theil decomposition method. *The Annals of Regional Science*, *37*(1), 55–77. https://doi.org/10.1007/s001680200107

Ayala, L., Jurado, A., & Pedraja, F. (2009). Inequality and welfare in intra-territorial income distribution. In J. Cuadrado-Roura (Ed.), *Regional policy, economic growth and convergence*. Berlin: Springer. https://doi.org/10.1007/978-3-642-02178-7_11

Bourguignon, F., Ferreira, F. H. G., & Menéndez, M. (2007). Inequality of opportunity in Brazil. *Review of Income and Wealth*, *53*(4), 585–618. https://doi.org/10.1111/j.1475-4991.2007.00247.x

Carpantier, J. F., & Sapata, C. (2013). An ex-post view of inequality of opportunity in France and its regions. *Journal of Labor Research*, *34*(3), 281–311. https://doi.org/10.1007/s12122-013-9161-5

Chasco, C. (2014). Objective index of quality of life in Spain. In A. C. Michalos (Ed.), *Encyclopaedia of quality of life and well-being research* (pp. 4432–4436). Dordrecht: Springer.

Checchi, D., & Peragine, V. (2010). Inequality of opportunity in Italy. *Journal of Economic Inequality*, *8*(4), 429–450. https://doi.org/10.1007/s10888-009-9118-3

Cowell, F. A. (1980). On the structure of additive inequality measures. *Review of Economic Studies*, *47*(3), 521–531. https://doi.org/10.2307/2297303

Cowell, F. A. (1995). *Measuring inequality* (2nd ed.). Hemel Hempstead: Prentice Hall/Harvester Wheatsheaf.

Ferreira, F. H., & Gignoux, J. (2011). The measurement of inequality of opportunity: Theory and an application to Latin America. *The Review of Income and Wealth*, *57*(4), 622–657. https://doi.org/10.1111/j.1475-4991.2011.00467.x

Foster, J., & Shneyerov, A. (2000). Path independent inequality measures. *Journal of Economic Theory*, *91*(2), 199–222. https://doi.org/10.1006/jeth.1999.2565

Foster, J. E. (1985). Inequality measurement. In H. Peyton Young (Ed.), *Fair allocation* (pp. 31–68). Providence, RI: American Mathematical Society.

Jurado, A., & Perez-Mayo, J. (2010). Dimensión territorial de la pobreza en España. *Revista Española del Tercer Sector*, *16*, 43–66.

Jurado, A., & Perez-Mayo, J. (2012). Construction and evolution of a multidimensional well-being index for the Spanish regions. *Social Indicators Research*, *107*(2), 259–279. https://doi.org/10.1007/s11205-011-9835-4

Lefranc, A., Pistolesi, N., & Trannoy, A. (2008). Inequality of opportunities vs. inequality of outcomes: Are Western societies all alike? *Review of Income and Wealth*, *54*(4), 513–546. https://doi.org/10.1111/j.1475-4991.2008.00289.x

Márquez, M. A., Lasarte-Navamuel, E., & Lufin, M. (2016). Isolating neighborhood components of regional inequality: Illustration for the Spanish case. *International Regional Science Review*, *41*(5), 483–509. https://doi.org/10.1177/0160017616665670

Márquez, M. A., Lasarte-Navamuel, E., & Lufin, M. (2017). The role of neighborhood in the analysis of spatial economic inequality. *Social Indicators Research*, Online First, 1–29. https://doi.org/10.1007/s11205-017-1814-y

OECD (2018). *A broken social elevator? How to promote social mobility*. Paris: OECD Publishing. https://doi.org/10.1787/9789264301085-en

Ooghe, E., Schokkaert, E., & Van de Gaer, D. (2007). Equality of opportunity versus equality of opportunity sets. *Social Choice and Welfare*, *28*, 209–230. https://doi.org/10.1007/s00355-006-0165-4

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

Sastre, M., & Trannoy, A. (2002). Shapley inequality decomposition by factor components: Some methodological issues. *Journal of Economics*, *77*, 51–89. https://doi.org/10.1007/BF03052500

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica*, *48*(3), 613–625. https://doi.org/10.2307/1913126

Shorrocks, A. F. (1999). Decomposition procedures for distributional analysis: A unified framework based on the Shapley value. *The Journal of Economic Inequality*, *11*(1), 99–126.

Silber, J. (1989). Factor components, population subgroups and the computation of the Gini index of inequality. *Review of Economics and Statistics*, *71*, 107–115. https://doi.org/10.2307/1928057

Van de Gaer, D. (1993). Equality of opportunity and investment in human capital. Ph.D. Dissertation, Catholic University of Louvain.

Wendelspiess Chávez Juárez, F., & Soloaga, I. (2014). Iop: Estimating ex-ante inequality of opportunity. *Stata Journal*, *14*(4), 830–846. https://doi.org/10.1177/1536867X1401400408

How to cite this article: Perez-Mayo J. Inequality of opportunity, a matter of space?. *Reg Sci Policy Pract*. 2019;11:71–87. https://doi.org/10.1111/rsp3.12181

APPENDIX

TABLE A1 Statistical information about Spanish regions

	Average income per capita	Per capita GDP	Unemployment general rate	Youth unemployment rate	Female activity rate	Firms in high-tech sectors	Industrial output index	Average annual wages	School dropout
	2016	2017	2017	2017	2017	2016	2017	2015	2017
Total	10,708	24,999	17.22	38.57	53.24	4888	97.039	23,106.30	18.3
1: Andalucia	8,398	18,470	25.51	49.05	50.60	438	108.057	21,381.03	23.5
2: Aragon	11,649	27,403	11.65	30.49	52.78	205	89.991	22,327.18	16.4
3: Asturias	12,060	22,046	13.71	36.79	46.59	109	107.105	22,826.06	14.8
4: Balearic Is.	12,222	25,772	12.43	30.79	58.14	55	93.546	21,395.42	26.5
5: Canary Is.	8,702	20,425	23.46	45.30	55.96	57	99.906	19,856.61	17.5
6: Cantabria	10,670	22,513	13.56	35.58	50.97	63	99.036	21,856.81	8.9
7: Castilla-Leon	10,815	23,555	14.08	36.24	49.03	192	89.080	21,296.10	16.7
8: Castilla-La Mancha	8,731	19,681	20.77	44.40	51.32	106	95.554	20,670.55	22.1
9: Catalonia	12,660	29,936	13.41	30.39	56.89	1301	97.705	24,321.57	17.0
10: Valencia	9,265	22,055	18.17	41.65	53.14	547	93.925	20,935.41	20.3
11: Extremadura	8,674	17,262	26.22	48.58	47.73	44	95.913	19,564.49	19.2
12: Galicia	10,439	22,497	15.67	34.48	49.11	282	98.331	20,624.55	14.9
13: Madrid	12,647	33,809	13.34	34.66	58.05	927	97.046	26,448.04	13.9
14: Murcia	8,273	20,585	18.03	39.47	52.39	145	107.577	20,928.98	23.1
15: Navarre	13,408	30,914	10.24	31.24	53.86	160	90.830	24,863.98	11.3
16: Basque Country	14,345	33,088	11.31	27.26	50.98	676	93.319	27,571.31	7.0
17: Rioja (La)	11,589	26,044	12.00	29.57	53.46	44	86.161	21,756.79	12.9

Source: Spanish National Statistical Institute.

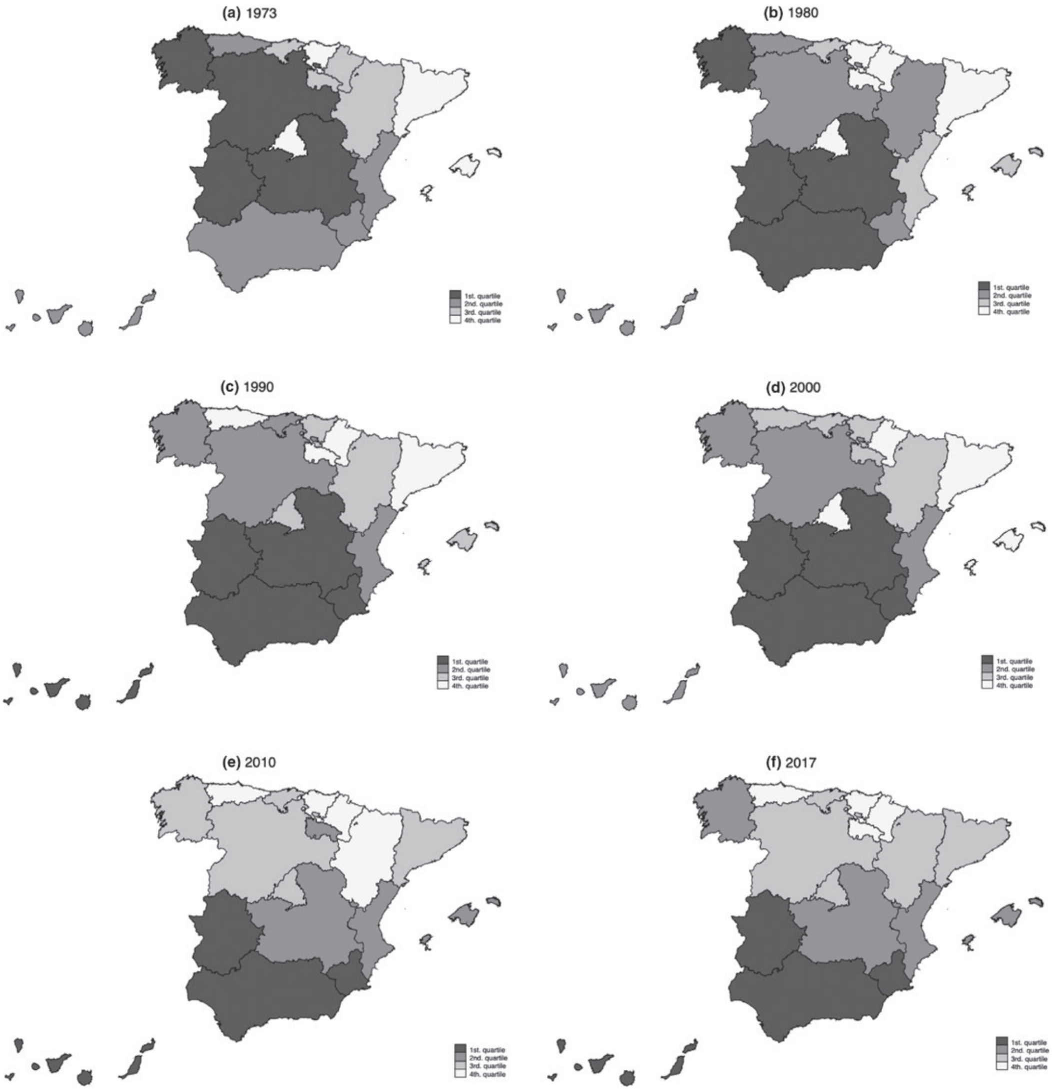

FIGURE A1 Regional distribution of poverty rates in Spain 1973-2017 (% of thenational rate)
[i]The map has been modified for sake of readability. The actual situation of Canary Islands is further away in the South. The maps are quantile maps to ease the comparison between periods.
Source: Jurado and Perez-Mayo (2010) and INE (Spanish National Statistical Institute)

DOI: 10.1111/rsp3.12181

Resumen. La desigualdad se ha convertido en un problema muy importante en los países desarrollados en las últimas décadas. Entre las diferentes definiciones de desigualdad, la desigualdad de oportunidades se destaca como uno de los aspectos principales. ¿En qué medida las circunstancias personales que no se pueden controlar son la raíz de resultados de desigualdad? El espacio, definido como el lugar donde viven las personas, puede limitar el abanico de oportunidades de las personas o, incluso, de personas con las mismas circunstancias personales, excepto la región, y el mismo abanico de oportunidades pueden ofrecer resultados muy diferentes dependiendo de la región. Esta cuestión se mide en España mediante el uso de microdatos de EU-SILC 2011.

抄録: この数十年間で、不平等は先進国における非常に大きな問題となった。不平等の定義は様々であるが、その中でも機会の不平等は主要な問題のひとつとして際立っている。制御することのできない個人の環境が、どのくらい不平等な結果の原因となっているのかは疑問である。空間を人が生活するところと定義すると、空間は、人が、または地域を除いて境遇が同じ人々までもが、向き合う一連の機会を制約することができ、同じ一連の機会から、地域によって大きく異なる結果を得ることができる。2011年のEUの所得・生活状況調査のマイクロデータを使用して、この問題をスペインの事例において測定する。