Hacienda Pública Española/Review of Public Economics

Manuscript 2062

INEQUALITY OF OPPORTUNITY IN SPAIN: NEW INSIGHTS FROM NEW DATA

LEOPOLDO CABRERA

GUSTAVO ALBERTO MARRERO

JUAN GABRIEL RODRIGUEZ

PEDRO SALAS-ROJO

INEQUALITY OF OPPORTUNITY IN SPAIN: NEW INSIGHTS FROM NEW DATA[1]

Leopoldo Cabrera
Universidad de la Laguna (Spain), CEDESOG and ISTUR
Tel: +34 922 317304. E-mail: lcabre@ull.edu.es.

Gustavo A. Marrero
Universidad de la Laguna (Spain), EQUALITAS and CEDESOG
Tel: +34 922 317123. E-mail: gmarrero@ull.es.

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

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

Abstract

Modern theories of justice consider Inequality of Opportunity (IO), the part of overall inequality explained by individual circumstances (factors beyond the individual control, like socioeconomic background), as the truly concept of unfair inequality. In addition, recent empirical studies have found that IO harms growth. Then, given the big increase in income inequality in Spain during the last decade (now one of the highest levels in the EU), how large is IO in Spain? By using a novel database from the Centro de Investigaciones Sociológicas (CIS) questionnaire on '*Social inequality and social mobility in Spain*', we observe that the share of IO is 44% of overall inequality (Gini index). By circumstances, we find that about 90% of IO is due to parental education and occupation, the type of school attended, the gender of the household's head and the size of the household. In addition, it is found that a large share of IO is channeled through the occupation and, especially, the level of education of the individual. These findings are consistent with the low levels of relative mobility in education and occupation observed in the database for Spain (2017).

Keywords: inequality of opportunity; intergenerational mobility; educational channel; occupational channel; Spain.

JEL Classification: D63, J62, I24.

[1] We acknowledge suggestions of Florian Wendelspiess and the collaboration of the Centro de Investigaciones Sociológicas (CIS) to elaborate the Questionare CIS-3178 on '*Desigualdad Social y Movilidad Social en España*'. This paper has received financial support from the Ministerio de Economía y Competitividad of Spain (Marrero through project ECO2016-76818-C3-2-P, and Rodríguez and Salas-Rojo through project ECO2016-76506-C4-1-R). Marrero also acknowledges financial support from Gobierno de Canarias for support through the ProID2017010088 (María del Carmen Betancourt y Molina program) R&D project, co-funded by the Operative Program FEDER 2014-2020. Salas-Rojo also acknowledges financial support from Ministerio de Ciencia, Innovación y Universidades through FPU grant 17/735. All views, and any remaining errors or omissions are our sole responsibility.

1. Introduction

Income inequality has significantly increased in Spain since 2002. According to The Standardized World Income Inequality Database (SWIID), the Gini index of disposable income went from 30.0 in 2002 to 34.2 in 2016 (Solt, 2019). This increase of more than four points has been the largest increase in inequality in Europe since 2002 (see Figure 1 for an international comparison). In fact, Spain shows after Bulgary (with a Gini index equal to 35.0) the largest inequality of disposable income in Europe in 2016. It is interesting to note that this increase of inequality in Spain has occurred before, during and after the Great Recession (1.8 points of increase from 2002 to 2007 and 2.4 points of increase from 2007 to 2016), so it cannot be entirely blamed for the upward inequality trend observed in Spain during these years. At this point, we wonder to what extent these data on Spain are worrisome. The literature has traditionally focused on inequality, but modern theories of justice have shown that Inequality of Opportunity (IO) is a better concept to deal with the issues of justice and efficiency. Following this strand, the main goal of the paper is to study IO and its channels of transmission in Spain using a wide range of circumstances that, to the best of our knowledge, has never been explored in the IO literature.

There are two main reasons to focus on IO. First, the modern theory of justice recognizes that an individual's income is a function of variables *beyond* and *within* the individual's control, called circumstances (race, socioeconomic background, health endowments, gender or place of birth) and effort (related with the number of hours worked, investment in human capital or occupational choice), respectively.[2] As a result, overall inequality is actually a composite measure of IO, which is the part of total inequality explained by the unequal distribution of individual circumstances, and inequality of effort (IE), which represents differences in individual choices (Roemer, 1993 and 1998; Van de Gaer, 1993; Fleurbaey, 2008). Only the first type of inequality, IO, is the one that is truly important from the standpoint of social justice since the individual is not responsible of it.

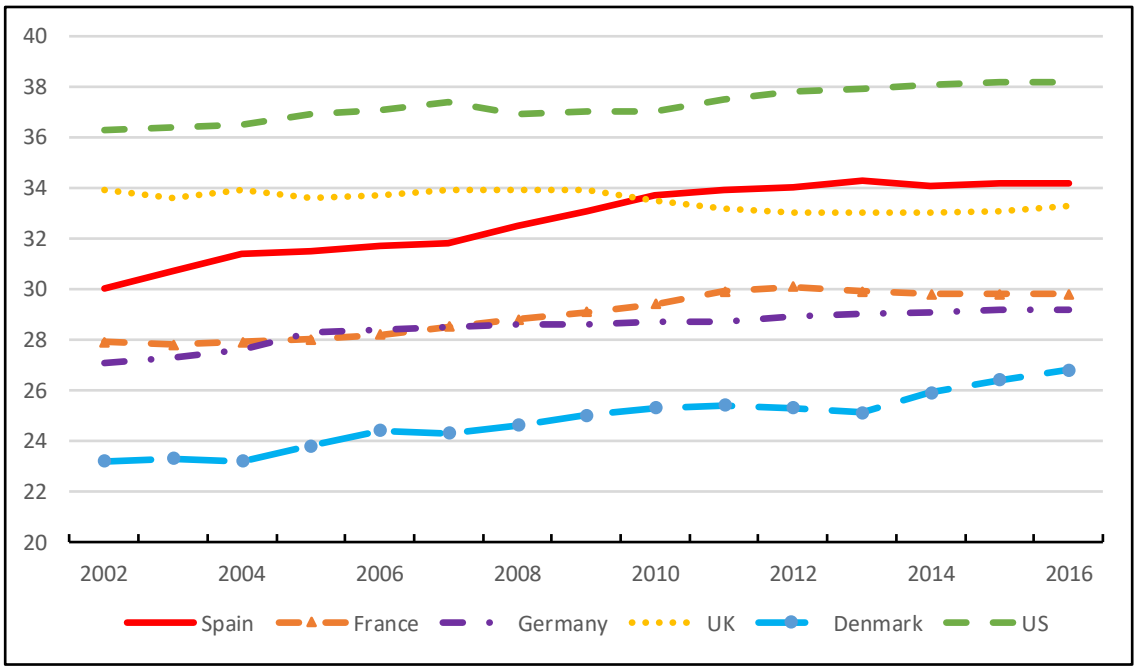
Second, despite all efforts, whether inequality is good or bad for growth is still a challenging question. This lack of consensus is attributed to the co-existence of a variety of channels –some of them growth-enhancing and others growth-deterring– through

[2] Talent could also be considered a circumstance, although this variable may reflect past effort of a person. On the other hand, luck could be considered an additional source of income (Lefranc et al., 2009).

which inequality affects growth.[3] To exit from this impasse, the literature has recently explored the *cholesterol hypothesis* (Ferreira, 2007). This hypothesis states that the part of total inequality generated by individual circumstances (IO) is growth-deterring, while the type of inequality generated by the difference in the willingness to exert effort (IE) is growth-enhancing. As a result, the final effect of total inequality on growth depends on the type of inequality that dominates. Studying this hypothesis, recent empirical work has found that IO is harmful for growth (Marrero and Rodríguez, 2013 and 2019; Bradbury and Triest, 2016; Marrero et al., 2016). Consequently, correcting a country's IO would not only result in a fairer society, but it would also spur economic growth (Marrero and Rodríguez, 2016).

Figure 1. International comparison of inequality of disposable income.

(Gini index, SWIID)



In this paper, we focus on measuring the level of IO in Spain. Since previous results have shown that Spain is one of the European countries with highest levels of IO, Spain is a relevant case of study. For instance, Rodríguez (2008) found that Spain (1990) was the

[3] The positive channels are related to the incentives for saving and investing (Kaldor, 1956; Barro 2000), asymmetric information (Mirrlees, 1971), and the generation of productivity premiums (Goldin and Katz, 2008; Mankiw, 2013). The negative channels are related to imperfect capital markets (Galor and Zeira, 1993; Banerjee and Newman, 1993), political economy issues (Alesina and Rodrik, 1994; Stiglitz, 2012), and the development process (Dasgupta and Ray, 1987).

country with the highest IO among the following set of countries: Belgium (1991), Denmark (1992), Great Britain (1990), Italy (1992), France (1993), Sweden (1990), United States (1990), Norway (1994) and France (1993). Later, Marrero and Rodríguez (2012) and Palomino et al. (2019) found that Spain still had a significantly large IO for European standards in 2004 and 2010.

Among the existing methods to estimate IO (see Ramos and van de Gaer, 2016), we apply the ex-ante parametric approach (Bourguignon et al, 2007; Ferreira and Gignoux, 2011) for two reasons. First, this method allows us to compare our results with previous studies for Spain (Marrero and Rodríguez, 2012; Palomino et al., 2019) and second, it allows us to exploit the large set of circumstances included in the database under consideration. When the number of observed circumstances is high, some types (groups of people with the same circumstances) may present a small number of observations and, as a result, the non-parametric estimates may be inaccurate.[4] For equality of opportunity, individual circumstances should not affect the distribution of income, i.e., the part of total income explained by individual circumstances should show zero inequality. Our analysis is mainly focused on the Gini index but, for robustness, we also use the Mean Logarithmic Deviation (MLD).

A first contribution of this paper is the database used. The module '*Social Inequality and Social Mobility in Spain*' carried out by the Centro de Investigaciones Sociológicas (CIS) during 2017 in Spain was designed by us in collaboration with some other researchers (see Marrero et al., 2017 and Betancort et al., 2019).[5] The collaboration with the CIS allowed us to have information for a wide range of circumstances that, as said, have never been explored in the IO literature. Apart from gender, parental education and occupational status of the father, we have the structure of the household where the individual grew up; the health perception during childhood and adolescence; the type of school where the participant studied (public/semi-private/private); parental reading and cultural activities; and, availability of sport, transport, sanitary and other similar public infrastructures during childhood and adolescence.

[4] In contrast to the parametric method, the non-parametric approach (Checchi and Peragine, 2010) is not a regression-based approach for computing IO. See Marrero and Rodríguez (2011) for an empirical comparison between the parametric and non-parametric approaches for the US (1970-2009).

[5] The CIS is a dependent entity of the Spanish Ministry of Presidency. The main scope of this organization is to improve the scientific knowledge of the Spanish society by gathering data and supporting training and research in the field of social sciences. For more information about this institution see http://www.cis.es/cis/opencms/EN/8_cis/.

As a second contribution, we measure intergenerational mobility in education and occupation for Spain in 2017. Note that intergenerational mobility is actually a particular case of IO since parental education or parental occupation are the only circumstances under consideration. We find a high upward educational mobility, as 47.5% of individuals exceed their parental educational levels. On the other hand, downward mobility is very small, 10.3%. Despite these results, we show that reaching the highest educational rank is strongly conditioned by family origin, i.e., relative mobility is low. For example, the percentage of individuals with university studies is 64.7% when at least one parent has university studies, while this share is only 16.36% when the parents have only basic studies. With respect to parental occupation, 26.8% of respondents have an occupation of greater rank than their father's (upward mobility) and 22.3% less (downward mobility). Meanwhile, the percentage of respondents in the group of highest occupations (managers, administrators, technicians and high-level professionals), 20.2%, almost doubles (38.2%) if the father belongs to that group and significantly decreases (12.2%) when the father belongs to the unqualified group of workers.

Using the Gini index, the share of overall inequality explained by our set of circumstances is 44.1%, i.e., almost half income inequality corresponds to IO. In addition, despite its shortcomings (Brunori et al., 2019a), we calculated for comparability the IO ratio for the Mean Logarithmic Deviation (MLD). This value is 17.7%, which is significantly higher than previous estimations. For instance, Palomino et al. (2019), applying the same parametric ex-ante approach and inequality index (MLD) on a smaller set of circumstances from the EU-SILC database, found an IO ratio of 12.5% for Spain in 2010. By circumstances, the Shapley value decomposition confirms that parental education and growing in a large family are the most important circumstances. Father’s occupation and the type of school also present a high contribution to overall IO, while gender and cultivated parents are less important. The contributions of the urban environment and healthy childhood are negligible. These contributions are robust to the inequality index under consideration.

Our last contribution is the analysis of education and occupation as channels of transmission of IO in Spain. In line with the proposal of Palomino et al. (2019), we first regress household income on the level of education and occupation of the household head. Then, for the predicted level of income explained by education (conditioned on the occupation) and occupation (conditioned by the level of education), we replicate the

previous analysis to calculate IO and the contribution of each circumstance. For the educational-adjusted household income (conditional on occupation), our set of circumstances explain 75.5% and 32.35% of total inequality according to the Gini index and the MLD, respectively. Moreover, a Shapley value decomposition finds that parental education has the largest contribution (36.8% for the Gini index and 45.1% for the MLD). The occupation of the father, the type of school, having born on a large family with cultivated parents also contribute significantly. For the occupational-adjusted household income (conditional on education), individual circumstances explain 56.7% (Gini index) and 15.5% (MLD) of total inequality. Once more, parental education and having born on a large family are important contributors. However, the most important circumstance is now the type of school attended, while fathers' occupation weakens its effect.

In the next section, we explain the database and explore the set of circumstances. In Section 3 we present the analysis of intergenerational mobility in education and occupation. The main results of the paper, the IO estimates and contributions of circumstances, are explained in Section 4. We investigate the transmission channels of IO in Section 5, while Section 6 presents some policy recommendations. Finally, we summarize our main results and conclusions in Section 7.

2. A new database

This section describes the module '*Social Inequality and Social Mobility in Spain*' ('*Desigualdad Social y Movilidad Social en España*') carried out during 2017 by the Centro de Investigaciones Sociológicas (CIS). The design of the questionnaire (CIS-3178) is based on Marrero et al. (2017) and Betancort et al. (2019), where a genuine database for the Canary Islands to estimate IO was first used.[6]

2.1. The database

The database used in this paper is new and unique for Spain. It is representative for the Spanish population in 2017, and it contains a rich set of circumstances that allows us to

[6] The details of the collaboration agreement were published in the Boletín Oficial del Estado, number 246 (see https://www.boe.es/boe/dias/2017/10/12/pdfs/BOE-A-2017-11708.pdf). Moreover, the entire database, the details of the questionnaire, and the sample design are available at: http://www.cis.es/cis/opencm/ES/2_bancodatos/estudios/ver.jsp?estudio=14350.

estimate and characterize IO for Spain with more precision than ever before, and also analyze the channels of transmission through education and occupation.

The sample focused on people over 18 years of age is representative by age and gender: 2,500 interviews were carried out with proportional affixing and finally 2,482 valid interviews were obtained.[7] The sampling procedure was multi-stage and stratified by conglomerates (municipalities, sections and individuals). The sample error is ± 2% for the whole sample with a confidence level of two standard deviations.

The questionnaire gathered information on three types of variables. First, a set of common variables obtained by any survey conducted by the CIS, such as age, family structure, place of residence, etc. Second, information related with the achievements of the individual or the household, such as (net) income, the highest level of achieved education, occupation and health status. Finally, the most important contribution of this questionnaire, the collection of individual circumstances as described below.[8]

2.2. Income and circumstances

Our analysis is based on household income, which comprehends all sources of income perceived by the household after adjusting by the number of household members.[9] Nonetheless, we replicate the whole analysis of the paper for the income of the household head, although the results are very similar (see Appendix). Income values are not reported in levels but in ranges, so we attribute the corresponding middle point of the range to each observation. The other two key output variables refer to individual education and occupation.[10] Following the IO literature, the age range considered to estimate IO is

[7] The interviews held from the 13rd to the 24th of November (2017) were face-to-face.
[8] Parental education and occupation correspond to the respondent when she was 16 years old. Being this way correct, it has the problem that some information is lost since not all the interviewees met their parents and some do not know the education/occupation of their father/mother at this age.
[9] In order to make household incomes comparable, we adjust the reported values by the square root scale which divides household income by the square root of household size (see Buhmann et al., 1988 and Coulter et al., 1992). The use of this equivalence scale is common in the studies of inequality for the case of Spain (see for example Perrote et al., 2003).
[10] Education levels are measured by the number of years studied and by the highest level of education attained (we use International Standard Classification of Education, ISCED-11). To characterize the population according to the type of occupation, we use the International Standardized Classification of Occupation (ISCO-08) following the CIS adjustment suggestions. Therefore, the professional category is used to give a representation of the social-occupational state of the individual, distinguishing between lower, middle and upper classes. The lower class refers mainly to unskilled workers, skilled workers in agriculture and fisheries, skilled workers in the construction industry, skilled workers in the industry / mechanics and workers. The middle class refers mainly to administrative staff, hospitality and catering workers, personal services and security workers, shop clerks, drivers and mobile machinery operators,

between 25 and 60 years old, in order to avoid a potential life-cycle bias in our estimates and, at the same time, maximize the sample size. After filtering the sample for age, the total number of observations is 1,561. Table 1 presents the main statistics of our dependent variable, adjusted household income. The average income scales up to 971€, with a standard deviation that reaches 556€. In addition, we observe that our selection is reasonable, as there are no big differences between cohorts and the median income is consistent with the permanent income hypothesis.

Table 1. Adjusted household income by cohorts.

Adjusted Household Income	Mean	Standard Deviation	Min	Max
All sample	970.98	555.89	130.19	3464.68
25-29 years old	968.61	550.32	130.19	3464.68
30-34 years old	1030.24	592.83	184.12	3031.67
35-39 years old	976.35	535.49	170.46	2683.73
40-44 years old	1037.05	513.33	134.61	2683.73
45-49 years old	877.19	515.36	170.46	2348.32
50-54 years old	951.92	639.73	184.12	3464.68
55-60 years old	865.38	523.12	201.69	3464.68

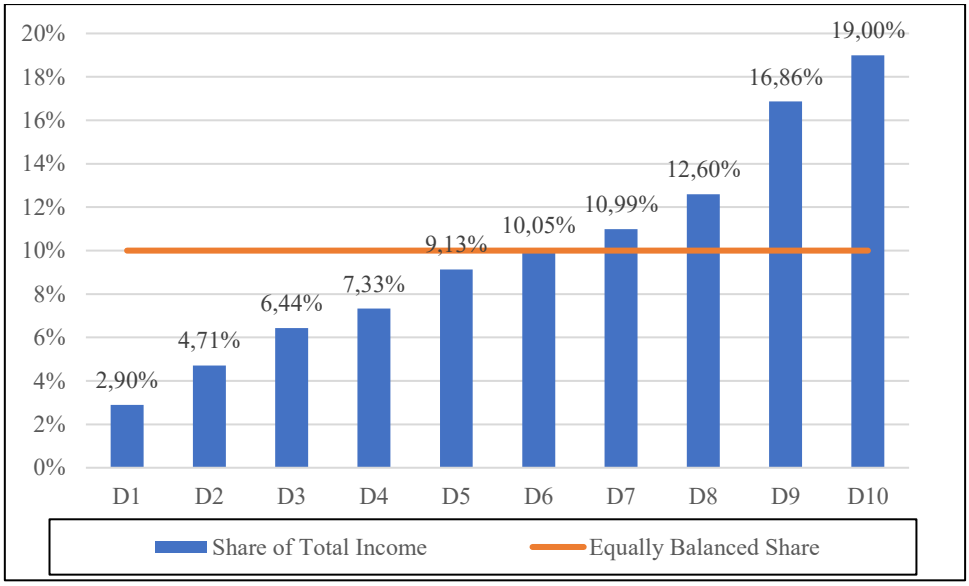
Deepening into the analysis on overall inequality, Figure 2 deploys the distribution of household income by deciles. It is observed that half of total income is gathered by the three top deciles. Particularly, the top decile (households with monthly income above 2,152€) accumulates 19.00% of total household income, which is almost seven times higher than the value of the first decile (households with monthly income up to 328€). This result is consistent with the findings in Ayala (2016) for Spain in 2014.[11] However, our data reveals a higher effect on the tails: the distance between the ninth and the first

armed forces and police. The upper class refers mainly to directors, managers and entrepreneurs, professionals and technicians, professional assistants and technicians.

[11] The database used by Ayala (2016) was *Encuesta de Presupuestos Familiares*.

deciles (p90/p10) is 5.82 while it is 4.39 in Ayala (2016). Nevertheless, if we look at the distance between the ninth decile and the median (p90/p50), our results are a bit lower: 1.84 against 1.97 in Ayala (2016). Finally, for our sample of individuals between 25 and 60 years old, our estimated Gini Coefficient reaches 0.310 points, which is very close to that obtained by Ayala (2016) for 2014 (0.315) and smaller than that shown by the INE for 2017, which is 0.341.

Figure 2. Household monthly income in Spain by deciles.



Concerning the set of circumstances, we have information for a wide range of innovative and relevant variables that, to the best of our knowledge, have never been explored in the IO literature. According to their nature, these variables have been classified into three groups. First, '*basic circumstances*', which includes gender, the structure of the household in which the individual grew up (whether the respondent has two or more siblings, or not), and health perception during childhood and adolescence (healthy or unhealthy). Second, '*socioeconomic circumstances'*, measured by the type of school where the participant studied (public/private/semi-private), the highest educational level achieved by parents and, finally, the occupational status of the father.[12] Finally, '*perception circumstances'*, a set of variables measuring respondent's perception of her family and school environment during adolescence. Among this set of circumstances, we

[12] The use of mother's occupation would reduce the size of the sample by 800 observations because there are too many missing values due to the category of "Housewives".

have '*cultivated parents*' which considers whether the respondent grew on a household where the parents habitually read or promoted cultural activities, and '*environment*' that reflects if the respondent grew with sport, transport, sanitary and other similar public infrastructures.

Table 2 summarizes the description of the set of circumstances and provides the sample mean and standard deviation for our sample. With respect to the set of basic circumstances, our sample is evenly distributed between women and men. Moreover, the 35% of the sample has more than two siblings and the vast majority of our respondents perceived a healthy childhood (only a 12% declared to have had an 'unhealthy' childhood). The socioeconomic circumstances reflect that most respondents went to public school (74%), being the rest evenly distributed among private and semi-private schools. On average, the highest educational level achieved by any parent is 7.14 years, but the large standard deviation (4.46 years) remarks the wide dispersion of this variable. Around 80% have a father belonging to the middle class, while 14% belonged to the high class and the remaining 6% to the low class. With respect to the circumstances of perception, 27% of respondents remember to have grown on a cultural-enhancing environment and around 75% remember to have grew with the necessary infrastructures.

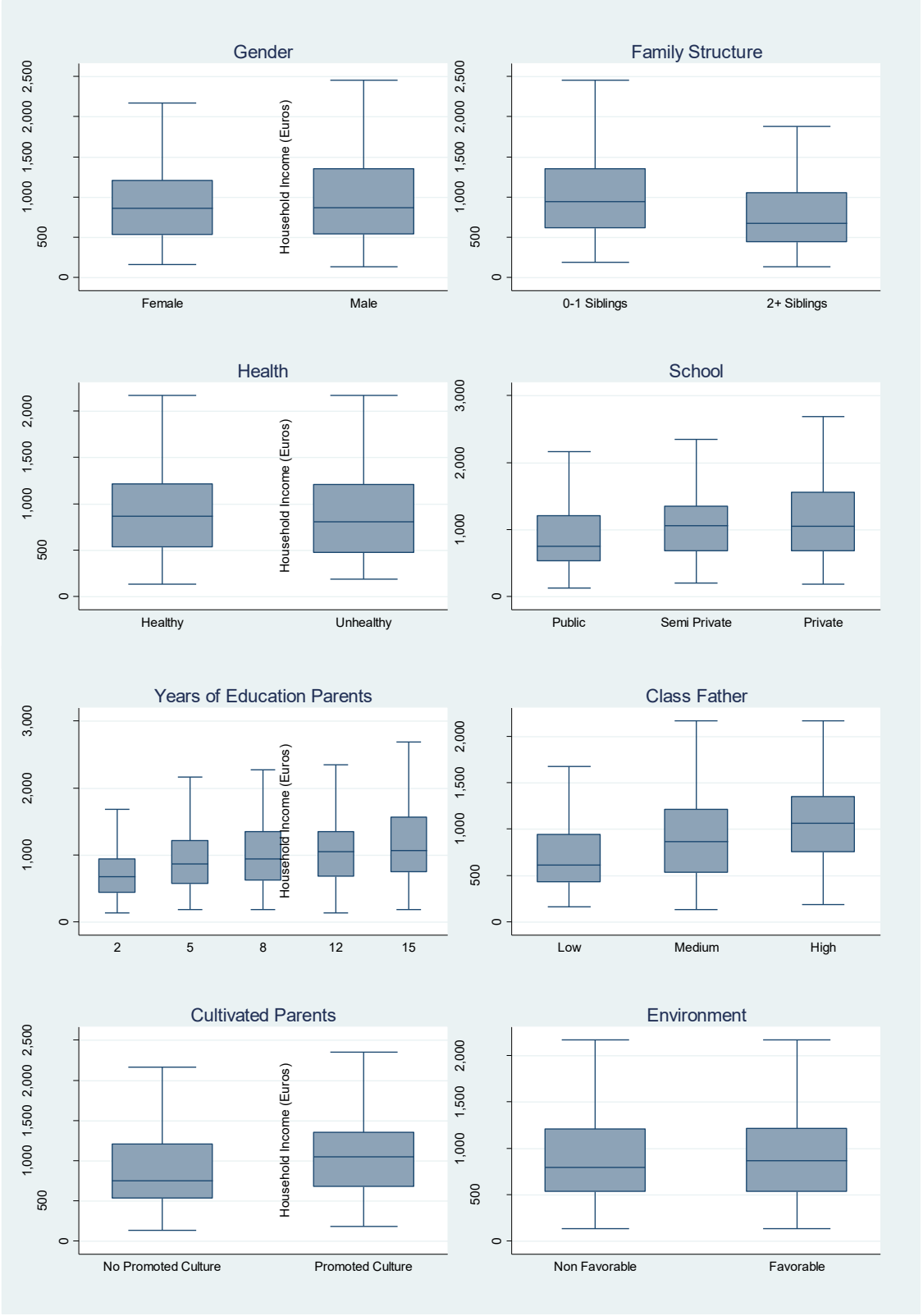
Figure 3 presents preliminary evidence of the importance that our set of circumstances have on household income. We summarize the results in a set of eight box plots showing the household income distribution across circumstances. On average, we observe that households with a male head earn more than those with a female head. The family structure seems to play a relevant role, as those households with their head with one or no siblings earn, on average, more than those that grew on a large family (two or more siblings).

Concerning the health during childhood, the result is not surprising, as the group of people with a healthy childhood show a higher average income. The type of school is relevant: those individuals who went to public schools earn less income than those who went to semi-private and private schools. On the other hand, there is an income premium for those with more educated parents. Another circumstance that seems to be important is the occupational status (or class) of the father because it is highly positively associated with the income of the household. To have parents who frequently read and promote cultural activities is related to higher income, as it is to have grown on a favorable urban environment.

Table 2. Individual circumstances.

Variable (encoding)	Description	Mean	Sd
Basic circumstances			
Gender	(1) 'Male'; (0) 'Female'	0.50	0.50
Household Structure	Number of siblings of the respondent at the age of 16: (1) 'Two or more'; (0) 'One or none'	0.35	0.48
Health Perception	Perceived health status during childhood: (1) 'Healthy'; (0) 'Unhealthy'	0.88	0.32
Socioeconomic circumstances			
Type of School: Private	Type of school attended by individual during childhood: (1) 'Private'; (0) 'Otherwise'.	0.14	0.34
Type of School: Semi-Private	Type of school attended by individual during childhood: (1) 'Semi-Private; (0) 'Otherwise'.	0.12	0.33
Type of School: Public	Type of school attended by individual during childhood: (1) 'Public; (0) 'Otherwise'.	0.74	0.44
Highest level of parental education	Highest education level reached by father/mother, measured as the number of years completed.	7.16	4.46
Class of the father	Class of the father when the respondent was 16: (0) 'Low'; (1) 'Medium'; (2) 'High'.	1.08	0.44
Perception Circumstances			
Cultivated Parents	Remembers that her parents read and promoted cultural activities: (1) 'Parent did promote such activities'; (0) 'Parents did not promote such activities'.	0.27	0.44
Environment	Remembers that she grew on a place with sport, sanitary and other similar infrastructures well connected by road: (1) 'She grow on such environment'; (0) 'She did not grow on such environment'.	0.75	0.43

Figure 3. Distribution of monthly household income by circumstances.



3. Intergenerational mobility in education and occupation

We initiate our analysis of IO by studying intergenerational mobility. As said in the introduction, intergenerational mobility considers only one circumstance (parent's achievement) to measure the distribution of output (usually income, education or occupation). Because we do not have information on parental income, here we focus on the transmission of education and occupation in Spain. In both cases, the sample is restricted to the group of individuals who are between 25 and 60 years old, which is a usual age to have obtained the highest educational attainment. This implies a sample size of 1,340 observations.[13] For mobility in education, we consider the educational achievement of children and the highest level of education attained by the parents (Golthorpe, 2013). Following the ISCED classification system (UNESCO, 2012), we recoded the levels of studies to work with four categories.[14] Once the transition matrix is defined, upward mobility is obtained by summing up the shares over its main diagonal (immobility). Similarly, downward mobility is calculated by summing up the shares below the main diagonal.

Table 3 shows that there exists a high upward educational mobility: 47.46% of individuals exceed their parental educational level. Meanwhile, downward mobility is quite small: 10.30%. The remaining proportion of people, 42.24%, reached the same level of studies than their parents. Despite the high absolute intergenerational mobility in education, relative mobility is not so positive. For instance, the 27.61% of interviewees who have university studies are distributed very unevenly across the levels of parental education. Thus, this percentage increases to 64.73% ((134/207)*100) when at least one parent has university studies, but it is only 16.36% ((132/807)*100) when the parents only have basic studies. It seems clear that reaching a high educational achievement is significantly conditioned by family origin.

[13] There are almost 100 respondents aged 18 to 24 years (4% of the total sample) who claim to be active, working, unemployed or looking for their first job. On the other hand, from the age of 60, the majority of respondents declare being retired or pensioner. We have replicated this analysis of intergenerational mobility for the subsample used in the measurement of IO (see below) but the main results remain the same (they are available from the authors upon request).

[14] The four categories are: 0-1: zero or just primary education; 2: lower secondary education; 3-4: upper secondary education and post-secondary but non-tertiary education; 5-8: tertiary education (short-cycle tertiary education, bachelor's, masters and doctoral studies, or equivalent). Here, we have considered exogenous thresholds, but it could be interesting to analyze the effect of endogenous thresholds. Because this is not the main scope of the paper, we leave this analysis for future research.

Table 3. Studies of the respondent vs highest studies achieved by parents.

Educational Mobility			Studies of the respondent (4 categories)				Marginal Distribution
			ISCED 0-1	ISCED 2	ISCED 3-4	ISCED 5-8	
Highest studies achieved by any parent	ISCED 0-1	n	328	121	226	132	**807**
		%	(24,48%)	(9,03%)	(16,87%)	(9,85%)	(60,22%)
	ISCED 2	n	23	33	53	45	**154**
		%	(1,72%)	(2,46%)	(3,96%)	(3,36%)	(11,49%)
	ISCED 3-4	n	12	30	71	59	**172**
		%	(0,90%)	(2,24%)	(5,30%)	(4,40%)	(12,84%)
	ISCED 5-8	n	6	11	56	134	**207**
		%	(0,45%)	(0,82%)	(4,18%)	(10,00%)	(15,45%)
Marginal Distribution		**n**	**369**	**195**	**406**	**370**	**1340**
		%	(27,54%)	(14,55%)	(30,30%)	(27,61%)	(100,00%)

Note: ISCED=0-1: primary education or less; ISCED=2: lower secondary education; ISCED=3-4: upper secondary education and post-secondary non-tertiary education; ISCED=5-8: tertiary education (short-cycle tertiary education, bachelor's, masters and doctoral studies, or equivalent).

Our results are similar to those obtained in other studies for Spain and other developed countries (Marrero et al., 2017; IECA, 2018; and EU, 2018). In particular, they are consistent with those collected for Spain at Eurobarometer 471 of 2018. According to these data, upward educational mobility in Spain was around 60%, the largest value of upward mobility in the European Union (see graph QA9T, EU, 2018: 17). Moreover, in accordance to our results, a recent study by the OECD (2018) concludes that "an expansion of access to education, particularly tertiary education, does not automatically result in greater equity in educational attainment. For that to happen, disadvantaged students need to benefit as much as or more than advantaged students. In recent decades, some 41% of adults attained a higher level of education than their parents did, on average across countries that participated in the Survey of Adult Skills (PIAAC). However, the children of families with higher levels of education were more likely than the children of families with lower levels of education to benefit from the expansion of tertiary education".

Next, we analyze occupational mobility based on the interrelation between the occupation of the interviewed and that of the father (Table 4). As said, we do not use information on the mother's occupation because it would substantially reduce the sample size. To measure the occupation, we start from the ten large occupational groups of the ISCO-

08.[15] We reduce these groups to four to reflect the professional situation of individuals and generate occupational mobility matrices with a reasonable dimension.[16]

Table 4. Occupation of respondent and occupation of the father.

Occupational Mobility			Occupation of the Respondent				Marginal Distribution
			ISCO08=9	ISCO08=4-8	ISCO08=3	ISCO08=1-2	
Occupation of the Father	ISCO08=9	n	26	36	7	12	81
		%	(1,94%)	(2,69%)	(0,52%)	(0.90%)	(6,05%)
	ISCO-08=4-8	n	122	559	117	155	953
		%	(9,10%)	(41,72%)	(8,73%)	(11.57%)	(71,12%)
	ISCO-08=3	n	7	55	26	32	120
		%	(0,52%)	(4,10%)	(1,94%)	(2.39%)	(8,96%)
	ISCO-08=1-2	n	8	68	39	71	186
		%	(0,60%)	(5,07%)	(2,91%)	(5,30%)	(13,88%)
Marginal Distribution		**n**	**163**	**718**	**189**	**270**	**1340**
		%	(12,16%)	(53,58%)	(14,10%)	(20,15%)	(100%)

Note: ISCO-08=9: no qualified; ISCO-08=4-8: semi-qualified and qualified; ISCO-08=3: technicians, support professionals; ISCO-08=1-2: managers and professionals.

For this sample, we observe that 20.15% of respondents occupy the highest social class (managers, administrators, managers, technicians and high-level professionals). This percentage almost doubles, reaching 38.17% ((71/186)*100) when the father also belongs to the highest social class. The opposite occurs with the lowest social class (the unqualified). In this class 12.16% of the interviewees are located. This proportion almost triples when the respondent has a father already placed in such a position, 32.10% ((26/81)*100). For this sample, 26.79% of respondents have an occupational position of greater rank than their father's (upward mobility), 22.31% less (downward mobility) and

[15] We follow the international classification of occupations and the assignment to the 10 large occupational groups of the uniform international classification of occupations ISCO-08 (http://www.ilo.org/public/english/bureau/stat/isco/), International Labor Organization (ILO). In the case of military occupations, given their small number, the observations are adjusted according to the studies they indicate and their membership (officers or troops).

[16] The four occupational groups are defined as: 1, no qualified (ISCO-08=9); 2, semi-qualified and qualified (ISCO-08=4-8); 3, technicians, support professionals (ISCO-08=3); 4, managers and professionals (ISCO-08=1-2). The correspondence between these groups and the three parental classes considered in Table 2 is the following: low class = group 1, middle class = groups 2 and 3, high class = group 4.

50.90% the same (immobility). These results are consistent with the findings in Carabaña (1999), Marqués (2015) and Requena and Stanek (2015).[17]

To end this section, we calculate a set of mobility indices that summarize the degree of absolute mobility (upward and downward) in our transition matrices (M) for education and occupation. First, we calculate the Prais (1955)-Shorrocks (1978) mobility index which measures the average probability across all categories that an individual will leave the category of her parents. Second, the Bartholomew index represents the average number of categories crossed by all sons (with respect to the categories of their parents) (Bartholomew, 1982). Third, the mobility index proposed by Sommers and Conlisk (1979), that relies on the second largest eigenvalue (λ_2) of M which can be regarded as the distance between M and perfect mobility (there are no observations in the main diagonal of the matrix). Finally, we calculate the determinant index (Shorrocks, 1978) that uses the product of all eigenvalues as a measure of mobility.

Table 5. Intergenerational mobility in education and occupation.

Index of mobility	Education	Occupation
Prais-Shorrocks $M_1 = \frac{1}{(K-1)} * (K - trace(M))$	0.773 (0.021)	0.831 (0.024)
Bartholomew $M_2 = \frac{1}{K*(K-1)} * \sum_i \sum_j p_{ij} \lvert i-j \rvert$	0.283 (0.008)	0.296 (0.011)
1-Second largest eigenvalue $M_3 = 1 - \lvert\lambda_2\rvert$	0.567 (0.025)	0.699 (0.122)
Determinant index $M_4 = 1 - \lvert\det(M)\rvert = 1 - \prod_{i=1}^{m} \lambda_i$	0.996 (0.003)	0.998 (0.002)

Note: These four indexes are defined for a (KxK) transition matrix M with generic elements p_{ij}. Bootstrap standard errors are based on 500 replications. This analysis was performed by using the Stata command described in Savegnago (2016).

[17] Using data from the INE Sociodemographic Survey (1991) with 157,000 records, Carabaña (1999) found an absolute intergenerational mobility (upward and downward) of 70%. Marqués (2015) found the following absolute mobility for the male/female populations: 71/72% in the UK (43/52% upward and 28/20% downward); 82/72% in France (43/50% upward and 39/22% downward); 74/88% in Sweden (40/70% upward and 34/18% downward); 71/80% in Portugal (46/53% upward and 25/27% downward); 70/79% in Italy (46/53% upward and 24/26% downward); and 70/77% in Spain (46/53% upward and 24/24% downward).

The results in Table 5 highlight two empirical facts. First, as said, absolute mobility is high for both education and occupation. Only when we weight each class transition by the number of class boundaries that have been crossed (the Bartholomew index) we find that mobility is not high. Second, total mobility is always higher for occupation than for education. Despite these results, it is important to recall that social origin is important since relative mobility is low. To explain in detail this empirical fact we study inequality of opportunity in the next section.

4. Inequality of Opportunity in Spain

Equality of opportunity requires compensating persons for a variety of circumstances whose distribution is morally arbitrary (the *compensation principle*) and simultaneously adjusting the relationship between their outcome and their effort (the *reward principle*). Thus, an equal-opportunity policy with respect to a given outcome (income, utility, health) should allocate the resource so that it makes the degree to which an individual achieves the outcome a function only of her effort, and therefore independent of her circumstances because individuals are only responsible for their own efforts (Roemer, 1993 and 1998; Fleurbeay, 2008).

There are two main alternative methods to measure inequality of opportunity, namely, the *ex-post* approach (Roemer, 1993) and the *ex-ante* approach (van de Gaer, 1993). On the one hand, the *ex-post* approach states that there is equality of opportunity if all individuals who exert the same effort obtain the same outcome. Unfortunately, the distribution of effort is typically not observed so the Roemer's pragmatic approach is usually adopted. According to this proposal, two persons with different circumstances have tried equally hard if and only if they are on the same rank of their respective outcome distributions, that is, they have exerted the same degree of effort. On the other hand, there is equality of opportunity as defined by the *ex-ante* approach if all individuals face the same set of opportunities regardless of their circumstances. Because the use of the ex-post approach implies making serious assumptions on how effort is measured and for comparability reasons, here we focus on the second approach.

4.1. **Estimation of IO**

To estimate Inequality of Opportunity in Spain (2017), among the existing methods (see Ramos and van de Gaer, 2016, for a recent survey), we have leaned towards the ex-ante parametric approach described in Ferreira and Gignoux (2011). Two reasons justify this

choice. First, this method allows us to compare our results with previous studies for Spain and Europe (Marrero and Rodríguez, 2012; Palomino et al., 2019). Second, it allows us to easily exploit the large set of circumstances included in the database.[18] According to this methodology, it is estimated the following reduced form that relates individual's income, Y_i, with her circumstances, C_i:

$$\ln(Y_i) = \beta C_i + \varepsilon_i, \quad (1)$$

where ε_i is the part of household income not explained by the set of observed circumstances.

For equality of opportunity, personal circumstances should not affect the distribution of income. Consequently, inequality of fitted income, $\hat{Y}_i = exp(\hat{\beta} C_i)$, should be zero. By applying an inequality index to the predicted vector of incomes (the smoothed income distribution), we obtain a parametric estimate of absolute IO. If the measure of IO is divided by total inequality, we obtain a relative IO measure that reflects the share of overall inequality due to the set of circumstances considered. Here we wish to remark that, given the impossibility of observing all relevant circumstances in practice, our results should be interpreted as lower-bound measures of IO. Indeed, including more circumstances will provide more information which, in turn, will increase the value of IO. A deeper discussion on this issue can be found in Brunori et al. (2019b).

To test the robustness of our results, we use two different inequality indexes. Typically, the literature on IO has used the MLD because this index is the only additively and path independent decomposable inequality index (Foster and Shneyerov, 2000). For this reason, in order to compare our results with previous findings, we use the MLD index. However, our analysis will be mainly focused on the Gini index, which is the most used index of inequality. Following Brunori et al. (2019a), we find a great advantage of using this index in the IO framework. As said, IO is measured from a smoothed distribution that by construction does not contain extreme values since they are removed by the parametric estimation in (1). However, because the MLD is much more sensitive to extreme values than the Gini index, the reduction of inequality by going from the original to the smoothed distribution is much higher for the MLD than for the Gini index. As a result, the share of overall inequality explained by IO is significantly lower for the MLD index. In addition,

[18] This analysis was performed by using the Stata command described in Chávez-Juárez and Soloaga (2014).

we are not particularly interested in the decomposition of overall inequality into IO and IE since we just want to calculate the size of IO. Therefore, the fact that the Gini index is not additively decomposable is not a big problem.

4.2 Results of IO

Table 6 presents the OLS results of regressing the natural logarithm of household income against various sets of circumstances. For illustrative purposes, we consider six alternative models. In the first model, we include a number of basic circumstances: gender, the type of household in which the individual grew up, and health status during childhood. Subsequent models add additional relevant circumstances to this initial model. Thus, the second model adds the type of school attended by the individual during adolescence. The third model includes both parental education and occupation. The four model considers in addition to the basic circumstances, those relating to having or not "cultivated parents". In the fifth model we add to the basic circumstances the "environment" as defined in the previous section. Finally, the sixth model considers all circumstances together.

In general, results go in line with the intuitions presented in Figure 3. We start commenting the results for the basic set of circumstances. Being the household head a man is positive correlated with higher levels of income in the household. In particular, when conditioning on the whole set of circumstances, being a man has an income advantage between 9% and 12% more than being a woman. Those who belonged to a large family seem to have suffered a significant penalty, as they show the largest negative coefficient in our estimations (between 25% and 31% according to the different models estimated). In principle, this circumstance may be influenced by the individual's socio-economic framework, but it remains significant (and with similar coefficient) after controlling for the socio-economic circumstances (columns 3 and 6). Finally, the individual's self-perceived health status during her adolescence has a negative sign but it is not a significant circumstance in all estimated models.

The type of school attended is relevant. Thus, those individuals who attended to semi-private and private schools have, on average, a higher income (columns 2 and 6). In fact, attending to a private school is, among all circumstances, one of the most important positive contributors to income. With respect to a public school, attending to a semi-private school has an average advantage of about 13% in terms of income, while attending

to a fully private school provides about 20% more income. These results are consistent with the existing literature, which suggests that the educational performance of students in Spain is better in private schools and that the social network created in school influences future income levels.

Table 6. OLS regression of household income on circumstances (alternative models).

Dep. Variable: Household per capita adjusted Income						
Gender	0.091**	0.085**	0.124***	0.098***	0.091**	0.119***
	(0.035)	(0.035)	(0.037)	(0.035)	(0.035)	(0.037)
Family size	-0.312***	-0.277***	-0.257***	-0.302***	-0.309***	-0.260***
	(0.038)	(0.037)	(0.040)	(0.038)	(0.038)	(0.040)
Health status	-0.130	-0.123	-0.181	-0.169	-0.132	-0.190
	(0.184)	(0.177)	(0.182)	(0.179)	(0.184)	(0.182)
Semi-private school		0.190***	0.130**			0.124**
		(0.053)	(0.057)			(0.058)
Private school		0.297***	0.219***			0.208***
		(0.055)	(0.059)			(0.059)
Parental education			0.022***			0.021***
			(0.005)			(0.005)
Father medium class			0.156**			0.157**
			(0.071)			(0.071)
Father high class			0.223**			0.214**
			(0.089)			(0.089)
Cultivated parents				0.213***		0.045
				(0.041)		(0.045)
Environment					0.031	-0.022
					(0.041)	(0.043)
Constant	6.917***	6.696***	6.536***	6.862***	6.892***	6.557***
	(0.027)	(0.030)	(0.077)	(0.030)	(0.042)	(0.080)
R^2	0.07	0.09	0.14	0.09	0.07	0.15
Observations	1203	1203	1203	1203	1203	1203

*Note: standard errors in parenthesis. * p<0.1; ** p<0.05; *** p<0.01.*

Next, we comment on the estimated results for socioeconomic status during childhood (columns 3 and 6). Parental education significantly contributes to higher levels of household's income, whilst having a father belonging to the medium and high class increases income. Both, parental education and the social class of the father are highly significant in both models. Acording to our estimates, having a father from the 'medium class' provides a 15% more income than having a father from the "low class", whilst

having a father from the 'high class' increases expected income by more than 20%. In addition, conditioning on all other circumstances (included father's occupation), one more year of parental education implies an advantage of about 2% more income.

In column 4 we show that having cultivated parents could be a significant advantage. However, this variable becomes non-significant in column 6, when all variables (specially those representing socioeconomic background) are included in the model. Finally, a positive environment does not seem to make any difference on household income levels (columns 5 and 6).

4.3. IO and its decomposition: the role of circumstances

In this section we first show the estimates of both absolute and relative IO using the entire set of circumstances. Then, we calculate the contribution of each circumstance by applying the Shapley value decomposition, which is the only method that solves the tension between marginality and consistency (Chantreuil and Trannoy, 2013; Sastre and Trannoy, 2002; Rodríguez, 2004; and Shorrocks, 2013). By assuming that all possible combinations of factors have the same probability, the Shapley value calculates the contribution of each circumstance as the average of all its possible contributions.

In Table 7 we show that the share of IO measured by the Gini index is 44% of overall inequality. Despite that the ex-ante approach is very conservative since all the dispersion within types (groups of people with the same circumstances) is not considered for the calculation of IO, the share of estimated IO is very large. As expected and commented above, this percentage is significantly smaller for the MLD (around 18%). Here, it is important to notice that even the last value is significantly higher than previous IO estimates. For instance, Palomino et al. (2019), using the EU-SILC database and the same parametric ex-ante approach, found for the MLD and a reduced set of circumstances an IO ratio of 12.1% and 12.5% for Spain in 2004 and 2010, respectively. Because all estimates of IO are lower bounds (as said above), by adding more information on individual circumstances we provide more accurate estimates of IO.

With respect to the Shapley value decomposition, it seems in first place that the contributions of circumstances are robust to the inequality index under consideration. Parental education and growing on a large family are the factors that contribute more to IO. The significant effect of parental education on IO is a traditional result in this literature. However, the importance of the size of a family on IO is a new interesting

result. It seems that a larger family impedes parents to invest more time and resources in their children, what is reflected in their outputs later. The occupation of the father and the type of school in which the respondent studied are also important individual circumstances. The latter circumstance is not usually observed so we find another factor that studies on IO should also consider. The contributions of gender and having cultivated parents are relevant but not as large as previous contributions. Finally, urban environment and being a healthy child are negligible circumstances according to our estimates.

Table 7. Relative contribution of each circumstance.

	IO (Gini)		**IO (MLD)**	
	Absolute	**Relative**	**Absolute**	**Relative**
Index	0.14	44.09%	0.03	17.68%
Standard Deviation	(0.00)	(0.88)	(0.00)	(0.44)
	Shapley Decomposition			
	Relative Contribution		**Relative Contribution**	
Gender	6.22%		3.91%	
Size Family	26.82%		31.96%	
Health Status	1.25%		1.46%	
Type of School	14.47%		12.42%	
Parental Education	28.01%		31.52%	
Class of the Father	12.42%		12.68%	
Cultivated Parents	9.22%		5.61%	
Environment	1.59%		0.44%	

Note: standard deviations in parenthesis are based on 50 bootstrap replications.

5. The channels of education and occupation

Now we analyze the relevance of education and occupation (of the individual himself) as channels of transmission of IO in Spain. For this task, we propose a novel procedure, although in line with Palomino et al. (2019), who highlight the relevance of these two channels for Europe. First, we regress household incomes (in logs) on the education (Edu_i) and occupation (Ocu_i) of the household head,

$$\ln(Y_i) = \alpha_0 + \beta Edu_i + \gamma Ocu_i + \varepsilon_i, \quad (2)$$

where ε_i is the part of household income not explained by both channels.[19]

By doing this, we recover the part of household income explained by her education (conditioned on her occupation) and vice versa (household income explained by her occupation conditioned by her level of education):

$$\widehat{Y\_Edu_\iota} = \exp\left(\hat{\beta}\, Edu_i\right), \tag{3}$$

$$\widehat{Y\_Ocu_\iota} = \exp(\hat{\gamma}\, \mathrm{Ocu}_i). \tag{4}$$

Then, we apply for these two conditional predicted levels of income the same procedure as described in Section 4: estimate equation (1), calculate the resultant levels of IO and, finally, perform the Shapley value decomposition.

Table 8 presents the OLS results of equation (1) for these two new income measures. Model 1 uses as dependent variable the conditional income predicted by the years of schooling of the respondent, and Model 2 uses the conditional income predicted by the individual's occupation. In general, results are similar to those exposed in Table 6 for overall income, although there are some differences. For example, being male is not statistically associated with the income explained by education but is positive and strongly significant for the income adjusted by occupation.[20] It seems that the gender income gap does not come from the education system but rather from the posterior incorporation into the labor market.

A higher parental education and parental occupational status generally lead to higher levels of income in both cases, but its effect is more intense and significant for the household's educational adjusted income. Now, for both measures of income, having cultivated parents is positive and significantly correlated with household's income: having cultivated parents is associated with about 3% and 2.5% more educational and occupational adjusted income, respectively. In both models, the type of school is again significant and positive. Finally, the penalty of growing on a large family remains

[19] Our empirical results are: $\ln(Y_i) = 6.156 + 0.0471 \cdot Edu_i + 0.1998 \cdot Ocu_{2i} + 0.5174 \cdot Ocu_{3i} + \varepsilon_i$, where Edu_i represents the years of education of the individual, and Ocu_{2i} and Ocu_{3i} represent the medium and high occupation category, respectively (the omitted category is the lowest occupational class). The coefficient of determination, R^2, is 24.32%, which implies that there is still a large fraction of household's income explained by other channels. Nonetheless, these other channels could be related with the quality of education and occupation, although these variables are not observed here.

[20] In our sample, 32.43% and 56.03% of women have secondary and tertiary education, respectively. For men, these values are 62.61% for secondary education and 29.32 for tertiary education.

significant, although it is more intense and significant for the occupational adjusted income.

Table 8. Adjusted income by education and occupation on circumstances.

	Model 1 - Household Income (education)	**Model 2 - Household Income (occupation)**
Gender	-0.016	0.035***
	(0.012)	(0.010)
Size Family	-0.075***	-0.039***
	(0.013)	(0.010)
Health Status	-0.037	-0.032
	(0.036)	(0.037)
Semi – Private School	0.053***	0.051***
	(0.018)	(0.017)
Private School	0.052***	0.047***
	(0.016)	(0.015)
Parental Education	0.014***	0.005***
	(0.001)	(0.001)
Father Medium Class	0.090***	0.034*
	(0.030)	(0.019)
Father High Class	0.121***	0.040
	(0.035)	(0.025)
Cultivated Parents	0.030**	0.024**
	(0.013)	(0.012)
Environment	0.002	-0.012
	(0.015)	(0.011)
Constant	6.496***	6.082***
	(0.032)	(0.021)
$\mathbf{R^2}$	0.27	0.12
N	1203	1203

Note: standard errors in parenthesis. * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table 9 presents the IO and IO-ratio estimates for the educational-adjusted household income (conditional on occupation). Both, absolute and relative IO measures are higher using this definition of income instead of overall income. Thus, the share of IO is 75.5% for the Gini index and 32.4% for the MLD. This result shows that the part of total income explained by education (measured as years of education in our case) varies across individuals mainly because they possess different circumstances, and not because they exert a higher degree of effort. With respect to the Shapley value decomposition, it is observed that the bulk of IO is explained by parental education: 36.8% for the IO Gini and 45.1% for the IO MLD. This result is consistent with the results in Table 7 and the

analysis of relative mobility in education performed in Section 2: individual education is determined at a large extend by parental education.[21] We find a set of additional relevant circumstances: the occupation of the father (explaining about 17%), the type of school (13%-14%), the size of the family (15%), and having cultivated parents (between 9%-11%).

Table 9. IO of educational-adjusted household income (conditioned on education).

	IO (Gini)		**IO (MLD)**	
	Absolute	**Relative**	**Absolute**	**Relative**
Index	0.23	75.54%	0.05	32.35%
Standard Deviation	(0.02)	(0.64)	(0.00)	(0.08)
	Shapley Decomposition			
	Relative Contribution		**Relative Contribution**	
Gender	1.19%		0.12%	
Size Family	15.52%		15.19%	
Health Status	0.06%		0.02%	
Type of School	14.24%		12.83%	
Parental Education	36.78%		45.07%	
Class of the Father	16.62%		16.74%	
Cultivated Parents	11.74%		8.80%	
Environment	3.85%		1.23%	

Note: standard deviations in parenthesis calculated based on 50 bootstrap replications.

Table 10 shows the results for the occupational-adjusted household income. The relative IO of this variable is 56.6% for the Gini and 15.5% for the MLD. With respect to the Shapley value decomposition, the bulk of IO is now explained by parental education and the type of school, each circumstance contributing about 20%. The occupation of the father explains between 16% and 19% and the family size between 18% and 19%. Gender is now significant and contributes to explain occupational income differences by about 7%-9%. The rest of circumstances contribute very little.

[21] By using data from EU-SILC (2011), Requena (2016) shows that education is still an important channel for upward mobility.

Table 10. IO of occupational-adjusted household income (conditioned on occupation).

	IO (Gini)		IO (MLD)	
	Absolute	**Relative**	**Absolute**	**Relative**
Index	0.18	56.66%	0.03	15.49%
Standard Deviation	(0.01)	(0.35)	(0.00)	(0.19)
	Shapley Decomposition			
	Relative Contribution		**Relative Contribution**	
Gender	8.84%		6.91%	
Size Family	18.20%		19.45%	
Health Status	1.54%		2.25%	
Type of School	21.05%		23.42%	
Parental Education	21.06%		19.49%	
Class of the Father	15.55%		19.28%	
Cultivated Parents	10.76%		8.37%	
Environment	3.00%		0.83%	

Note: standard deviations in parenthesis calculated based on 50 bootstrap replications.

6. Conclusions and policy implications

This article has estimated inequality of opportunity for Spain in 2017 using a novel database from the questionnaire CIS-3178 on 'Social inequality and social mobility in Spain'. A clear strength of this database is the inclusion of a wide set of circumstances, such as gender, the size of the household where the individual grew up, health perception during adolescence, type of school, parental education and occupation and, family and urban environments during adolescence.

Our results point out that IO explains an important share of total inequality: 44% according to the Gini index and 18% using the MLD. These values are far above previous IO estimates for Spain. The remaining part of overall inequality is explained by unobserved circumstances, pure effort (the part of total effort that do not depend on the set of circumstances) and luck. Making this distinction is relevant because IO has been found to harm growth (for example by generating misallocation of human capital and talent in the education sector and high-qualified occupations), while the part of total inequality related to effort is growth enhancing due, among other things, to the incentives to develop individuals' talent and to exert higher levels of effort in both the educational sector and the labor market.

Which are the most important circumstances affecting the generation of IO in Spain? And which are the main channels through which circumstances end up affecting income? About the first question, we find that about 90% of IO (regardless the inequality measure under consideration) is due to parental education and occupation, the type of school attended during childhood, the size of the household and the gender of the household's head. With respect to the second question, we show that a large part of IO is channeled through the individual's occupation and education. These findings are consistent with the low levels of relative educational and occupational mobility observed in the data for Spain. In particular, we find that the percentage of individuals with university studies is 64.7% when the father has also university studies, while it reduces to only 16.4% when the father has basic studies. Moreover, the percentage of respondents in the group of managers, administrators, technicians and high-level professionals is 38.2% when the father belongs to that upper-class group, and it is only 14.8% when the father belongs to the unqualified group of workers. Therefore, it seems that reaching the highest educational and occupational ranks is still significantly conditioned by the family origin in Spain.

To end this section, we comment now on a set of public policies consistent with our results that could help to reduce IO in Spain. Nonetheless notice that for a better understanding of the causes and channels of IO in Spain and, therefore, for a more accurate evaluation of the capacity of certain economic policies to promote social mobility and equality of opportunity, we need longitudinal databases and appropriate experiments to evaluate each intervention.

Our findings suggest that early interventions to reduce disparities between people with different gender, family sizes, socioeconomic background, type of school attended during childhood, and home environment would help to reduce IO in Spain. Moreover, because education and occupation are found to be important channels of transmission of IO, in order to reduce IO in Spain, the public sector should implement policies that minimize the impact of the aforementioned circumstances in both the education system and the labor market.

With respect to the education system, an equality of opportunity reform should drastically reduce school dropout and increase secondary enrollments. As school failure in Spain is highly concentrated at the low social strata, reducing drastically school failure will

prevent premature departure from the education system of current children and future parents. It is important to level standards upwards for the quality of college education, and to reduce the gap between public schools and private and semi-private schools (Rodríguez, 2017). Otherwise, children from initially rich families will stay rich because they attend high-quality private and semi-private schools, while children from initially poor families will stay poor because they attend low-quality public schools (Ferreira, 2001).

With respect to the labor market, interventions must focus on reducing barriers of entry and discrimination between groups. Thus, individuals with the same levels of education and skills, but different set of circumstances, must face the same opportunities to entry in the labor market. Moreover, according to our results, increasing public spending on child and health care would help to reduce the disadvantage observed in women and large family households.

A potential extension of our research could focus on the interactions between the education system and the job market. For example, it could be important to analyze the lack of opportunities caused by a mismatch between the labor supply and the labor demand caused by the qualification of workers and the skills required by the job. mproving the offer in secondary education (for example, through a large supply of vocational training cycles) and tertiary education (through new and more flexible degrees and postgraduate courses) could bring the labor demand closer to the necessities of the labor market which, in turn, could significantly reduce IO in Spain.

References

Alesina, A. and Rodrik, D. (1994): "Distributive politics and economic growth". *Quarterly Journal of Economics*, 109: 465-490.

Ayala Cañón, L. (2016). *La Desigualdad En España: Fuentes, Tendencias y Comparaciones Internacionales*. Madrid: FEDEA.

Banerjee, A. and Newman, A. (1993): "Occupational choice and the process of development". *Journal of Political Economy*, 101: 274-298.

Bagües, M. and Esteve-Volart, B. (2009): "Top civil service: meritocracy or nepotism". Unpublished working paper.

Barro, R. (2000): "Inequality and Growth in a Panel of Countries". *Journal of Economic Growth*, 5: 5-32.

Bartholomew, D.J. (1982): *Stochastic Models for Social Processes* (3rd ed.). Wiley, London.

Betancort, M., Darias, S., Cabrera, L., Nieves, C., Marrero, G.A., Rodríguez, J.G., and Sanchez, D. (2019): "Inequality of opportunity in an outermost region: the case of Canary Islands". *Island Studies Journal*, 14: 23-42.

Bradbury, K. and Triest, R. K. (2016): "Inequality of opportunity and aggregate economic performance". *The Russell Sage Foundation Journal of the Social Sciences*, 2(2): 178-201.

Bourguignon, F., Ferreira, F. H., and Menendez, M. (2007): "Inequality of opportunity in Brazil". *Review of Income and Wealth* 53: 585–618.

Brunori, P., Palmisano, F. and Peragine, V. (2019a): "Inequality of opportunity in Sub-saharan Africa". *Applied Economics*, 51: 6428-6458.

Brunori, P., Peragine, V. and Serlenga, L. (2019b): "Upward and downward bias when measuring inequality of opportunity". *Social Choice and Welfare*, 52: 635-661.

Buhmann, B., Rainwater, L., Schmaus, G. and Smeeding, T. (1988): "Equivalence scales, well-being, inequality and poverty: sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database". *Review of Income and Wealth*, 34: 115-142.

Coulter, F. A., Cowell, F.A. and Jenkins, S.P. (1992): "Difference in needs and assessment of income distributions". *Bulletin of Economic Research*, 44: 77-124.

Carabaña, J. (1999). Dos estudios sobre movilidad intergeneracional. Madrid, Fundación Argentaria, Visor.

Chantreuil F. and Trannoy A. (2013): "Inequality decomposition values: the trade-off between marginality and efficiency". *Journal of Economic Inequality*, 11: 83-98.

Chávez-Juárez, F. W. and Soloaga, I. (2014) "iop: Estimating ex ante Inequality of Opportunity". *The Stata Journal*, 4: 830-846.

Checchi, D. and Peragine, V. (2010): "Inequality of opportunity in Italy", *Journal of Economic Inequality*, 8: 429-450.

CIS (Centro de Investigaciones Sociológicas) (2017): *Social Inequality and Social Mobility in Spain*' ('*Desigualdad Social y Movilidad Social en España*'). *CIS 3178*. http://www.cis.es/cis/opencm/ES/2_bancodatos/estudios/ver.jsp?estudio=14350

Dasgupta, P. and Ray, D. (1986): "Inequality as a determinant of malnutrition and unemployment: theory". *Economic Journal*, 96: 1011-1034.

Echeverría-Zabalza, J. (1999): *La movilidad social en España* (1940-1991). Madrid. Istmo.

European Union (2018): Inter-generational mobility in social status, education and occupation. *Special Eurobarometer* 471, December 2017. Summary (Chapter IV). https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/survey/getsurveydetail/instruments/special/surveyky/2166

Ferreira, F. (2001): "Education for the masses? The interaction between wealth, educational and political inequalities". *Economics of Transition*, 9: 533-552.

Ferreira, F. (2007): "Inequality as cholesterol", in *Poverty in Focus: The Challenge of Inequality*. Brasilia: International Poverty Centre (IPC), United Nations Development Programme, 20-21.

Ferreira, F. H. and Gignoux, J. (2011): "The measurement of inequality of opportunity: Theory and an application to Latin America". *Review of Income and Wealth*, 57: 622-657.

Fleurbaey, M. (2008). *Fairness, responsibility, and welfare*. Oxford University Press.

Foster, J. E. and Shneyerov, A. A. (2000): "Path independent inequality measures". *Journal of Economic Theory*, 91: 199-222.

Galor, O. and Zeira, J. (1993): "Income Distribution and Macroeconomics". *Review of Economic Studies*, 60: 35-52.

Goldin, C. and Katz, L.F. (2008): *The Race between Education and Technology*. Cambridge, MA: Harvard University Press.

Goldthorpe, J. H. (2013). "Understanding – and Misunderstanding – Social Mobility in Britain: The Entry of Economists, the Confusion of Politicians and the Limits of Educational Policy". *Journal of Social Policy*, 42 (3): 431-450.

Instituto de Estadística y Cartografía de Andalucía - IECA (2018). Encuesta de Movilidad Social en Andalucía.

International Labour Organization (2012): *International Standard Classification of Occupations. ISCO-08*. https://www.ilo.org/public/english/bureau/stat/isco/isco08/

Kaldor, N. (1956): "Alternative theories of distribution". *Review of Economic Studies*, 23: 83-100.

Lefranc, A., Pistolesi, N. and Trannoy A. (2009): "Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France". *Journal of Public Economics*, 93: 1189-1207.

Li Donni, P., Rodríguez, J. G., and Rosa Dias, P. (2015): "Empirical definition of social types in the analysis of inequality of opportunity: a latent classes approach". *Social Choice and Welfare*, 44: 673-701.

Mankiw, N.G. (2013): "Defending the one percent". *Journal of Economic Perspectives*, 27: 21-34.

Marqués, I. (2015): *La movilidad social en España*. Los libros de La Catarata, Madrid.

Marrero, G. A. and Rodríguez, J. G. (2011): "Inequality of Opportunity in the U.S.: Trends and Decomposition". *Research on Economic Inequality*, 19: 217-46.

Marrero, G. A. and Rodríguez, J. G. (2012): "Inequality of opportunity in Europe". Review of Income and Wealth, 58: 597-621.

Marrero, G. A. and Rodríguez, J. G. (2013): "Inequality of opportunity and growth". *Journal of Development Economics*, 104: 107-122.

Marrero, G. A. and Rodríguez, J. G. (2016): "Inequality…of opportunity and economic performance" in *Economic Mobility* (eds., A. Brown, D. Buchholz, D. Davis and A. González), Federal Reserve Bank of St. Louis and Board of Governors of the Federal, 385-419.

Marrero, G. A., J. G. Rodríguez, and Van Der Weide, R. (2016): "Unequal Opportunity, Unequal Growth". *The World Bank Working Papers Series*, WP no.7853.

Marrero, G.A., Betancort, M., Cabrera, L.J., Darias, S., Pérez, C.N., Rodríguez, J.G. and Sánchez, D. (2017): *Desigualdad de oportunidades y movilidad intergeneracional en Canarias*. Santa Cruz de Tenerife: Fundación Cajacanarias.

Marrero, G. A. and Rodríguez, J. G. (2019): "Inequality and growth: the cholesterol hypothesis". *ECINEQ WP*, 501.

Mirrlees, J. (1971): "An exploration in the theory of optimum income taxation". *Review of Economic Studies*, 38: 175-208.

OECD (2018): *Equity in Education: Breaking Down Barriers to Social Mobility*. PISA, OECD Publishing, Paris. Available at: https://doi.org/10.1787/9789264073234-en.

Palomino, J.C., Marrero, G.A. and Rodríguez, J.G. (2019): "Channels of inequality of opportunity: the role of education and occupation in Europe". *Social Indicators Research*, 143: 1045-1074.

Perrote, I., Rodríguez, J.G. and Salas, R. (2003): "La inequidad horizontal y la redistribución vertical en el impuesto sobre la renta de las personas físicas: un análisis de robustez". *Hacienda Pública Española/Review of Public Economics*, 166: 49-60.

Prais, S.J. (1955): "Measuring social mobility". *Journal of the Royal Statistical Society Series A*, Part I, 118: 56–66.

Ramos, X. and Van de Gaer, D. (2016): "Approaches to inequality of opportunity: Principles, measures and evidence". *Journal of Economic Surveys*, 30: 855-883.

Requena, M. and Stanek, M. (2015): Las clases sociales en España: cambio, composición y consecuencias, en Blanco, A.; Chueca, A. y Bombardieri, G. (dir. y ed.), *Informe España 2015*, 487-518. Madrid. Fundación Encuentro/Centro de Estudios del Cambio Social.

Requena, M. (2016): *La educación como ascensor social*. Dossier Observatorio Social de la Caixa, 19-28.

Rodríguez, J. G. (2004): "Descomposición Factorial de la Desigualdad de la Renta". *Revista de Economía Aplicada*, 12: 25-46.

Rodríguez, J. G. (2008): "Partial equality-of-opportunity orderings". *Social Choice and Welfare*, 31: 435-456.

Rodríguez, J.G. (2017): "Equality of opportunity and inclusive growth" in *Progressive Lab for Sustainable Development* (Eds. B. Caracciolo, Ch. Cheuvart, C. Dragomirescu-Gaina and V. Ntousas), 301-325.

Roemer, J. E. (1993): "A pragmatic approach to responsibility for the egalitarian planner". *Philosophy Public Affairs*, 20: 146-166.

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

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

Savegnago, M. (2016): "igmobil: A Command for Intergenerational Mobility Analysis in Stata". *The Stata Journal*, 16: 386-402.

Shorrocks, A.F. (1978): "The measurement of mobility". *Econometrica* 46: 1013 –1024.

Shorrocks, A. F. (2013): "Decomposition Procedures for Distributional Analysis: a Unified Framework Based on the Shapley Value". *The Journal of Economic Inequality*, 11: 99-126.

Solt, F. (2019): "Measuring Income Inequality Across Countries and Over Time: The Standardized World Income Inequality Database." SWIID Version 8.1, May 2019.

Sommers, P.S. and Conlisk, J. (1979): "Eigenvalue immobility measures for Markov chains". *Journal of Mathematical Sociology*, 6: 253-276.

Stiglitz, J.E. (2012): *The Price of Inequality: How Today's Divided Society Endangers Our Future*. W.W. Norton and Company.

UNESCO, Institute for Statistics, (2012): *International Standard Classification of Education. ISCED-2011*.

Van de Gaer, D. (1993): "Equality of opportunity and investment in human capital". PhD thesis, Leuven: Katholieke Universiteit Leuven.

Appendix: results for the income of the household head

Table A2. Personal Income.

Personal income	Mean	Standard Deviation	Min	Max
All sample	962.75	822.78	1	6001
25-29 years old	763.16	602.93	1	2701
30-34 years old	954.42	830.17	1	5251
35-39 years old	1020.64	837.75	1	6001
40-44 years old	1080.85	761.90	1	3751
45-49 years old	880.07	771.93	1	3751
50-54 years old	1066.96	997.99	1	6001
55-60 years old	910.67	841.56	1	6001

Figure A1. Personal monthly income distribution in Spain (by deciles).

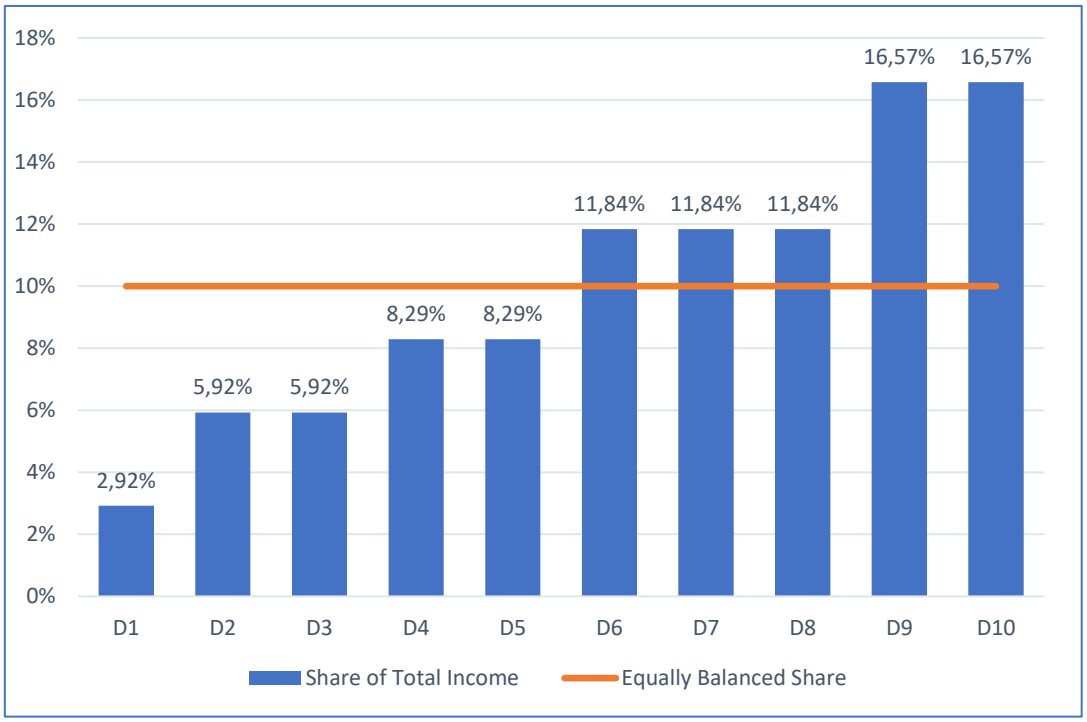


Figure A2. Circumstance's influence over monthly personal income.

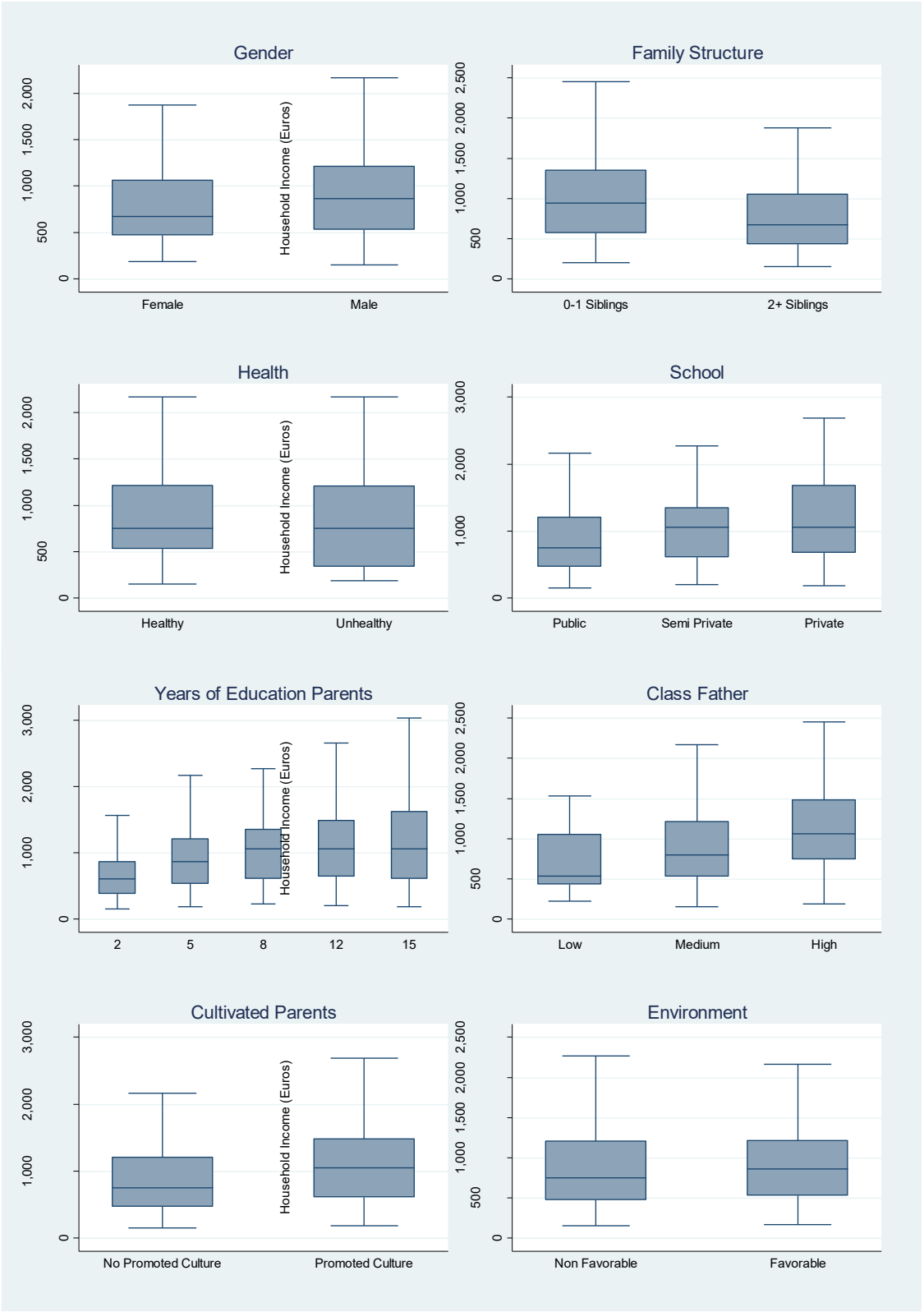


Table A2. OLS regression of personal income on circumstances (alternative models).

	Personal Income of the Household Head					
Gender	0.129**	0.188***	0.236***	0.181***	0.128**	0.241***
	(0.060)	(0.060)	(0.064)	(0.061)	(0.060)	(0.065)
Size family	-0.360***	-0.295***	-0.282***	-0.333***	-0.356***	-0.280***
	(0.059)	(0.056)	(0.062)	(0.060)	(0.059)	(0.063)
Health status	-0.119	-0.094	-0.210	-0.184	-0.120	-0.216
	(0.263)	(0.269)	(0.299)	(0.258)	(0.265)	(0.299)
Semi – private school		0.191**	0.128			0.129
		(0.089)	(0.101)			(0.103)
Private school		0.386***	0.340***			0.339***
		(0.087)	(0.091)			(0.092)
Parental education			0.025***			0.025***
			(0.008)			(0.009)
Father medium class			0.097			0.099
			(0.096)			(0.097)
Father high class			0.149			0.149
			(0.133)			(0.133)
Cultivated parents				0.236***		0.012
				(0.067)		(0.073)
Environment					0.070	-0.024
					(0.064)	(0.068)
Constant	6.882***	6.564***	6.457***	6.778***	6.828***	6.468***
	(0.055)	(0.061)	(0.112)	(0.063)	(0.074)	(0.116)
R^2	0.08	0.12	0.19	0.11	0.09	0.19
Observations	580	580	580	580	580	580

Note: standard errors in parenthesis. * p<0.1; ** p<0.05; *** p<0.01.

Table A3. Relative contribution of each circumstance.

	IO (Gini)		**IO (MLD)**	
	Absolute	**Relative**	**Absolute**	**Relative**
Index	0.19	59.94%	0.07	20.99%
Standard Deviation	(0.01)	(1.02)	(0.00)	(0.84)
	Shapley Decomposition			
	Relative Contribution		**Relative Contribution**	
Gender	15.88%		20.12%	
Size Family	13.11%		11.63%	
Health Status	7.37%		9.23%	
Type of School	26.50%		31.75%	
Parental Education	17.70%		15.30%	
Class of the Father	5.08%		2.55%	
Cultivated Parents	8.41%		5.87%	
Environment	5.95%		3.55%	

Note: standard deviations in parenthesis are based on 50 bootstrap replications.

Table A4. Adjusted income by education and occupation on circumstances.

	Model 1 - Personal Income of the Household Head (education)	Model 2 - Personal Income of the Household Head (occupation)
Gender	-0.053***	0.022
	(0.018)	(0.023)
Size Family	-0.085***	-0.077***
	(0.019)	(0.018)
Health Status	-0.033	-0.132**
	(0.064)	(0.067)
Semi – Private School	0.066***	0.064**
	(0.023)	(0.028)
Private School	0.072***	0.064**
	(0.022)	(0.030)
Parental Education	0.014***	0.003
	(0.002)	(0.003)
Father Medium Class	0.084**	0.061**
	(0.037)	(0.028)
Father High Class	0.132***	0.097**
	(0.045)	(0.042)
Cultivated Parents	0.042**	0.032
	(0.018)	(0.023)
Environment	0.011	-0.007
	(0.022)	(0.021)
Constant	6.403***	6.031***
	(0.041)	(0.038)
$\mathbf{R^2}$	0.37	0.15
N	580	580

Note: standard errors in parenthesis. * $p<0.1$; ** $p<0.05$; *** $p<0.01$.

Table A5. IO of educational-adjusted household income (conditioned on education).

	IO (Gini)		IO (MLD)	
	Absolute	**Relative**	**Absolute**	**Relative**
Index	0.23	73.62%	0,19	59.99%
Standard Deviation	(0.00)	(0.87)	(0.00)	(0.16)
	Shapley Decomposition			
	Relative Contribution		**Relative Contribution**	
Gender	7.37%		5.09%	
Size Family	11.99%		11.00%	
Health Status	1.53%		0.63%	
Type of School	15.62%		15.16%	
Parental Education	31.20%		34.70%	
Class of the Father	17.24%		21.23%	
Cultivated Parents	13.66%		11.82%	
Environment	1.39%		0.37%	

Note: standard deviations in parenthesis are based on 50 bootstrap replications.

Table A6. IO of occupational-adjusted personal income (conditioned on occupation).

	IO (Gini)		IO (MLD)	
	Absolute	**Relative**	**Absolute**	**Relative**
Index	0.19	61.06%	0.05	17.09%
Standard Deviation	(0.01)	(0.54)	(0.00)	(0.58)
	Shapley Decomposition			
	Relative Contribution		**Relative Contribution**	
Gender	5.90%		5.32%	
Size Family	19.61%		21.87%	
Health Status	2.39%		1.60%	
Type of School	24.63%		27.92%	
Parental Education	19.56%		16.46%	
Class of the Father	15.97%		20.29%	
Cultivated Parents	9.21%		5.89%	
Environment	2.73%		0.65%	

Note: standard deviations in parenthesis are based on 50 bootstrap replications.