CULTURE CONSUMPTION AFTER RETIREMENT IN SPAIN

Eduardo Polo Muro

Supervisor: Javier Gardeazabal

Abstract

This paper studies culture consumption after retirement. Using data from Spain, for the period 2006-2014, we analyze changes in culture consumption expenditures. To measure these changes, we use a Regression Discontinuity design. This paper exploits the discontinuity on the probability of retirement at age equal to 65 to draw statistical inference. The endogenity of retirement requires using instrumental variables to estimate the effect of retirement on cultural goods demand. We show that some testable necessary conditions are fulfilled. We find a causal effect of retirement on culture consumption. This effect is a fall in the demand of culture after retirement for retired individuals. A placebo test provides internal validity to our results.

Keywords: Fuzzy-regression-discontinuity retirement potential-outcomes causal-inference culture-consumption

Consumption theory has been one of the most fruitful topics in economics. This field has been covered extensively both in theoretical and empirical studies. Durable and non-durable goods have also been extensively analyzed as they cover the basics needs of individuals. Thus, consumption theory has provided a huge number of papers of very high relevance. Some of them have also analyzed the relationship between retirement and changes in consumption. Banks, Blundell and Tanner (1998) analyzed whether British households save enough to keep constants their consumption after retirement, using data from 1968 to 1992. Browning and Lusardi (1996) analyzed household savings in the framework of a Life-cycle model for US households, using data from 1962 to 1996. Bernheim et al (2001) analyzed how the wealth of the household changes after retirement in the US, using data from 1987 to 1990.

For the Spanish case, Luengo-Prado and Sevilla (2012) exploited the 1985 to 2004 waves of the Spanish expenditure survey. Linking these works related to the consumption field to causal inference, Battistin et al (2009) used a regression discontinuity design to estimate the causal effect of retirement on consumption. Regression discontinuity design has also been used to analyze the effect of financial aid or class size on college outcomes (Van der Klaaw, 2002 or Angrist and Lavy, 1999).

Using the data provided by INE in the Household Expenditure Survey, we find an average annual drop of around 166€-175€ in culture consumption after retirement. This drop can be explained, at least partially, by an increase in the number of individuals who drop their culture consumption totally. Our data and procedure pass some validity checks. This gives more strength to our results. We also find that, for those individuals that do not attend college, the drop is around 102€. This results is significant at 10% confidence level. For those individuals with at least this level of education, we do not find a significant change in their behavior.

This paper is divided as follows: Section 1 explains the Regression discontinuity idea, covering the 2 main cases, sharp RD and fuzzy RD as well, and the relationship that these designs have with potential outcomes framework. We also cover the main propositions to estimate the causal effects in these frameworks. Section 2 is devoted to explain how retirement can be understood as a treatment, and also to explain briefly the baseline ideas about retirement in Spain. Section 3 is mainly devoted to describe our data sources and also their most relevant characteristics. Section 4 explains the estimation methods, and presents the result. We also show that our findings are consistent with those reported in the literature. We also show that our data fulfills the basic assumptions needed in the RD design. This section also includes a robustness check. Finally, in Section 5 we add another estimation using 2 subsamples from our sample, using the educational level as to split it.

1 Regression discontinuity design In potential outcomes framework

Regression discontinuity (RD) design has become a useful tool for researchers in the last few years, since Thisthlewaite and Cook (1960) introduced it in statistics literature. Several papers, as Van der Klaauw (2002), Angrist and Lavy (1999), Battistin et al (2009) or Lee (2008) have exploited RD design to estimate the causal effect of a certain treatment on an outcome of interest. This paper links RD design with potential outcomes framework (Rubin, 1974), where, following the usual notation, we have 2 potential outcomes, Y_1, Y_0, but only one of them is observable for each individual, depending on whether the individual is treated or not. These 2 potential outcomes are related in this way

$$Y_i = \begin{cases} Y_{1i} & if\ D_i = 1 \\ Y_{0,i} & if\ D_i = 0 \end{cases} \tag{1}$$

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i})D_i \tag{2}$$

Where D_i is a binary indicator, which takes value equal to 1 if individual i is treated, and 0 otherwise. So, we can observe either Y_1 or Y_0, but not both for the same individual. Thus, we are not allowed to estimate the causal effect of treatment at individual level, but we can estimate the Local Average Treatment Effect ($LATE$, henceforth), applying RD design as we are going to explain.

1.1 RD design

This design is based on the idea that individuals receive some treatment if a covariate, say x, exceeds a known threshold, and they do not receive treatment otherwise. Thus, we have

$$D_i = \begin{cases} 1 & if\ x_i \geq x_0 \\ 0 & if\ x_i < x_0 \end{cases} \tag{3}$$

Where x_i is the observed value of the key variable, and x_0 is the known cutoff point. A typical example is the case of fellowships, which are granted to students when they overpass a previously determined mark. This case is known as sharp RD, and only individual over the cutoff point are allowed to be treated. Thus, treated individuals, have an outcome equal to Y_1, and untreated individuals, have an outcome equal to Y_0.

Fuzzy RD

It may be the case that treatment is not mandatory for individuals even when they surpass the threshold, or attainable for some individuals who have not overran cutoff point. This situation supposes that some individuals above the threshold do not revive treatment, and some individuals under it actually receive treatment. Despite this fact, the probability of receiving treatment must be higher for individuals over the threshold point. Thus, the framework that we need in a Fuzzy RD is

$$P\left(D_i = 1 \,|\, x_i\right) = \begin{cases} F(\alpha_1 + \beta_1(x_i - x_0)) & if \quad x_i \geq x_0 \\ F(\alpha_0 + \beta_0(x_i - x_0)) & otherwise \end{cases} \tag{4}$$

where $F(\alpha_1 + \beta_1(x_i - x_0)) > F(\alpha_0 + \beta_0(x_i - x_0))$. Fuzzy RD design usually appears when the individuals have some influence over their own treatment status, but they are not totally allowed to manipulate it, which lead us to a case in which the probability of being under treatment is lower than 1 for those are over the threshold, and higher than 0 for those are under it. Fuzzy RD design needs a discrete change in probability of being treated at the cutoff point.

1.2 Estimation

The equation of interest is

$$Y_i = \eta + \tau D_i + \gamma' X_i + \varepsilon_i \tag{5}$$

where Y_i is the outcome of interest of the individual i, τ is the LATE, D_i is already defined, X_i is a vector of covariates and ε_i is the error term. Hahn et al (2001), in one of the most relevant papers in this field, shown that, under some conditions, RD design can be estimated using Wald estimator

$$\tau = \frac{lim_{x\downarrow c} E\left[Y_i \mid x_i = x\right] - lim_{x\uparrow c} E\left[Y_i \mid x_i = x\right]}{lim_{x\downarrow c} E\left[D_i \mid x_i = x\right] - lim_{x\uparrow c} E\left[D_i \mid x_i = x\right]} \tag{6}$$

where D_i is the treatment indicator. This Wald estimatior is the ratio of the differences in outcome between those individuals just over the threshold and just under it, and the differences in the probability of treatment just over the threshold and just below it. That is why we refer to the average treatment effect (ATE) as local, due to our analysis relies on the differences between the observations around the threshold (locally). This also implies that our result has a limited external validity: only can be applied for those we call compliers, individual who want to be treated when they are eligible and do not want it when they are not eligible.

The Wald estimator is a valid estimator for the local Average Treatment Effect when the following assumptions are fulfilled (Hahn et al (2001) and Lee & Lemieux (2009)) :

- $E\left[\tau_i \mid x_i = x\right]$, is continuous at x_0.This means that the treatment effect must be the same for individuals just around the threshold, if they actually receive treatment.
- $(\tau_i,\ D_i(x))$ is jointly independent of x_i near x_0. Likelihood of receiving treatment around the threshold is the same for all individuals. These is specially important, because this independence condition is what ensures that treatment is as good as randomly assigned, and what makes RD design a good framework for causal inference.
- Monotonicity is assumed. Those who want to retire when they are not eligible must keep their intention when they are actually eligible as well.
- Outcome (and other covariates) must be continuous at x_0, what means that there are no jumps in the covariates and outcome apart from the forcing variable.

Hahn et al (2001) established the link between IV procedure and Wald estimator to measure the treatment effects. Angrist & Pischke, in their book “Mostly Harmless Econometrics” (2009) also showed that RD design leads to 2 stages least squares (2SLS) estimation using a polynomial function. In the first step, we should estimate a regression of the form

$$D_i = \delta + \theta_1 z_i + \theta_2 x_i + \theta_3 x_i^2 + ...\theta_{p-1} x_i^p + u_i \tag{7}$$

where z_i is the variable used to instrument the endogenous treatment status, and x_i, x_i^2, ..., x_i^p are the powers of the covariate. The eligibility rule,

$$z_i = 1\left(x_i \geq x_0\right) \tag{8}$$

is the instrument.

2 Retirement status

2.1 Retirement as treatment

This paper analyzes the causal effect of retirement decision on culture consumption. Thus, we think of retirement as a treatment. So, there are 2 subsamples, retired and non-retired individual, or treated and untreated individuals. We use a binary variable, called R_i, which takes value equal to 1 when individual i is retired and 0 otherwise. Eligible individuals are those where age is above 65, $Z_i = 1(s_i \geq 0)$, where s_i refers to distance to/from cutoff point, that is,$s_i = Age_i - 65$.

2.2 Retirement in Spain

In 1994 Spanish government passed an important act about retirement, pensions and social security, which should be consider as the baseline over what new laws rely. One of the most important changes in this act happened in 2002, and another one in 2013. However, as the data runs from 2006 to 2014, the law that affects our data is law 35/2002. We detail the main features of the retirement in Spain after this act come into force. Retirement age in Spain was still 65. However, retirement is not mandatory, but working after this age gives the individuals some fiscal advantages, and a higher pension, when they actually retire. Moreover, individuals can retire since they are 61, or even before, under some conditions.

Working further than 65

- If a worker asked for the retirement after 65, he/she is entitled to a 100% of the regulatory base, plus and extra 2% more for each additional year of work.
- Workers and employers do not have to pay social security contribution.

These advantages apply only in case that the worker has contributed for at least 35 years.

Early retirement

- Age equal or greater than 61.
- Apply for a job in the SEPE (the name of Spanish employment office) during, at least, 6 month before asking for retirement
- Having, at least, 30 years of contribution.
- That the worker did not quit the job voluntarily

These conditions are for general arrangement workers. In other sectors, for instance, hazardous activities, retirement at 61 is allowed without the penalties that we are going to detail below. In case of early retirement, for each remaining year until the standard retirement age, the pension will be decreased in the next proportion:

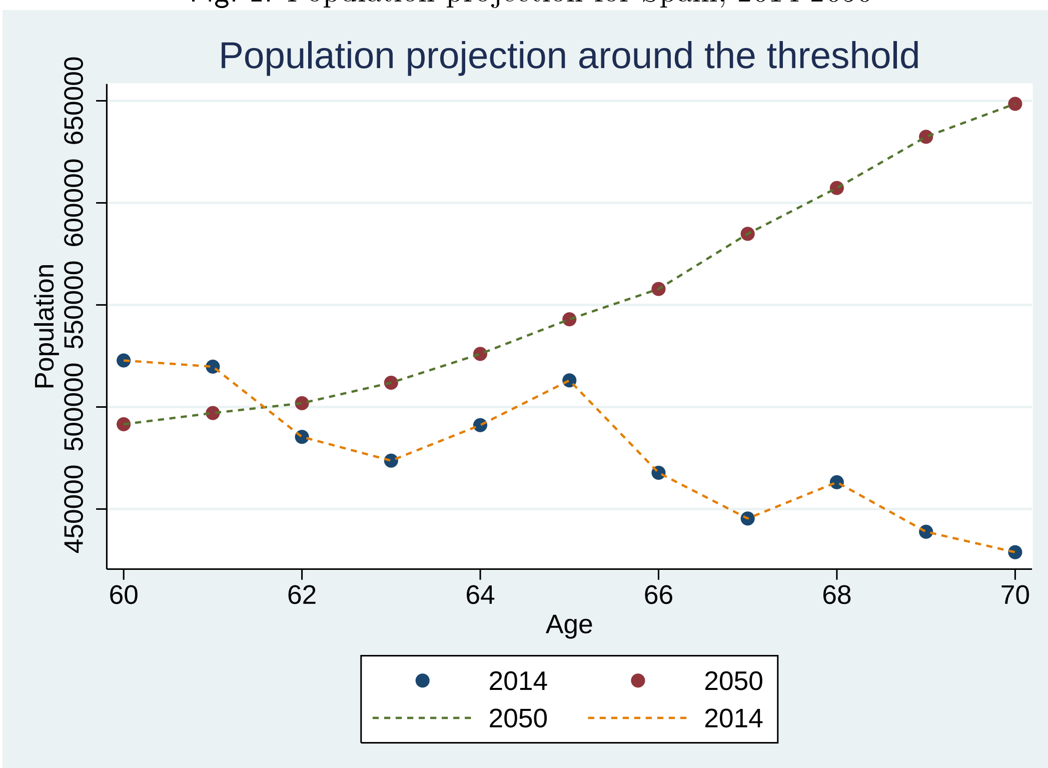
- With 30 years contributed: 8% each year
- With at least 31 and a maximum of 34 years contributed: 7.5% each year
- With at least 35 and a maximum of 37 years contributed: 7% each year
- With 38 or 39 years contributed: 6.5% each year
- With 40 or more years contributed: 6% each year

Retirement in Spain, a controversial issue

Retirement payment is a controversial issue in Spain, due to some remarkable factors: The aging population, shown in Figure (1), reduces the labor force, and increases the potential receivers for payment. The labor market in Spain, with a very high unemployment rate (see Figure (2)). These are ingredients threatening the sustainability of the pension system. To these factors, we should add one more: Large number of individuals choose to retire early, as we will see in figure 3.

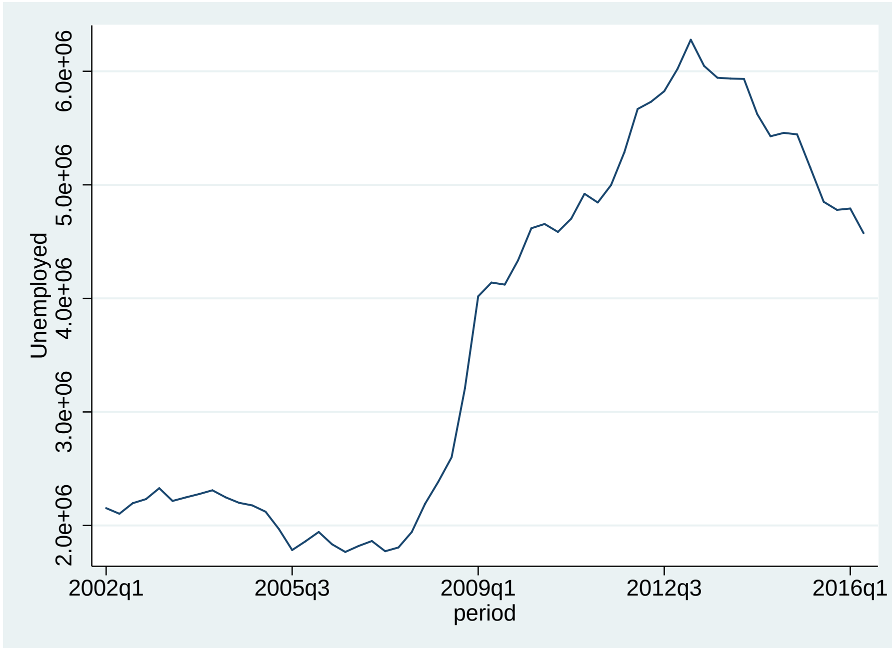
Some data about retirement status in Spain are provided in next section, fully devoted to analyze our data.

Fig. 1: Population projection for Spain, 2014-2050



Source: INE.Population projection for Spain, 2014-2016. Population around 65 with a bandwidth of -5/+5.

Fig. 2: Unemployment in Spain



3 Data

3.1 - Household expenditure survey

Methodology

INE (Instituto Nacional de Estadistica) provides, since 2006, a new model of *Household Expenditure Survey* (HES, henceforth), adjusting it for new purposes and recommendations from EUROSTAT. The units of analysis are households, divided in different strata, that are representative of a determined number of families.

The household data is gathered using a mixed method of personal survey and the use of notebooks. Each member of the household uses an expenditure notebook, where they must record each one of the expenses that they do. Households also have a notebook for common expenses. The personal survey covers information about household's expenses and a wide range of variables, such as education level, income, household size, labor market status and so on. For more detail, these surveys are available in INE's webpage.

Data

INE provides some reports with the most relevant data, but the microdata is available as well. These microdata are provided in 3 files for each year: household, members, and expenditure. The household file provides information about the common characteristics for the household members (For instance, house size), and about the head of the household. Members file, instead, supplies information about each member of the household, and lastly, expenditure file which gives information about the expenditure made by each household on performing arts, museums / exhibitions and books -among many others-. It is remarkable the fact that expenditures in each good are raised to population and temporal factors: each expense shows the aggregate expense for whole year of all the families that are represented by this household.

We use data from the 2006 to 2014 waves. Table (1) shows how many households were interviewed each year, and how many households are represented. We consider our data as pooled cross sections, despite they come from different years. In our final dataset, we have, for each household, it common characteristics, all the individual data as covariates, and each of the expenditures.

Tab. 1: Nº of Surveys and represented population

YEAR	**Nº of Surveys**	**Represented Population**
2006	19435	16179780
2007	21542	16643003
2008	22077	17067748
2009	22346	17384274
2010	22203	17644384
2011	21680	17897736
2012	21808	18091838
2013	22057	18212214
2014	22146	18303178

3.2 Identification strategy

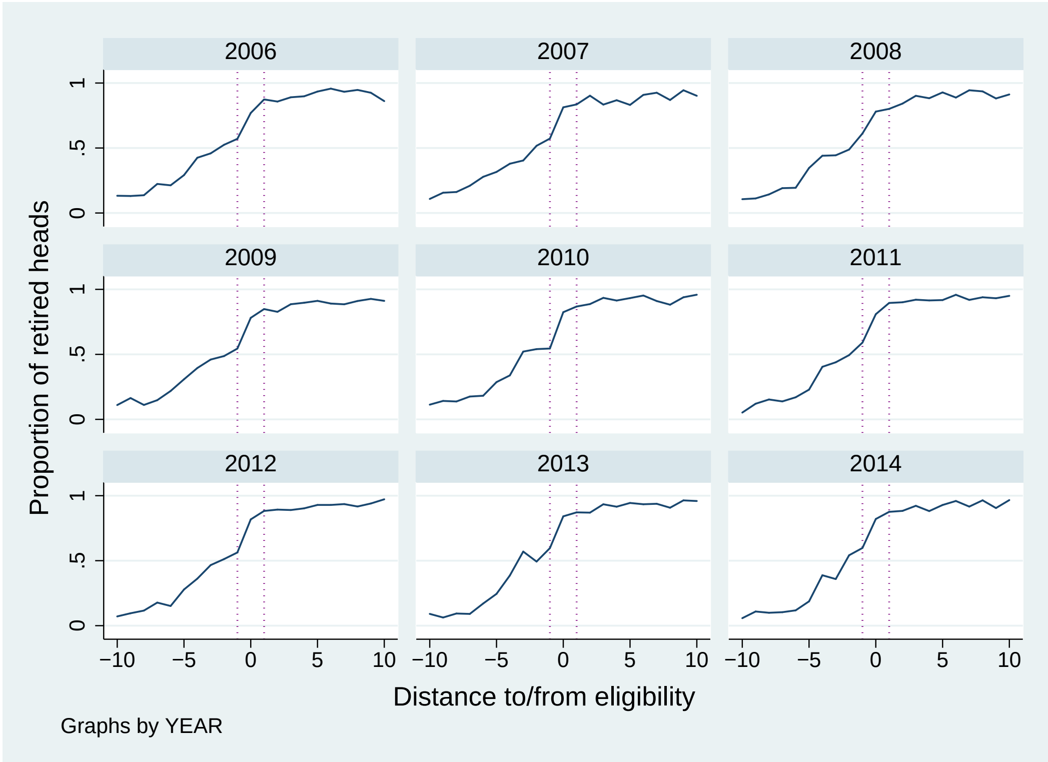
In order to assess the causal effect of retirement on culture consumption using a RD design, we should isolate a subsample around the threshold. From Section 1 we know that we need to establish the bandwidth which we are working with. This bandwidth is 10 years, the same that Battistin et al (2009) proposed, so we select individuals from 55 to 75 years old. Our analysis focuses on males, as the share of retired women is very low to consider it representative, as we can see in Table (2). To compute this table, we have used only men and women that self reported as head of the household or the couple of the head, and weight it with factor variable, so not entire population is taken into account in this table.

Tab. 2: Labor Market Status, by Year and Gender

	Male		Female	
YEAR	**Non-Retired**	**Retired**	**Non-Retired**	**Retired**
2006	0.543	0.198	0.201	0.055
2007	0.535	0.192	0.214	0.057
2008	0.537	0.188	0.212	0.061
2009	0.520	0.190	0.228	0.061
2010	0.513	0.190	0.235	0.061
2011	0.500	0.192	0.246	0.060
2012	0.490	0.195	0.254	0059
2013	0.474	0.199	0.262	0.063
2014	0.467	0.199	0.269	0.062

Thus, we select only households where males reported being head of the household, or households where males reported being the couple of the female household head. Once we have identified these heads, we compute the distance to/from the threshold, which is 65 years old in this case. Thus, we have a distance variable that goes from -10 to +10. Eventually, we identify the labor market status of each head. We consider that all non-retired heads are untreated, regardless whether they

Fig. 3: Proportion of retired heads



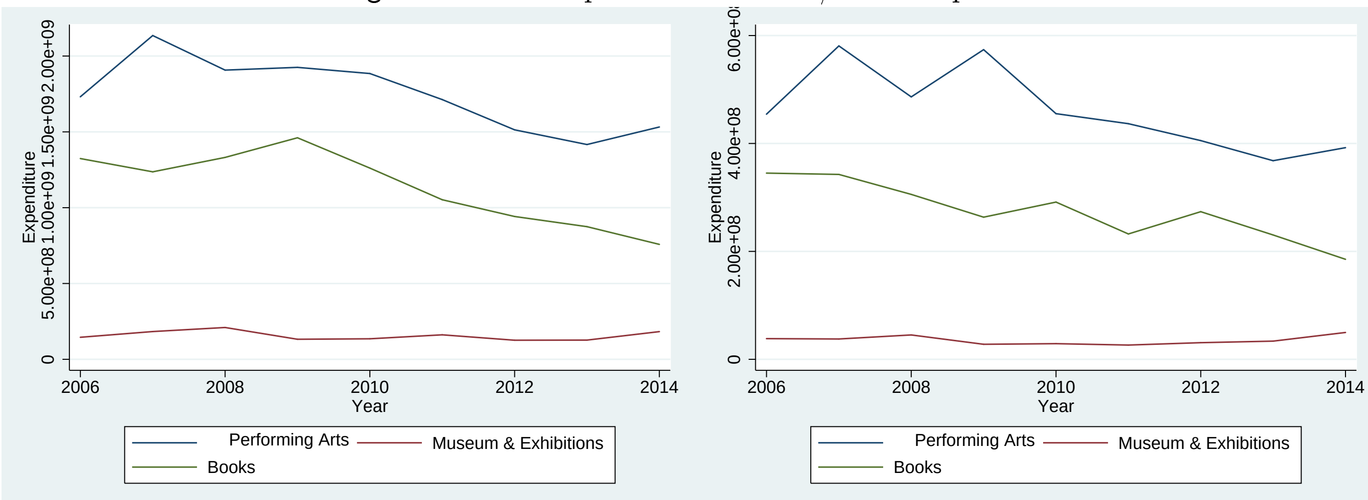
are employed, unemployed, or any other status, and only heads reporting retirement status are consider as treated. The number of households under these conditions is 54276.

Figure (3) shows the share of retired male heads over total male heads as a function of the distance to 65.

3.3 Expenditure data

Following COICOP (Classification Of Individual Consumption by Purpose), INE provides a section devoted to leisure and culture. For our analysis, we have put together 3 subsections: Performing arts, museums and exhibitions and books. Figure 4.a shows the evolution of total expenditure in these areas throughout years, and figure (4).b shows the same evolution but only for our heads subsample. Thankfully, INE does not change the questions of the survey in the section devoted to culture consumption across our waves, which makes easier our work.

Fig. 4: Culture expenditure: Total / Subsample



To better understand our estimations results, Table (3) shows, for a subsample of all households included in the selected bandwidth, and by labor status, how many people spent at least 1€ in any of these 3 types of culture, and how many do not. We also show, for those expenditure in culture, regardless the amount, the mean value of the expenditure. These values are computed weighting with the Factor variable, which indicates how many households are represented by each observation.

Tab. 3: Number of Households with culture expenditure and average expenditure

	Consumer		NonConsumer	
YEAR	**Nº of Households**	**Expenditure**	**Nº of Households**	**Expenditure**
Retired	7513406	364.95	13945560	0
NonRetired	8059703	521.63	8963708	0

We can appreciate in table (3) that, whereas the number of workers that consume culture and the number that do not remain similar, the retired population that do not consume culture is larger than retired population that do. This means that, when people are retired not only the consumption decreases, as we will see in section 4, but there exist a large number of individuals that reduce they consumption in 100%.

4 Estimation and Results

4.1 Estimation

Following Battistin et al (2009), we calculate for each year and each distance the average value of each one of these covariates, (hence, we call "cohort" to each one of the observations compounded by each year and distance to/from the threshold. As our data comes from 9 waves, and the established distance is -10/+10, dropping distance equal to 0 -explained later why-, we will have 180 cohorts). As the treatment indicator is binary, taking values equal to 1 or 0, the average value of treatment indicator for each cohort is equal to the share of retired individuals. We drop distance equal to 0 due to, for those who retire, it covers a year under both labor market status.

The regression that we compute is

$$Y_{t,s} = \beta_0 + \beta_1 R_{t,s} + \beta_2 s_t + \beta_3 s_t^2 + u_j \tag{9}$$

where $Y_{t,s}$ is the average outcome for each cohort (t and s subindex refers to year and distance. The combinations of each distance with each year are called cohorts). $R_{t,s}$ is the average value of treatment indicator, and s_t is the distance to/from the threshold for each year. We also include dummy variables for years. Due to the (limited) ability of individuals to manipulate their status, we need and instrument, that cannot be manipulable, to run a 2SLS. These instrument is denoted as $z_{t,s}$, and is the eligibility status (individuals / cohorts that over-passed the age of retirement), and must fulfill condition explained in (8). We use 2 different method to estimate the average value of the probability indicator. Before showing our results, we describe the methods:

Method 1 - OLS by cohort

First step of 2SLS is running a regression of the form

$$R_{t,s} = \delta_0 + \delta_1 z_{t,s} + \delta_2 s_t + \delta_3 s_t^2 + v_{t,s} \tag{10}$$

where $R_{t,s}$ is the average value of retirement indicator by cohort, $z_{t,s}$ is the instrument that we are going to use for estimating by Instrumental Variables, in this case, the eligibility rule (whether an individual over-passed the threshold), and s_t is the distance from/to the threshold (Thus, it is clear that we include the distance and the square of the distance). We also include dummy variables for years. This regression allows us to estimate the average probability of retirement for each cohort.

Thus, equation (10) is the first stage of 2SLS, and (9) the second one. This first stage gives us a R^2 equal to 0.97, and the instrumental variable and distance variables are significant at 5% confidence level. Thus, we run an IV regression as in (9), instrumenting $R_{t,s}$ with $z_{t,s}$.

Method 2 - Logistic model for individuals

This method is slightly different from the other one. The first step we estimate the probability of treatment for each observation, running a logit model of the next form

$$E(R_i \mid z,\, s) = \Lambda(\theta_0 + \theta_1 z_i + \theta_2 s_i + \theta_3 s_i^2) \tag{11}$$

where each variable is the same as in method 1, but R_i, is the observable labor market status for household head i. We also use observed values at individual level instead of average values for cohort for the rest of covariates. We include dummy variables for years as well. This regression is estimated over the entire sample, instead of the remained subsample after the identification strategy. We do this because we consider that, the larger the sample, the better the estimation.

Once we compute the probability of retirement for each individual, we estimate the average values of the variables that we are including in our model, as we have done in method 1 (average values for the outcome, for the labor market status and the probability of retirement for each cohort, by year and distance). But now, we use these average fitted values from (11) as instrument, following Angrist and Pischke (2009). This procedure is implemented when we have a dummy endogenous variable and we think that the Conditional Expectation Function is nonlinear.

The rest of covariates are the same as in (10). It is also remarkable that the instrument cannot be now a regressor. Thus, we can guarantee that our estimated probabilities are not correlated with residuals when we run the second part of our 2SLS. This second part of this method consist of running (9) using as instrument for $R_{t,s}$ average fitted values from (11). This first stage of this method gives us a R^2 equal to 0.98, and the instrumental variable and 2 distance variables are significant at 5% confidence level.

4.2 Results

Table (4) shows the result of estimation equation (9), using method 1 and method 2. We show the estimated values of the parameters (omitting year dummies, for simplicity), and between parentheses we include standard error. Stars indicate the confidence level. At the bottom of the table we included the number of observations and the value of the R^2. Both estimates show the same trend: after retirement, individuals drop their culture consumption, and following our results, this drops is about 166€-175€, being this drop annual. In both cases, results are significant at 5% confidence level, and R^2 is equal to 0.558 in both cases. Notice that the fall in culture consumption is not uniform, a large number of individuals drop completely their culture consumption. We have to remind that these estimations should have good internal validity, but limited external validity. It applies for compliers, that is, for those individuals who retire when they are eligible. It is also important to note that culture is not a necessary good, whose demand can be very elastic to changes on income, and the taste for culture has great influence in its consumption. Thus, we could say that culture consumption drops after retirement, maybe because a bad saving plan during the working age, or maybe because a change in preferences.

Our results are in line with those found in the literature. Battistin et al (2009) found a drop of 9.8% in non-durable goods and 14% in durable goods after retirement, when they applied RD design, and around 169€ - 241€ when they exploited a panel data on a determined subsample coming from the same sample. Paulin and Duly (2002), found that pre-retired men expend significantly more ($311) than retired men ($178) in leisure goods, using data from US, what is probably the research more similar to what we have done. For Spain, Luengo-Prado and Sevilla (2014) found a drop in consumption between 1,8% and 13%, depending on the kind of good. They also found that some of their results were not significant, but they provided evidence against a constant expenditure after retirement. Beblo and Schreiber (2010) studied German household data, but focused on the expenditure in housing after retirement, with an interest population bounded between 55 and 75 years. They also concluded that retirement causes a drop in this expenditure. Smith (2005) use

Tab. 4: Estimation results

	(1) Method 1	(2) Method 2
causal effect	-175.3**	-166.3**
	(-2.06)	(-2.24)
distance	-0.393	-0.852
	(-0.09)	(-0.23)
dist. squared	-0.370	-0.352
	(-1.38)	(-1.43)
_cons	252.9***	246.8***
	(4.10)	(4.53)
N	180	180
R^2	0.558	0.559

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

data from BHPS to analyze the changes in spending after retirement for some groups (different age of retirement, or whether they retired voluntary or not), and their results show again a drop in expenditure.

4.3 Internal validity check

As we have said in section 1, RD design identifies the causal effect under some assumptions. In this section, we check whether conditions are met.

Discontinuity in probability of being under treatment

As we have already said in section 1, the probability of being under treatment must have a jump at the threshold. We have plotted in figure (3) the share of treated individuals by years for all distances to/from the cutoff point, but now we propose 2 models, 1 linear and the other nonlinear, to estimate the probability of being treated.

Method 1 - OLS by cohort

Equation (10), used for the first stage of the 2SLS, can be used to estimate the probability of retirement for each cohort. In this method, we assume a linear model but introducing an square of the distance to obtain a better fit of the model. Thus, fitted values from this regression, and observed values, are plotted in figure (5).a .

Method 2 - Logistic model for individuals

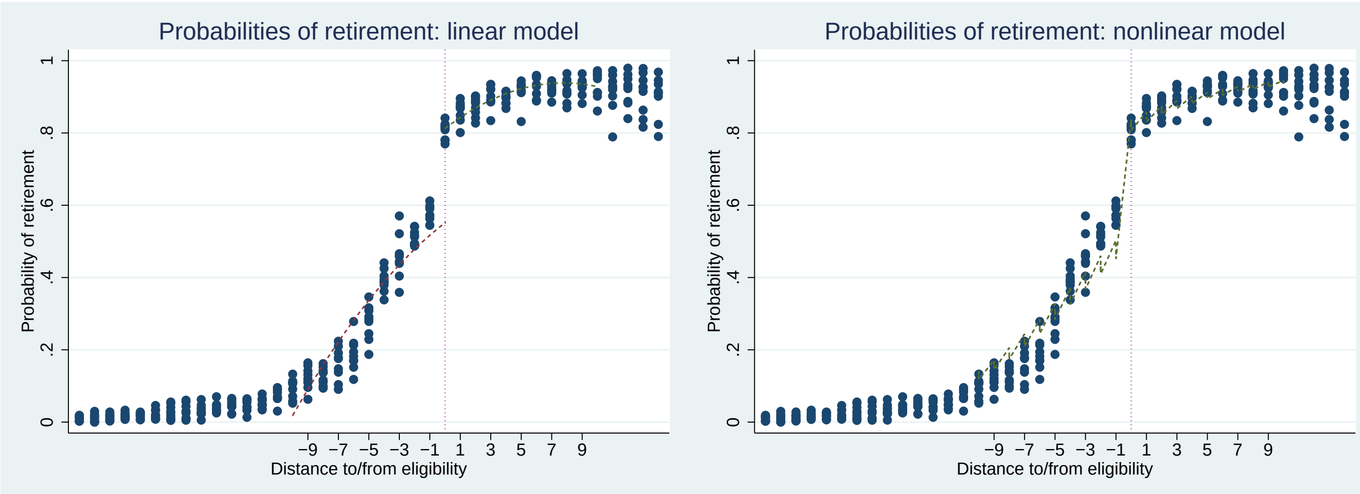
To estimate the probabilities in method 2, we run a regression like in equation (11). Once we have obtained fitted values from this regression, we calculate the average value of the probability of retirement by cohort. Thus, in this method first the probability at unit level is estimated, and then are computed the cohort average values.

We show, in table (5) the average probabilities for each distance, using each method. We only show values for -5/+5, due to the highly important issue here is looking at the jump in probabilities. Figures (5).a and (5).b show the same result graphically, which makes easier for us to understand the idea.

Tab. 5: Probability of retirement, by Distance

Distance	**Linear Model**	**Non-Linear Model**
-5	0.335	0.300
-4	0.386	0.342
-3	0.433	0.385
-2	0.476	0.428
-1	0.515	0.470
0	0.810	0.819
1	0.840	0.841
2	0.867	0.859
3	0.889	0.875
4	0.906	0.889
5	0.920	0.901

Fig. 5: Probabilities of retirement: Linear / Nonlinear model



Smoothness in outcome

As Lee (2008) shows, the jump should only appear in the indicator variable, but not in the outcome. Following the same procedure used in method 1, we have estimated by OLS the following regression

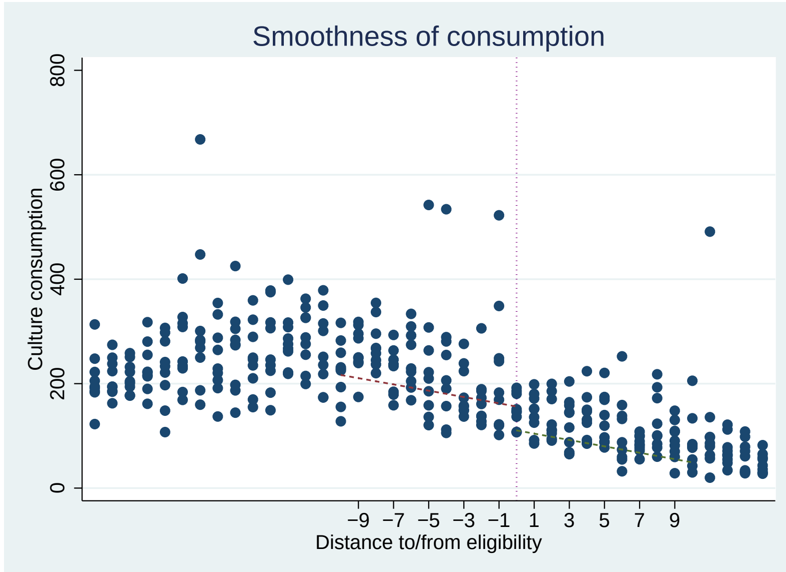
$$Y_{t,s} = \vartheta_0 + \vartheta_1 z_{t,s} + \vartheta_2 s_t + \vartheta_3 s_t^2 + \varepsilon_{t,s} \tag{12}$$

where $Y_{t,s}$ is the average expenditure in culture for each cohort, and the rest are the same covariates used in (10). Figure (6) shows observed and estimated outcome, where, clearly, we did not appreciate any jump. Thus, $E\left[Y_0 \mid X = x\right]$ and $E\left[Y_1 \mid X = x\right]$ are continuous in x at x_0 (Lee , 2008).

4.4 Robustness check

Following the robustness check proposed by Lee (2008), and also used by Battistin et al (2009) we are going to repeat the same procedure implemented in (9), but only with the method 1 (We have already seen that the results do not differ much between methods) but now we are going to use as outcome a set of covariates that are determined before the eligibility status, that is, that cannot be affected by the eligibility status (first condition for selection of this covariates). The

Fig. 6: Smoothness in consumption



second condition is, as explained by Battistin et al (2009), that these covariates must affect the consumption level. Following this idea, we select as outcome some educational levels (as binary covariate), age, residence area type, number of rooms of the house, and home size. For simplicity, we report only some of the estimated coefficient of the treatment indicator and their standard errors.

Shown variables are

- If individual has high studies (university at least) or not (dummy variable)
- If individual has Spanish citizen or not (dummy variable)
- If the residential are is urban and mid-class area. (dummy variable)
- Number of rooms. (Continuous variable)

What we should expect here is that retirement has not affected ex-ante determined variables. A significant drop would invalidate of our method. However, what we see is that not even one of these predetermined outcomes are affected by retirement, what gives credibility to our procedure.

Tab. 6: Robustness check

	(1) Education	(2) Cityzen	(3) Resid. area	(4) Nº rooms
causal effect	0.00110	-0.0171	-0.00642	-0.00642
	(0.17)	(-0.75)	(-0.26)	(-0.26)
_cons	0.00269	0.908***	0.0736***	0.0736***
	(0.59)	(53.28)	(3.98)	(3.98)
N	180	180	180	180
R^2	0.177	0.998	0.094	0.094

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5 Education and culture

It is well known by researcher that focus on culture economics the fact that there exists a relationship between culture consumption and education level, in part explained by the highest income of individuals with more years of education, but also explained by the idea of relationship between education and social class and education and the taste for culture. DiMaggio & Mukhtar (1982-2002), who analyze participation in high-culture arts, segregated their data by educational level. Cultural Statistic (2016), published by Eurostat, use educational level as variable to analyze the cultural participation as well. Beilby-Orrin & Gordon (2006) also introduce educational level as indicator for culture attendance. These literature leads us to expect that educational level has some influence in culture consumption. Thus, we want to redo our research introducing the educational level.

What we have done here is the same procedure es in method 1, but instead of applying it for entire heads sample, we have split it up into 2 subsamples: High education subsample for those who report having at least a University degree, and those who did not. Results differs between 2 subsamples. While for low education levels the result are close to results in section 4, with a drop in consumption level, which is significant at 10%, and R^2close to that obtained in section 4, for the high education subsample we have a huge drop in consumption but which is not significant.

Tab. 7: Estimation by Education subgroups

	(1)	(2)
	Coll.Att.	Not Coll.Att.
causal effect	-698.7	-102.2*
	(-1.33)	(-1.75)
distance1_65	26.88	-2.830
	(1.03)	(-0.99)
distance1_65_2	-0.964	-0.205
	(-0.70)	(-1.09)
_cons	720.3**	172.0***
	(2.11)	(3.95)
N	180	180
R^2	0.105	0.524

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

6 Conclusions

In this paper we have analyzed whether there are changes in culture consumption after retirement. We have focused on Spain, using micro data from 2006 to 2014. As culture consumption, we have included expenditure in performing arts, museum and exhibitions, and books. We have exploited a regression discontinuity design to measure the causal effect of retirement on culture consumption. We have isolated a subsample from original data, selecting those household where male heads are around the retirement age. We have also proposed 2 methods to estimate the causal effect.

Thus, our estimations show an annual drop of 166€-175€ in cultural goods consumption after retirement. Our procedure has also passed some robustness checks. These results are consistent with those we have found in the literature. We have done the same analysis splitting up our subsample by educational level. As results of this, we have found that those individuals with lower educational level have a significant drop in their culture consumption. However, those with higher educational level have not shown a significant change.

The reasons that explain this drop in culture consumption could be a good topic for a further research. It could be explained by the savings-puzzle, changes in taste, the requirement of changes in allocation of the expenditure or better consumption plan.

References

[1] Angrist, Joshua and Lavy, Victor. 1999. “Using Maimonides Rule to Estimate the Effect of Class Size on Scholastic Achievement”. The Quarterly Journal of Economics. 144(2): 533-75.

[2] Angrist, Joshua and Pischke, Jörn-Steffen. 2009. “Mostly Harmless Econometrics: An empiricist‘s Companion”, Princeton University press.

[3] Banks, James, Blundell, Richard and Tanner, Sarah. 1998. "Is there a Retirement-Savings Puzzle?". American Economic Review. 88(4): 769-88.

[4] Battistin, Erich, Brugiavini, Enrico Rettore and Weber, Guglielmo. 2009. “The retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach”. American Economic Review. 99(5): 2209-2226.

[5] Beblo, Miriam and Schreiber, Sven. 2010. “The Life-Cycle Hypothesis Revisited: Evidence on Housing Consumption after Retirement”. SOEPpapers on Multidisciplinary Panel Data Research. 339.

[6] Beilby-Orrin, Helen and Gordon John C. 2006. “International Measurement of the Economic and Social Importance of Culture”. OECD project on the measurement of Culture and Art-related activities.

[7] Bernheim, Douglas, Skinner, Jonathan and Weinberg, Steven. 2001. ”What Accounts for the Variation in Retirement Wealth Among U.S. Households?”. American Economic Review. 91(4): 832-57

[8] Browning, Martin and Lusardi, Annamaria. 1996. "Household Saving: Micro Theories and Micro Facts". Journal of Economic Literature. 34(4): 1797-1855.

[9] Dimaggio, Paul and Mukhtar, Toqir. 2004. “Arts participation as cultural capital in the United States, 1982–2002: Signs of decline?”. Poetics. 32: 169-94.

[10] Eurostat. 2016. “Culture Statistics”. Statistical Books.

[11] Hahn, Jinyong, Petra Todd and Wilbert van der Klaauw. 2001. “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design.” Econometrica, 69(1): 201–09.

[12] Lee, David S. 2008. “Randomized experiments from non-random selection in U.S. House elections”. Journal of Econometrics. 142: 675-97.

[13] Lee, David S. and Lemieux, Thomas. 2009. “Regression Discontinuity Designs in Economics”. National Bureau of Economic Research. Working Paper 14723.

[14] Luengo-Prado, María José and Sevilla, Almudena. 2013. “Time to Cook: Expenditure at Retirement in Spain”. The Economic Journal- 123:10.1111/ecoj.2013.123.issue-569, 764-789.

[15] Paulin, Geoffrey D. and Duly, Abby L. 2002. “Planning ahead: consumer expenditure patterns in retirement”. Monthly Labor Review. 125(7): 38-58.

[16] Rubin, Donald B. 1974. “Estimating Causal Effect of Treatments in Randomized and Non-randomized Studies”. Journal of Educational Psychology. 66(5): 688-701.

[17] Smith, Sarah. 2006. “The Retirement-Consumption Puzzle and Involuntary Early Retirement: Evidence from the British Household Panel Survey.” Economic Journal, 116(510): C130–48.

[18] Thistlethwaite, Donald L. and Donald T. Campbell. 1960. “Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment.” Journal of Educational Psychology, 51(6): 309–17.

[19] Van der Klaaw, Wilbert. 2002. “Estimating the Effect of Financial AID Offers on College Enrollement: A Regression-Discontinuity Approach”. International Economic Review. 43(4): 1249-1287.