Working Paper Series

Opportunity advantage: a new approach to comparing income distributions

Carmen Herrero
Antonio Villar

ECINEQ WP 2020 - 516

ECINEQ 2020 - 516
January 2020

www.ecineq.org

Opportunity advantage: a new approach to comparing income distributions [‡]

Carmen Herrero
University of Alicante & Ivie

Antonio Villar
Pablo de Olavide University & Ivie

Abstract

This paper proposes a new approach to the comparison of income distributions based on the notion of opportunity advantage. This is a measure of how likely is that a representative individual of a society gets a higher income than that of another. Opportunity advantage can be regarded as the evaluation of the income opportunities that a society offers to an individual, relative to other societies, from the 'veil of ignorance'? viewpoint. We show that this notion can be given a precise formalization which results in a complete, transitive and cardinal evaluation of income opportunity. The evaluation so obtained describes the willingness to stay in a society relative to move somewhere else. We provide an empirical application to the analysis of the economic recovery in Spain and its regions. The results show that this criterion offers new insights on the impact of the crisis on households.

Keywords: income distribution; opportunity advantage; ranking distributions: veil of ignorance.

JEL Classification: D31, D63, I30.

[‡]The first author acknowledges financial support from the Spanish Ministerio de Econom'a, Industria y Competitividad under projects PROMETEO 2109/037, ECO2015-65820P, and IMeRA. The second author recognizes financial support from the Spanish Ministerio de Econom'a, Industria y Competitividad under project ECO2015-65408-R (MINECO/FEDER).

1 Introduction

Ranking income distributions is one of the most basic tools of welfare economics and probably one of the most serious steps to go *beyond the GDP* when assessing the economic situation of different societies. This analytical exercise requires having a criterion that permits one comparing income distributions regarding societies that may have different population sizes and different mean incomes. As alternative criteria may yield diverse rankings, the specific value judgments assumed are key for the evaluation.

Let us motivate the discussion by considering a simple example involving the income distributions of two societies, $\boldsymbol{y}_A, \boldsymbol{y}_B$, with ten members each:

$$\mathbf{y}_A = (1,2,2,2,2,2,3,3,5,5), \quad \mathbf{y}_B = (1,1,1,1,3,3,3,4,5,5)$$

How should we rank those two distributions?

On the one hand, if we are only concerned about per capita values (a GDP approach, so to speak), then both distributions are indifferent because they have the same mean income (2.7). On the other hand, if income inequality enters the evaluation, $\boldsymbol{y}_A$ is to be regarded as better than $\boldsymbol{y}_B$. The reason is that distribution $\boldsymbol{y}_B$ obtains from distribution $\boldsymbol{y}_A$ by a series of regressive transfers. That is, the first distribution Lorenz dominates the second one and, therefore, any social evaluation function that satisfies the standard properties of monotonicity, anonymity and quasi-concavity will rank distribution $\boldsymbol{y}_A$ above $\boldsymbol{y}_B$ (Roschild & Stiglitz, 1973, Dasgupta, Sen & Starret, 1973). This property also applies to distributions with different means, under generalized Lorenz dominance (Shorrocks, 1983).[1]

We propose here to compare income distributions from a different perspective, that takes into account the spread of incomes but focusses on the *opportunities* that those distributions offer, rather than on their *fairness.* A positive approach rather than a normative one. This approach can be formulated in the following terms. Suppose you leave your country of origin and are given the option of choosing between those two societies, *A* and *B*, without knowing the position you will end up in the social ladder (an instance of the "veil of ignorance"). Which one would you choose?

[1] The quasi-concavity of the social evaluation function entails the well-known Dalton principle of transfers, which requires that a transfer from a rich to a poor agent, which does not modify their relative ranking, will enhance social welfare. This normative requirement is broadly accepted as one of the most basic ethical criteria for the evaluation of income distributions, even though it is not free of criticisms. In particular regarding the lack of concern for the impact of the transfers in the local environment of the two agents involved and its close relationship to the much more arguable property of additive separability.

A casual test performed with our students shows that while an outstanding majority rank $\boldsymbol{y}_A$ over $\boldsymbol{y}_B$ when required to evaluate income distributions from an ethical perspective, many of them reversed their ranking when asked which society they would prefer to live in. Why is that so? The common answer was that society *B* offers better chances of getting higher incomes, hence better opportunity. Some of them specified that it was more likely getting more than the average income, as 60% of the population in *A* is below the mean income whereas 60% of the population is above in *B*.

We present in this paper an approach to ranking (actually *valuing*) income distributions based on the idea of *opportunity advantage*, which can be regarded as an extension of the former reasoning about having more chances to get a higher income under the veil of ignorance. When comparing two income distributions, $\boldsymbol{y}_A$ and $\boldsymbol{y}_B$, the income opportunity of $\boldsymbol{y}_A$ with respect to $\boldsymbol{y}_B$ is a function of the probability that a representative agent of society *A* obtains a higher income than a representative agent of society *B*. This simple notion permits one not only rank $\boldsymbol{y}_A$ and $\boldsymbol{y}_B$ but also to have an estimate on how much better one income distribution relative to the other. This criterion for pair-wise comparisons can be consistently extended in order to get a general evaluation formula with suitable ethical and operational properties. The resulting evaluation describes the *relative* desirability of each income distribution regarding the societies with which it is confronted. Interestingly enough, the formula is computationally easy and has a clear intuitive meaning. Note that this change of focus implies that the discussion will shift from income distribution vectors to vectors of population shares by income levels.

The paper is organised as follows. Section 2 presents the notion of opportunity advantage and the associated formal model. The scenario refers to the evaluation of income opportunities of different societies made of homogeneous agents. Section 3 discusses the implications of moving from this scenario to that of a single society whose heterogeneous members can be classified into different population subgroups according to some non-income characteristics. This permits one to approach equality of opportunity from a different perspective (that we call *opportunity bias*). Section 4 contains an empirical application in which we measure the extent of the Spanish recovery after the crisis, by comparing the income opportunities for the Spanish regions in 2016 relative to 2007. A few final comments are gathered in Section 5.

2 The opportunity advantage approach to income distributions

We propose here a criterion to evaluate income distributions that captures the relative opportunities provided by the different societies to their citizens. We call this evaluation criterion the ***opportunity advantage***. It is based in the comparison of the likelihood of getting higher incomes, by comparing the distributions of the populations into a common set of income levels or intervals (e.g. a discretized version of the density on a common support).

The key element of this evaluation protocol refers to the advantage for an individual of belonging to a given society with respect to belonging to some other. That is, the opportunity advantage of the income distribution of a given society derives from the comparison of how good is to be in society *i* relative to be somewhere else, and how good is to be somewhere else with respect to be in *i*. The "how good" is measured by the average probabilities of getting better or worse income levels. We can think of this criterion as a "willingness to stay" measure.

Let us formalise these ideas.

Consider a collection of societies $M = \{1, 2, ..., m\}$ whose income distributions are to be compared. Income distributions are given by vectors $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \dots, y_{in(i)})$, where $n(i)$ is the cardinal of society *i*. The members of this society can be classified into *G* different *income brackets*, $g = 1, 2, ..., G$. Let $\boldsymbol{a}_i = (a_{i1}, a_{i2}, \dots, a_{iG})$ denote the vector of society *i*'s population shares into those income brackets. That is, a_{ig}, $i = 1, 2, \dots, m, g = 1, 2, \dots, G$, is the share of individuals of society *i* in income bracket *g*. By construction, $\sum_{g=1}^{G} a_{ig} = 1$.

The key principle for the comparison between societies refers to the probability that a representative agent belongs to a higher income bracket. We shall, therefore, change the focus from income distributions, $\boldsymbol{y}_i$, to distributions of people into income intervals, $\boldsymbol{a}_i$. It is implicit in our formulation that income is the only trait that distinguishes individuals (we take up the case or heterogenous populations later on).

Remark: *Note that the formulation permits using scalar income values, rather than income intervals, by selecting the common support of the income distribution vectors.*

Yet we find it more intuitive to think of income brackets and the associated social classes that those define.

Let p_{ik} denote the probability that an individual from society *i* belongs to a higher income group than an individual from society *k*. Such a probability can be calculated as follows:

$$p_{ik} = a_{i1}(a_{k2} + \cdots + a_{kG}) + a_{i2}(a_{k3} + \cdots + a_{kG}) + \cdots + a_{i(G-1)}a_{kG}$$

Similarly, we define $e_{ik} = e_{ki}$ as the probability that and individual of society *i* belongs to the same income bracket that an individual of society *k*. By definition, we have: p_{ik}+p_{ki}+$e_{ik} = 1$. Given the symmetry of the probability of a tie in binary comparisons, we shall split equally that probability between both societies, so that we can think of the former equality as follows: $\left(p_{ik} + \frac{e_{ik}}{2}\right) + \left(p_{ki} + \frac{e_{ki}}{2}\right) = 1$. Let us simplify notation by letting $q_{ik} \equiv p_{ik} + \frac{e_{ik}}{2}$.

We want to evaluate the opportunities that different income distributions offer, that is, finding some values that inform about how does an income distribution fare with respect to the others. Let us start by supposing that all income distributions are tentatively regarded as equally worthy, so that each one receives an initial score $s_i^0 = 1$. What would be the net benefit for an individual to join society *i* rather than some other in that case? The answer is given by the following expression:

$$NB_i^0 = \frac{1}{m-1}\left(\sum_{k \neq i} q_{ik} - \sum_{k \neq i} q_{ki}\right) \quad (1)$$

That is, the difference between the average probability that an individual of *i* belongs to a higher income bracket than an individual of some other society and the average probability that an individual of *i* belongs to a lower income bracket than an individual of some other society. Clearly, it is convenient entering society *i*, rather than some other, when $NB_i^0 > 0$.[2] If $NB_i^0 = 0$ both alternatives are indifferent, in which case the value $s_i^0 = 1$ can be regarded as adequate. NB_i^0 can thus be interpreted as the difference between the proper value of the income distribution of this society, s_i^1 say, and the one initially assigned, that is, $NB_i^0 = s_i^1 - 1$. We can, therefore, rewrite equation (1) as follows:

[2] If we think of s_i^0 as a fee required to enter society *i*, NB_i^0 describes the difference between the willingness to pay in order to enter society *i*, s_i^1, and the entering fee.

$$s_i^1 = 1 + \frac{1}{m-1}\left(\sum_{k\neq i} q_{ik} - \sum_{k\neq i} q_{ki}\right) \qquad (1')$$

That is, the proper value of the income distribution of society *i* corresponds to a change of the initial value given by the average difference of the advantage of being in that society with respect to being somewhere else.

Note that $s_i^1 = s_i^0 = 1$ if and only if $\sum_{k\neq i} q_{ik} - \sum_{k\neq i} q_{ki}$. When this is not the case, we shall find a new vector of evaluations derived from equation (1'), $\boldsymbol{s}^1 = (s_1^1, s_2^1, \dots, s_m^1)$, and will have to re-calculate the net benefits of joining each society, accordingly. That is,

$$NB_i^1 = \frac{1}{m-1}\left(\sum\nolimits_{k\neq i} q_{ik}\, s_k^1 - \sum\nolimits_{k\neq i} q_{ki}\, s_i^1\right)$$

This corresponds to adjusting equation (1) to the different evaluations derived from the previous step. This re-evaluation of the net benefits implies a still newer vector of evaluations given by:

$$s_i^2 = s_i^1 + \frac{1}{m-1}\left(\sum\nolimits_{k\neq i} q_{ik}\, s_k^1 - \sum\nolimits_{k\neq i} q_{ki}\, s_i^1\right)$$

This obviously opens a third round of updates in the evaluation of income distributions, which in turn will require a fourth round, and so on.[3]

By repeating this procedure indefinitely, we end up by finding a stable evaluation vector, $\boldsymbol{v} = (v_1, v_1, \dots, v_m)$, with $v_i = s_i^{\infty}$, $i = 1, 2, \dots, m$, which provides a measure of the *opportunity advantage* of being in society *i* with respect to being in some other one. The i*th* entry of this vector will be given by:

$$v_i = \frac{\sum_{k\neq i} q_{ik} v_k}{\sum_{k\neq i} q_{ki}}\ ,\ \ i = 1,2, \dots, m \qquad (3)$$

To check that this sequence of evaluations actually converges to the non-negative vector $\boldsymbol{v}$ in equation (3), let us define an *m* square matrix $\boldsymbol{Q}$ whose (i, j) off-diagonal element is $q_{ij}, i \neq j, i, j = 1,2, \dots, m$, and each element in the diagonal is given by $D_i = (m-1) - \sum_{k\neq i} q_{ji}, i = 1, 2, \dots, m$. By letting $\boldsymbol{s}^t = (s_1^t, s_2^t, \dots, s_m^t)$ be a column vector, the former process can be described as follows:

[3] Note the similarity of this notion with that of page rank that applies Google to order the web pages. See also a very close application of this principle to the evaluation of scientific influence in Palacios-Huerta & Volij (2004) and the notion of worth and balanced worth in Herrero & Villar (2013, 2018).

$$\boldsymbol{s}^0 = (1, 1, \dots, 1) \;;\; \boldsymbol{s}^1 = \frac{1}{m-1}\boldsymbol{Q}\boldsymbol{s}^0 \;;\; \dots ;\; \boldsymbol{s}^t = \frac{1}{m-1}\boldsymbol{Q}\boldsymbol{s}^{t-1} \;;\; \dots$$

As $\boldsymbol{Q}$ is a Perron matrix (non-negative entries), all whose columns add up to $(m-1)$, we know that this process converges to a point $\boldsymbol{v} = \lim_{t\to\infty} \boldsymbol{s}^t \in R_+^m$ (e.g. Berman & Plemmons, 1994). This vector is generically unique (after normalisation, as by construction it has one degree of freedom) and strictly positive. We shall refer to $\boldsymbol{v}$ as the ***opportunity advantage*** vector.

The opportunity advantage of society *i* is thus a measure of the income opportunity that this society offers to its members, relative to the other societies with which it is compared. It corresponds to the weighted average of the probabilities that a member of *i* achieves a higher income than a member of other societies, with weights given by their corresponding opportunity advantages, scaled down by the average probability that a member of *i* achieves a lower income than a member of the rest. This evaluation protocol provides a cardinal measure of the relative goodness of those income distributions from an opportunity viewpoint.

It is easy to check that whenever the income distributions of two societies, *i*, *k*, coincide, then $v_i^* = v_k^*$. Moreover, $v_i^* = 0$ if and only if $q_{ik} = 0$ for all $k \neq i$. Finally, let us mention that this evaluation criterion satisfies *monotonicity* in the following sense: If the population of a society shifts towards higher income levels, whereas the rest remain unaltered, then the new evaluation will be higher for this society. As a consequence, this evaluation can be regarded as a transitive, complete and cardinal extension of the first order stochastic dominance criterion (on a different space of that of the conventional income distributions, though).

3 Heterogeneous populations

The model presented in the former section was designed to provide an evaluation of the income opportunities *between* different societies with homogeneous agents. Income differences were the only aspect considered. The comparison was based on the chances of getting higher income for a newcomer to society under the veil of ignorance. We now consider a change of focus by addressing the problem of analysing income opportunities *within* a given society with heterogeneous population. The idea is that people in this society can be classified into a number of population subgroups, according to some characteristics different from income (e.g. age, gender, level of studies, region of residence). In this context

we can apply the opportunity advantage approach to evaluate the impact of those characteristics on income distributions by treating each population subgroup as a different society. Evaluating income opportunity in this context permits assessing the differences that exist in society due to those non-income characteristics that define the population subgroups. We shall refer to the application of opportunity advantage to the analysis of population subgroups within a society as ***opportunity bias***.

Note that opportunity bias still involves a positive approach to the evaluation of opportunity differences, even though it can be given a normative content depending on the context. That is, the existence of bias among population subgroups describes their opportunity differences, without any intrinsic fairness component. Opportunity bias may be regarded as a measure of unfairness in some cases, as when we compare income distributions between men and women with similar characteristics other than gender, say. Yet in other opportunity bias simply becomes a descriptive indicator of the different opportunities derived from some conditioning variable, as when we compare earning distributions of different employed workers depending on their university degrees (which might be regarded as a guide to decide on what studies to follow).

Discussing the normative content of opportunity bias opens the question of how this notion relates to equality of opportunity. Equality of opportunity is one of the most relevant approaches to distributive justice, characterised by a wide spectrum of views (see Fleurbaey 2008, Roemer & Trannoy 2015, 2016 for a discussion and detailed references). The bottom line behind the equality of opportunity principle is that external circumstances are to be taken into account when comparing outcomes, which is usually associated with the idea that people who are relatively disadvantaged deserve some kind of recognition or compensation. And, complementarily, that we should not be concerned for those outcome differences among people with the same circumstances, as long they derive from people's autonomous choices (e.g. differential effort).

The best-known version of the inequality of opportunity principle among economists is probably that based on the work of Roemer (1996, 1998). According to his approach an outcome distribution can be regarded as the result of two different effects: *effort* and *opportunity*. Effort has to do with responsibility and involves people's autonomous choices on a common "playing field". Opportunity refers to the agents' external circumstances, which may include genes, race, gender, family socioeconomic and cultural background, and other aspects for which agents cannot be held responsible. A fair society should care for the agents' differences in opportunity but not for those differences derived from autonomous personal

decisions.

Evaluating the degree of inequality of opportunity in a society from this perspective thus involves a double partition of their members. On the one hand, there are the *types*, which gather agents who share the same circumstances. On the other hand, there are the *effort groups*, which correspond to those population subgroups that exert a similar degree of effort. Within this framework, the outcome distribution of those agents of the same type can be regarded as determined by their effort decisions. In other words, agents of the same type have the same opportunity and all outcome differences within a type correspond to differences in people's effort decisions, which are ethically irrelevant. The relevant inequality refers, therefore, to that between effort groups, which implies that we can measure inequality of opportunity by recurring to some inequality index applied to those effort groups (e.g. Peragine 2002, 2004, Ruiz-Castillo 2003, Villar 2005).

The idea that outcome differences due to external circumstances are unfair is a powerful one. Opportunity bias can be used to measure the degree of fairness when population subgroups are defined, *à la* Roemer, by those people who share similar circumstances. The differences between the income distributions of those population subgroups reflect the different opportunities faced by individuals depending on their external circumstances. The implicit assumption in our approach is that the income distribution of each type can be regarded as a sufficient estimate of its opportunity, an idea very much in line with Sen's capability approach (Sen, 1985). Consequently, a fair society is one in which the opportunity advantages of the different population subgroups are all alike, that is, when the chances open to individuals do not depend on their external circumstances. This value judgement can be regarded as an instance of a basic non-discrimination principle by which we try to ensure that any new member of society will have access to its average chances, no matter which social group she ends up in (the application of a leximin social evaluation function on the space of opportunities).

Opportunity bias can be regarded as an approach rougher than the conventional equality opportunity one, regarding the evaluation of the unfairness due to differential circumstances, as there is no place for effort considerations in its simpler structure. Yet effort is a non-observable and type-dependent variable, which calls for the design of an index that permits comparing effort levels for agents of different types. The resulting measurement of inequality of opportunity is very sensitive to the way of defining the effort groups, always a challenging modelling choice.[4]

[4] To solve this problem Roemer (1998) assumes that effort is a single-valued variable that is positively correlated with outcome. So, even though the effort distribution is a characteristic of the

4 An empirical application: The extent of the economic recovery in Spain

The purpose of this empirical exercise is to analyze the degree of success of the economic recovery in Spain and its regions, by comparing the income opportunities faced by households in 2016 relative to those in 2007. The year 2007 is the one in which the crisis started. 2016 is the year in which the average income of the Spanish households, duly scaled by size and composition, achieves the level it had in 2008, the pick since the crisis started (see Figure 1 below). The basic question we shall address here is to what extent has the Spanish population recovered in 2016 the income opportunities they had in 2007. Needless to say, that depends not only on the per capita level of income but also on the income distribution.

The application of the opportunity advantage approach to the analysis developed here can be considered as an extension of the following extremely simple question: How is the probability that the income of a Spanish citizen in 2016 be equal to or greater than the per capita income in 2007? That difference in probability will measure the shifting of the population shares among social groups. We can interpret this from the "veil of ignorance" approach to assess income distributions in terms of the opportunities they offer to newcomers to society. From this perspective, a higher probability of obtaining at least the mean income corresponds to a situation with better income opportunities.

The idea of measuring the probability of obtaining at least the 2007 per capita income is equivalent to considering that society is composed of two large social groups: those who obtain income above the 2007 average and those who obtain income below it. Although this cutting point is very intuitive, it implies an arbitrary division of the society into just two social groups. Our analysis really consists of an extension of this idea that can be formulated as follows. We consider that each society is divided into a certain number of social groups, defined in terms of 2007 income brackets, and we calculate the probability that in 2016 a representative agent belongs to a "higher" income bracket than one in 2007.

Let us recall that Spain is a highly decentralised country in which half of the public expenditure is managed regionally and many areas of public authority such

type, we can take the quantiles of the effort distribution within types to compare the degrees of effort (i.e. two individuals of different types exert a comparable degree of effort if their outcomes belong to the same quantile of the outcome distribution of their corresponding types).

as health, education and other economic activities are also devolved to the regions. It is also well-known that regional differences in Spain are large. It is therefore interesting to analyse the impact of the crisis on income opportunities in the different regions. This permits one dealing with two different aspects of the problem. The direct one, regarding the change in opportunities between 2007 and 2016, and the opportunity bias aspect, by analysing income opportunities conditional on the region of residence.

The data used for our analysis come from the 2017 *Encuesta de Presupuestos Familiares* (Family Expenditure Survey, FES), elaborated by the Spanish Instituto Nacional de Estadística. The FES provides annual information on the nature and destination of consumption expenditures, as well as on various characteristics related to the living conditions of households. Even though the survey focuses on consumption expenditure, it also contains information on revenues that can be obtained by aggregating the income of the different household members. The sample size is approximately 24,000 homes per year.[5]

There are alternative ways of implementing the comparison of income opportunities in Spain and its regions between 2007 and 2016. In particular, we have to determine: (1) The precise notion of income (revenues or expenses); (2) Whether income data are expressed in constant or current values; (3) What are the units of analysis (individuals, families or consumption units); and (4) Which is our choice of income brackets that define the different social groups. Let us specify the chosen options.

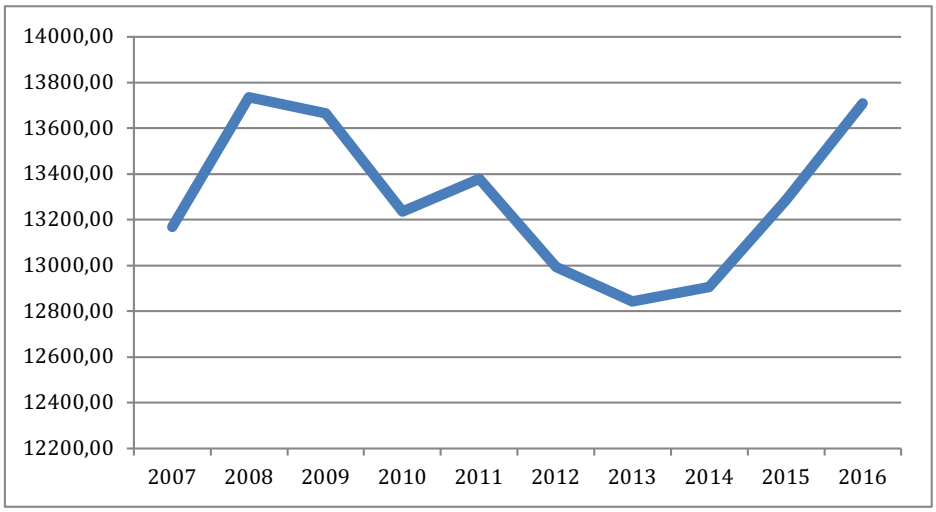
Regarding the income notion we select the yearly revenues, expressed in 2007 constant euros. It is well known that revenues and expenditures behave differently, especially during recessions, and represent alternative aspects of consumers' opportunities (e.g. Atkinson & Bourguignon, 2000, Deaton & Zaidi, 2002, Krueger & Perri, 2006, Brewer & O'Dea, 2012). We understand that in this context revenues provide a better evaluation of the agents' opportunities as they measure access to material goods while preserving wealth. This is a finer approximation to the notion of opportunity, in our view. As for the units of analysis, we take as reference the *consumption units*, which are households adjusted for their size and composition, following the European convention of giving weight 1 to the

[5] It is more usual to use the Encuesta de Condiciones de Vida (Life Conditions Survey) to get data regarding households' revenues. Yet we have preferred recurring to the FES due to the larger sample size (24,000 households versus 13,000), as we have to get results on a partition involving 17 regions and 8 income groups within each region (see below).

first adult of the family, weight ½ to the other adults in the unit, and weight 1/3 to children.

Figure 1 describes the evolution of the average income of consumption units for the period.

Figure 1: Per capita income, consumption units in Spain 2007-2016 (constant euros of 2007)



Source: Encuesta de Presupuestos Familiares (INE)

Concerning the choice of income brackets, we divide society in eight different groups depending on the equivalized per capita income of the corresponding consumption unit, whose values are described in Table 2. Those values are obtained through the following procedure. First, we calculate the per capita income of the representative consumption unit in Spain in 2007 (13,170€). Then we divide society into two different groups, those with income above the mean and those with income below the mean. Now each of those two groups is considered as a society on its own and the division process is repeated on each of them. 20,011€ and 8,712€ are the corresponding per capita income of the representative consumption units that are above and below the mean, respectively. We replicate once more the process to generate the eight groups.[6]

[6] The process can be repeated more times, but a higher disaggregation does not alter the outcomes. Indeed, the differences between four and eight income-groups are already very small (see Herrero, Villar & Soler, 2018).

Table 2: Social groups by income brackets in 2007 euros

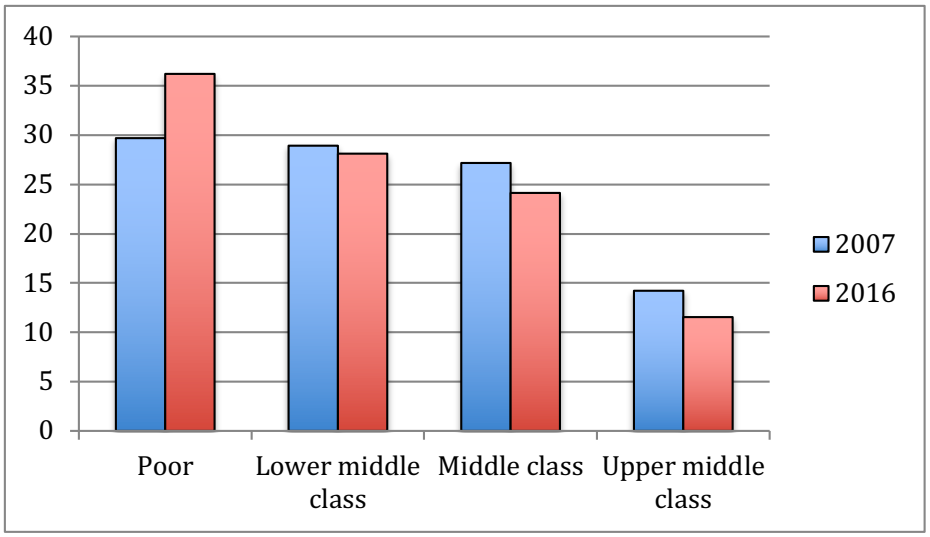
Income brackets	Thresholds
1	0 < y < 6,383
2	6,383 < y < 8,712
3	8,712 < y < 10,852
4	10,852 < y < 13,170
5	13,170 < y < 15,835
6	15,835 < y < 20,011
7	20,011 < y < 27,474
8	y > 27,474

The rationale of this way of dividing the income range to define the reference social groups can be found in the work of Esteban, Gradín & Ray (2006). They show that this way of finding the thresholds minimizes the error introduced when we substitute a complete income distribution by a reduced form in which we only consider a limited number of income intervals.

As already underlined in Section 2, we now shift the focus of our analysis from income distribution vectors, $\boldsymbol{y}_i$, to vectors of population shares into income brackets, $\boldsymbol{a}_i$. Detailed data on the distribution of the population among those income groups, both for 2007 and 2016 for Spain and its regions, are given in the Appendix.

It is interesting to give a simplified picture of what has happened in Spain between 2007 and 2016, recurring to a rougher aggregation than that used to calculate the change in income opportunity. Figure 2 shows the results on the change of population shares when we consider just four social groups: the poor (people in income brackets 1 and 2), the low middle class (income brackets 3 and 4), the middle class (income brackets 5 and 6), and the upper middle class (income brackets 7 and 8). The picture clearly illustrates how the crisis has pushed people towards the poor class, which implies that the probability of being poor for the average citizen has increased in 2016 relative to 2007. Indeed, as we shall see below, the opportunity advantage of the 2016 income distribution in Spain represents less than 87% of that in 2007, even though the mean income was higher in 2016.

Figure 2: Change in population shares between 2007 and 2016 in Spain



Source: Encuesta de Presupuestos Familiares (INE)

This picture already gives us a good hint of what has happened in Spain regarding income opportunities. Namely, catching up with the mean income has not allowed to recover income opportunities, due to the shifting of population to the lower levels of income (which is obviously possible only if income is now more concentrated in the upper levels).

Table 3 provides the data corresponding to the opportunity advantage of income distribution sin Spain and its regions in 2007 and 2016. We have normalised the values by taking the opportunity advantage of Spain in 2007 equal to 100. In this way the figures in the table tell us how far away each region is, in 2007 and 2016, from the average opportunities of a Spanish citizen in 2007.

Table 3: Relative opportunity advantage of income distributions, Spain and its regions (2007, 2016)

	2007	2016
España	100	86.6
Andalucía	65.8	60.0
Aragón	107.9	104.6
Asturias	121.4	110.1
Baleares	141.3	106.4
Canarias	86.6	63.6
Cantabria	95.9	93.2
Castilla León	97.9	93.9
Castilla Mancha	80.3	63.4
Cataluña	131.9	98.4
Comunidad Valenciana	89.9	82.3
Extremadura	52.0	51.9
Galicia	82.5	76.6
Madrid	148.1	125.3

Murcia	78.9	64.5
Navarra	147.8	120.0
País Vasco	155.3	152.0
Rioja	91.9	90.4
Coefficient of variation	0.290	0.285

The first row already tells us that the recovery of the mean income in 2016 still leaves income opportunities 13.4% below those in 2007.

Comparing the values in Table 3 across rows informs us about how much income opportunities have decreased within each region in 2016 relative to 2007. The regions in which this deterioration is higher are Baleares (almost 35 points of difference), Canarias (23), Cataluña (33.5), Madrid (22.7) and Navarra (27.7). Those with smaller relative reductions are Andalucía (5.8), Aragón (3.2), Cantabria (2.7), Castilla León (4), Extremadura (0.1), Galicia (5.9), País Vasco (3.3) and Rioja (1.4). Those data convey two main messages. First, that the crisis has had an impact on income opportunities that varies a lot between the Spanish regions. Second, that there is not a clear pattern on the type of regions regarding the size of the impact. The richer regions are most of those with a stronger opportunity reduction, but Canarias is far from being one of them. Those with smaller reductions are more varied and include Andalucía and Extremadura, which are relatively poor, but also País Vasco, which is the richest one.

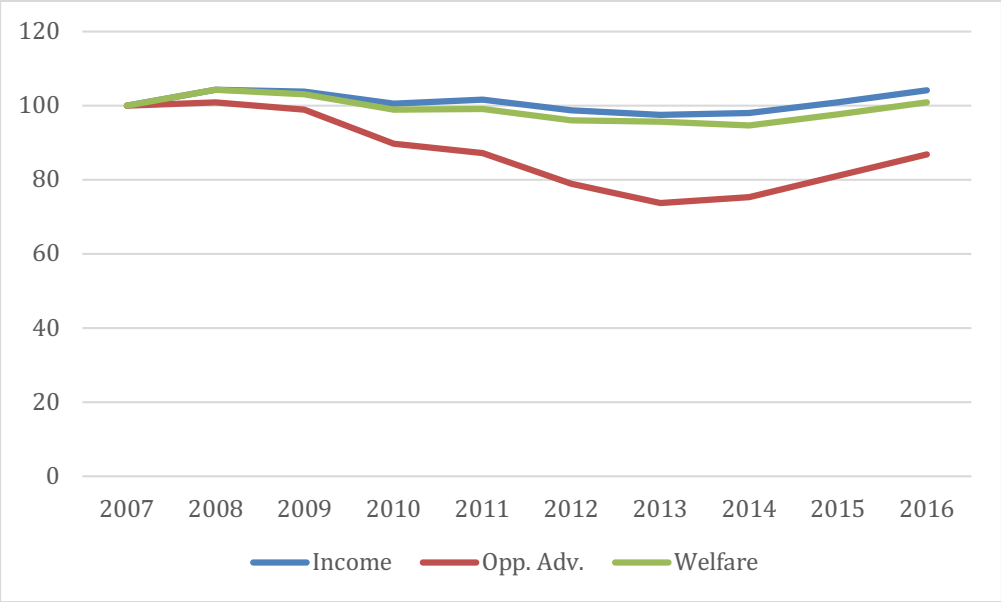
Comparing figures by columns we get an estimate of the opportunity bias, that is, how important is the region of residence regarding the opportunities it offers. In 2007 the values of income opportunity, relative to Spain, ranged from 155 in País Vasco to 52 in Extremadura (one third of the former). In 2016 those extreme values have moved very little (152 for País Vasco and 51.9 for Extremadura). In spite of the differential impact of the crisis commented above, Cataluña is the only region that has shifted from above to below the mean of Spain in 2007. The coefficient of variation has hardly changed.

Let us conclude this section by pointing out that the picture we get from the degree of success of the economic recovery in Spain is quite different when we look at the per capita income, the welfare measure that obtains when we deflate mean income by inequality, and the opportunity advantage. Here the welfare measure is given by:

$$w(\boldsymbol{y}) = \mu(\boldsymbol{y})[1 - G(\boldsymbol{y})]$$

where $G(\boldsymbol{y})$ is the Gini index of the income distribution. Figure 3 shows the evolution of the relative values of those three variables, by letting 100 the value of each variable in 2007. Income and welfare evolve quite similarly, as the changes in inequality have been relatively small (they range between 0.324 and 0.345). The gap between those measures along the period clearly points out that the average citizen has suffered along the period much more than it appears on the surface, from the income opportunity viewpoint.

Figure 3: Income, welfare and opportunity advantage in Spain 2007-2016 (relative values, Spain 2007 = 100 for both variables)



Source: Encuesta de Presupuestos Familiares (INE) and EU-SILC

5 Final remarks

Opportunity advantage is a criterion to evaluate income distributions based on the likelihood of getting higher incomes. It tries to capture the relative opportunities that the income distributions of different societies offer to their citizens, from a "veil of ignorance" perspective. The evaluation attached to each income distribution can be interpreted as the willingness to pay of the average citizen in order to stay in that society.

We shall conclude this paper by underlining some key aspects of this evaluation protocol.

(i) It provides a relative evaluation. That is, the opportunity advantage of a given income distribution is a measure that depends on the income distributions with which it is compared.

(ii) Opportunity advantage is not an *alternative* measure of income distributions, in the sense of a substitutive welfare indicator, but a *different* one. That is, it should be regarded as a complementary way of approaching the ranking of income distributions that measures something else than the standard welfare indicators (e.g. welfare measures of the type $W(\boldsymbol{y}) = \mu(\boldsymbol{y})[1 - I(\boldsymbol{y})]$, where $I(\boldsymbol{y})$ is a conventional income inequality measure).

(iii) The opportunity advantage approach induces a different way of looking at the data, as shown in Figure 2 and the Table in the Appendix. The focus is on the shares of populations on a set of income brackets, which provides information about how population shifts between social classes. This is a concern that has recently be expressed regarding the evolution of the middle classes (see for instance Gornick & Jäntti, Eds., 2010).

(iv) The structure of this evaluation protocol makes it clear that the choice of income brackets might be a key decision when it comes to implementing it. This will be the case when we choose a small number of intervals, whereas it will play no role for a large number of them (there is no restriction, either theoretically or computationally, to deal with as many income brackets as we want). Yet the contribution of Esteban, Gradín & Ray (2007) provide a non-arbitrary way of designing those income intervals which solve the problem.

REFERENCES

Atkinson, A. B., & Bourguignon, F. (2000). Introduction: Income distribution and economics. In A. B. Atkinson & F. Bourguignon (Eds.), **Handbook of income distribution** (Vol. 1, Chap. 1). Amsterdam: Elsevier Science Publishers B. V. North Holland.

Berman, A., & Plemmons, R. J. (1994), **Nonnegative matrices in the mathematical sciences**. Society for Industrial and Applied Mathematics.

Brewer, M., & O'Dea, C. (2012). Measuring living standards with income and consumption: Evidence from the UK (IFS Working Papers W12/12). Institute for Fiscal Studies.

Dasgupta, P., Sen, A., & Starrett, D. (1973). Notes on the measurement of inequality. *Journal of economic theory*, *6*(2), 180-187.

Deaton, A., & Zaidi, S. (2002). Guidelines for constructing consumption aggregates for welfare analysis (LSMS Working Paper No. 135). Washington, DC: The World Bank. http://www.worldbank.org/lsms.

Esteban, J.M., Gradín, C. & Ray, D. (2007), An extension of a measure of polarization, with an application to the income distribution of five OECD countries, **Journal of Economic Inequality** 5, n.o 1, 1-19

Fleuerbaey, M. (2008), **Fairness, Responsibility, and Welfare**. Oxford: Oxford University Press.

Gornick, J.C. & Jäntii, H. Eds. (2010), **Income Inequality and the Middle Class in Affluent Countries**, Stanford University Press.

Herrero, C. & Villar, A. (2013), On the Comparison of Group Performance with Categorical Data, **PLoS ONE**, Vol. 8, No. 12.

Herrero, C. & Villar, A. (2018), The Balanced Worth: A procedure to evaluate performance in terms of ordered attributes, **Social Indicators Research**, 140: 1279-1300.

Herrero, C., Villar, A. & Soler, A. (2018), Oportunidades de Empleo y Renta en España 2007-2016, Technical Report Fundación Ramón Areces.

Krueger, D., & Perri, F. (2006). Does income inequality lead to consumption inequality? Evidence and Theory, **Review of Economic Studies**, 73, 163–193.

Palacios-Huerta, I. & Volij, O (2004), The Measurement of Intellectual Influence, **Econometrica**, 72 : 963-977.

Peragine, V. (2002), Opportunity egalitarianism and outcome inequality: the rank-dependent approach, **Mathematical Social Sciences**, 44, 45–64.

Peragine, V. (2004), Measuring and implementing equality of opportunity for

outcome, **Social Choice and Welfare**, 22, 1–24.

Roemer, J. (1996), **Theories of Distributive Justice**, Cambridge: Massachusetts Harvard University Press.

Roemer, J. (1998), **Equality of Opportunity**, Cambridge: Massachusetts Harvard University Press.

Roemer, J. E., & Trannoy, A. (2015). Equality of Opportunities. Chapter 4 in AB Atkinson and F. Bourguignon: **Handbook of Income Distribution**.

Roemer, J. E., & Trannoy, A. (2016). Equality of opportunity: Theory and measurement. **Journal of Economic Literature**, *54*(4), 1288-1332.

Rothschild, M., & Stiglitz, J. E. (1973). Some further results on the measurement of inequality. *Journal of Economic Theory*, *6*(2), 188-204.

Ruiz-Castillo, J. (2003), The Measurement of Inequality of Opportunities, **Research in Economic Inequality**, 9 : 1-34.

Sen, A. (1985). Capabilities and commodities. *Amsterdam and New York: North-Holland*.

Shorrocks, A.F. (1983), Ranking income distributions, **Economica**, 50 : 3-17.

Villar, A. (2005), On the welfare evaluation of income and opportunity, **Contributions to Theoretical Economics** 5, 1 : article 3.

APPENDIX: POPULATION SHARES BY INCOME BRACKETS

We present here the data regarding the distribution of population shares in Spain and its regions, among the eight income brackets selected, both for 2007 and 2016. The data come from the Spanish Encuesta de Presupuestos Familiares 2017, elaborated by the Instituto Nacional de Estadística in 2017. Rodrigo Aragón and Ángel Soler, both technical staff at the Insitituto Valenciano de Investigaciones Económicas helped us to generate this table.

	2007	2016
Spain	1	1
Income bracket 1	0,1482	0,1936
Income bracket 2	0,1487	0,1684
Income bracket 3	0,1746	0,1377
Income bracket 4	0,1147	0,1434
Income bracket 5	0,1491	0,1143
Income bracket 6	0,1226	0,1270
Income bracket 7	0,0983	0,0767
Income bracket 8	0,0438	0,0388
Andalucía	1	1
Income bracket 1	0,2430	0,2902
Income bracket 2	0,2079	0,2021
Income bracket 3	0,1836	0,1481
Income bracket 4	0,0917	0,1197
Income bracket 5	0,1070	0,0809
Income bracket 6	0,0797	0,0823
Income bracket 7	0,0569	0,0521
Income bracket 8	0,0302	0,0248
Aragón	1	1
Income bracket 1	0,1170	0,1304
Income bracket 2	0,1109	0,1525
Income bracket 3	0,1990	0,1409
Income bracket 4	0,1238	0,1688
Income bracket 5	0,2119	0,1367
Income bracket 6	0,1236	0,1323
Income bracket 7	0,0917	0,0981
Income bracket 8	0,0221	0,0402
Asturias	1	1
Income bracket 1	0,1089	0,1149

Income bracket 2	0,1418	0,1626
Income bracket 3	0,1366	0,1334
Income bracket 4	0,1341	0,1545
Income bracket 5	0,1411	0,1490
Income bracket 6	0,1671	0,1375
Income bracket 7	0,1245	0,0987
Income bracket 8	0,0459	0,0493
Baleares	1	1
Income bracket 1	0,0880	0,1439
Income bracket 2	0,0977	0,1286
Income bracket 3	0,1424	0,1434
Income bracket 4	0,1244	0,1707
Income bracket 5	0,1871	0,1188
Income bracket 6	0,1793	0,1646
Income bracket 7	0,1247	0,0880
Income bracket 8	0,0564	0,0421
Canarias	1	1
Income bracket 1	0,1704	0,2901
Income bracket 2	0,1556	0,1763
Income bracket 3	0,1960	0,1489
Income bracket 4	0,1372	0,1198
Income bracket 5	0,1218	0,0935
Income bracket 6	0,1016	0,0845
Income bracket 7	0,0873	0,0622
Income bracket 8	0,0300	0,0248
Cantabria	1	1
Income bracket 1	0,1196	0,1616
Income bracket 2	0,1773	0,1460
Income bracket 3	0,1983	0,1485
Income bracket 4	0,1270	0,1690
Income bracket 5	0,1425	0,1409
Income bracket 6	0,1087	0,1398
Income bracket 7	0,0913	0,0630
Income bracket 8	0,0353	0,0312
Castilla y León	1	1
Income bracket 1	0,1485	0,1402
Income bracket 2	0,1577	0,1904
Income bracket 3	0,1753	0,1514
Income bracket 4	0,1072	0,1334
Income bracket 5	0,1525	0,1264
Income bracket 6	0,1269	0,1547
Income bracket 7	0,0870	0,0706
Income bracket 8	0,0449	0,0328
Castilla-La Mancha	1	1
Income bracket 1	0,1757	0,2681

Income bracket 2	0,1790	0,2120
Income bracket 3	0,1932	0,1406
Income bracket 4	0,1274	0,1154
Income bracket 5	0,1341	0,0864
Income bracket 6	0,0975	0,1015
Income bracket 7	0,0765	0,0602
Income bracket 8	0,0167	0,0158
Cataluña	1	1
Income bracket 1	0,0986	0,1635
Income bracket 2	0,1084	0,1501
Income bracket 3	0,1521	0,1051
Income bracket 4	0,1197	0,1765
Income bracket 5	0,1841	0,1478
Income bracket 6	0,1516	0,1468
Income bracket 7	0,1398	0,0800
Income bracket 8	0,0458	0,0301
Comunitat Valenciana	1	1
Income bracket 1	0,1397	0,1955
Income bracket 2	0,1687	0,1667
Income bracket 3	0,2065	0,1627
Income bracket 4	0,1173	0,1511
Income bracket 5	0,1531	0,1052
Income bracket 6	0,1191	0,1124
Income bracket 7	0,0674	0,0640
Income bracket 8	0,0282	0,0424
Extremadura	1	1
Income bracket 1	0,3340	0,2988
Income bracket 2	0,1907	0,2436
Income bracket 3	0,1819	0,1635
Income bracket 4	0,0893	0,1067
Income bracket 5	0,0762	0,0586
Income bracket 6	0,0624	0,0677
Income bracket 7	0,0465	0,0480
Income bracket 8	0,0190	0,0131
Galicia	1	1
Income bracket 1	0,1568	0,1742
Income bracket 2	0,1794	0,1898
Income bracket 3	0,2091	0,1644
Income bracket 4	0,1233	0,1865
Income bracket 5	0,1444	0,1306
Income bracket 6	0,0964	0,0872
Income bracket 7	0,0656	0,0478
Income bracket 8	0,0248	0,0196
C. de Madrid	1	1
Income bracket 1	0,0844	0,1260

Income bracket 2	0,0992	0,1259
Income bracket 3	0,1505	0,1242
Income bracket 4	0,1225	0,1378
Income bracket 5	0,1650	0,1186
Income bracket 6	0,1349	0,1773
Income bracket 7	0,1501	0,1201
Income bracket 8	0,0934	0,0701
Murcia	1	1
Income bracket 1	0,2142	0,2666
Income bracket 2	0,1707	0,1989
Income bracket 3	0,1690	0,1552
Income bracket 4	0,1091	0,1070
Income bracket 5	0,1146	0,1079
Income bracket 6	0,1256	0,0806
Income bracket 7	0,0658	0,0537
Income bracket 8	0,0310	0,0301
Navarra	1	1
Income bracket 1	0,0706	0,1159
Income bracket 2	0,0969	0,1043
Income bracket 3	0,1574	0,1599
Income bracket 4	0,1164	0,1516
Income bracket 5	0,1775	0,1419
Income bracket 6	0,1982	0,1913
Income bracket 7	0,1258	0,0994
Income bracket 8	0,0573	0,0356
País Vasco	1	1
Income bracket 1	0,0630	0,0680
Income bracket 2	0,0992	0,1143
Income bracket 3	0,1522	0,1172
Income bracket 4	0,1139	0,1580
Income bracket 5	0,1725	0,1515
Income bracket 6	0,1876	0,1721
Income bracket 7	0,1422	0,1346
Income bracket 8	0,0695	0,0843
La Rioja	1	1
Income bracket 1	0,1392	0,1705
Income bracket 2	0,1653	0,1572
Income bracket 3	0,1918	0,1350
Income bracket 4	0,1157	0,1771
Income bracket 5	0,1732	0,1252
Income bracket 6	0,1219	0,1389
Income bracket 7	0,0725	0,0729
Income bracket 8	0,0202	0,0231