Randomization Inferences about the Local Causal Effects of Retirement on Human Capital*

Eduardo Fé[†] [1] and Mario Pezzino[‡] [2]

[1] *Department of Economics, University of Strathclyde, U.K.*
[2] *Economics, University of Manchester, U.K.*

November 5, 2015

Abstract

We estimate the local causal effect of retirement on a range of measures of human capital with a new randomization inference framework for panel data regression and kink discontinuity designs. We find that retirement does not significantly affect any of the dimensions considered. We contextualise these results with a dynamic programming model of retirement. The model emphasises that effective retirement policies must be built on an understanding of the behavioural implications of retirement. More importantly, any causal effect of retirement is mediated through the contributions of an individual's time use and own job environment (as opposed to occupation) on human capital. However, these parameters can vary widely across individuals, so that heterogeneity in population-wide samples is likely to be substantial. Longitudinal information about time use and the extent to which individuals' jobs contribute to their human capital is rare, what severely restricts the ability to ascertain the extent to which existing estimates are sensitive to sample composition. As a result, data on time use together with new data on jobs' contributions to human capital are essential to obtain informative estimates of the causal effect of retirement.

Word count: 11,320

Key Words: *Human Capital, Panel Data, Randomization Inference, Regression discontinuity, Retirement.*

JEL Classification: *C21, C30, C90, J26, I18.*

---

*The authors thank Michael Caputo for his comments on the manuscript and Matias Cattaneo and Peter Ganong for helpful advice and discussion. The authors also benefited from the comments by participants at the 30th Annual Congress of the European Economic Association held at Universität Mannheim (August 2015).

[†] Department of Economics, University of Strathclyde. Sir William Duncan Building, 130 Rottenrow, Glasgow, G4 0GE, UK. email: eduardo.fe@strath.ac.uk.

[‡] Economics, University of Manchester. Oxford Road, Manchester, M13 9PL. email: mario.pezzino@manchester.ac.uk

1 Introduction.

The causal effect of retirement on human capital[1] is a topic of considerable policy relevance given that numerous countries around the world are modifying their retirement laws and, in particular, the qualifying criteria for state pensions and early retirement. The general goal of these policies is to create incentives that keep people at work until an older age and, ultimately, guarantee the sustainability of state pension schemes and improve individuals' economic prospects after retirement. However, the effectiveness of new legislation will depend on the implications of retirement for a population. For example, if retirement weakens mental hea lth or cognitive ability, individuals' increased vulnerability in the older age would be accompanied by a growth in the demand for health and social services. On the contrary, if retirement improves health or quality of life, there would be a strong incentive to retire from work as soon as possible, which would considerably dampen the effectiveness of those policies that intend to delay retirement.

The goals of this article are to empirically investigate the magnitude and sign of the causal effects of retirement on human capital and provide a theoretical framework within which to interpret the scope and limitations of results and those find in previous studies. An individual's level of human capital is a crucial determinant of retirement (Stern, 1989; Bound, 1991) and this complicates the estimation of the causal effects of retirement due to endogeneity. Therefore, to achieve the first goal, we propose panel data regression and kink regression discontinuity frameworks which enable the detection of sudden breaks and smooth changes in a conditional mean respectively. The environmental changes brought about by retirement can be dramatic, but might affect differently to alternative outcomes. For example, the sudden cessation of an occupation will affect the distribution of time across leisure and non-work activities. To measure this variation, a regression discontinuity design seems appropriate. On the contrary, the effect of retirement on cognitive ability or health might be progressive, in which case a kink regression discontinuity design might be better suited.

The new methods are panel data extensions of the Randomization Inference (RI; Fisher, 1935; Rosenbaum, 1996) frameworks for RDD and Kink Regression Designs (KRD) devised by Cattaneo, Frandsen, and Titiunik (2014) and Ganong and Jager (2014) respectively. In general, RI enables exact inferences on the parameters of interest with very small samples and

[1] Understood as knowledge, traits, health or skills which determine individuals' social and economic functioning

under very weak restrictions akin to the Independence-Exclusion-SUTVA trinity in Angrist et al. (1996). Because RI procedures are exact they successfully circumvent the size distortions typically expected under weak instrumental variables (Rosenbaum and Imbens (2005); Bound et al., 1995; Staiger and Stock, 1997). In a discontinuity setting these properties are highly attractive, given that data within the identifying window width tend to be scarce, so that asymptotic approximations are likely to be misleading. Equally important, the validity of the methods do not hinge on whether the running variable (age in our case) is measured discretely (as in our data) or continuously (thus, the pervasive discrete-running variable problem described in Lee and Card, 2008 is avoided). However, the assumptions underlying Cattaneo et al. (2014) and Ganong and Jager (2014) rule out inter-temporal correlations within individuals, which is restrictive for panel data environments. Thus we relax this hypothesis and enable a panel data approach that allows endogeneity due to time invariant unobserved heterogeneity. Unlike previous literature, we study a wide range of indicators of human capital (including, among others, health and cognitive functioning) while keeping national idiosyncrasies fixed by restricting the analysis to the English Longitudinal Study of Ageing (ELSA).

With these new methods, we find that retirement does not significantly affect any of the dimensions considered (including cognitive functioning). To better understand these results, we present a dynamic programming model which captures the two key features of retirement: the cessation of an occupation and the sudden increase in the availability of leisure time. This is the second contribution of the paper. The model outlines that the mechanisms through which retirement could affect individuals are the contribution of a particular job to human capital (as opposed to occupation *per se*) and an individual's re-allocation of leisure time after retirement. Our model emphasises that the effect of the same job on two different individuals can vary substantially depending on a large number of environmental factors (personal relations at work, attitudes towards work, individual preferences, goals and prospects, etc). But even if we could fix the contribution of a specific job to any one individual, the variation of human capital upon retirement is likely to be affected by a wide-ranging pattern of time use profiles which are likely to have a highly heterogeneous effect on human capital.

The empirical and policy implications of our model are profound. Because any causal effect of retirement is mediated through the contributions on human capital of an individual's time use and own job environment, longitudinal data on these dimensions are essential to tackle heterogeneity and obtain policy relevant estimates. Unfortunately, in practice data sets present very limited information on these dimensions. Whereas cross-sectional data on time use abound, existing longitudinal studies of adult populations only contain a few self-reported aggregates of time use outside work[2]. Longitudinal data measuring the contribution of an individual's job to her own human capital is also largely unavailable. Although job satisfaction and occupation are routinely measure in longitudinal studies, these are noisy measures of the relevant parameter. As a result, researchers' ability to ascertain the extent to which existing estimates are sensitive to sample composition is very restricted. This considerably limits the policy relevance of existing estimates and our understanding of whether specific findings are generalisable to a wider population or occupational group. However, a positive and compelling conclusion derived from the model, is that effective policy reforms should be built on an understanding of the behavioural implications of retirement, and those reforms that take these aspects into account will manage to mitigate the risk of policy-driven unintended consequences.

The remainder of the paper is organised as follows. Section 2 presents a discussion of the institutional setting of retirement in the UK. Section 3 introduces the econometric methods (with all the technical material deferred to an Appendix). Section 4 contains the empirical analysis. Section 5 discusses the results and studies a dynamic programming model that provides some intuition behind the empirical results. Section 6 concludes.

2 Data, Institutional Framework and Previous Studies.

Our data comes from the English Longitudinal Study of Ageing (ELSA), a bi-annual panel representative of the English population aged 50 or above. The first wave of the study was collected in 2002 and, at the time of writing, the sixth wave (collected in 2012) had just been released. However, we focus our attention on the period before October 2006. Until that date,

[2]For instance, the British Household Panel Survey (BHPS) only contains one variable regarding time use outside work, namely the number of weekly hours spent in housework (loosely understood as cooking, chores, etc.). Fé and Hollingsworth (2015) use data from an accompanying data set, the BHPS-Calibrated Time Use Data (Kan and Gershuny (2006)) to study variation of time use upon retirement. The BHPS-CTU is a longitudinal study containing calibrated time use data for each individual in the BHPS. However, calibration in this context implies calculating predicted values from a regression obtained from a the 1998-2001 Home on Line Survey (HoL), a longitudinal study which collected time use diaries from around 1000 individuals over three years. However, HoL is notorious for missing data and the sample size is too small to ensure that predicted values from this data set can indeed inform us about time use patterns in the participants of the BHPS.

individuals reaching age 60 (for women) or 65 (for men) could qualify for a state pension. On that date, the U.K. government transformed the State Pension Age (SPA) into a Default Retirement Age (DRA), what constituted a dramatic change to the institutional framework. In particular, the new law formally allowed firms to dismiss employees reaching the DRA (or 60 for women and 65 for men). Informally, the new legislation became an effective tool to discriminate workers facing retirement on the basis of productivity. The law was phased out in April 2011, but its existence makes retirement data from 2006-2011 incompatible with data prior to the introduction of the DRA. Using this data in our analysis would further invalidates our identification strategy, which requires comparability of the treatment and control groups in all respects except their eligibility for the SPA.

For the analysis, we exclude self-employed individuals or those permanently sick or disable. This results in a sub-sample of 15,915 observations from 7,449 individuals. At any point in time, an individual is categorised as retired if that is her self-reported job market status and she reports not to have undertaken any paid work during the two weeks prior to the interview. Our sample excludes all those individuals who re-enter the job market after retiring or who move to/from unemployment from/to retirement. After first-differencing the data and applying our definition of retirement, we are left with 6138 individuals providing 8,302 (first differenced) observations and 366 events of retirement.

The resulting data are used to study the effect of retirement on 10 summary indices (O'Brien, 1984; Anderson, 2008) of different aspects of human capital (cognitive functioning, quality of life, qualitative variation in expenditure, physical activity, engagement in socio-cultural activities, affective relationship with friends and mental fitness). Details about the construction of summary indices are given in the Appendix. The individual components of each index are listed in Table 5.

As in other popular data sets[3], in ELSA cognitive functioning is approximated by respondents' answers to a series of simple tasks aimed at revealing features of an individual's fluid and crystallised intelligence[4]. Most of these tasks involve memory recall exercises and solving numerical problems and thus are primarily suggestive about an individual's fluid intelligence

[3]See the Health and Retirement Survey in the U.S. or Understanding Society in the U.K. At European level, the The Survey of Health, Ageing and Retirement in Europe also has used similar measures.

[4]Fluid intelligence refers to the ability to solve problems in a logical way, regardless of acquired knowledge. Crystallised intelligence is characterised by a person's lifetime accumulation of knowledge and vocabulary. Similar items were used in the studies by Bonsang et al. (2012) and Mazzonna and Peracchi (2012).

-which from the point of view of retirement is, potentially, the most vulnerable dimension of cognition[5]. For our analysis we keep those items which appear repeated in consecutive waves and construct two separate summary indices. The first index measures the overall performance in all the tests, while the second index focuses on the performance in memory tasks alone.

Quality of Life is measured with CASP-19, a 19 item survey which intends to measure an individual's level of Control, Autonomy, Self-realisation and Pleasure in life. Each item appears in the survey as a four-point Likert response ranging from "Often" to "Never". Items were recoded so that higher scores equal lower quality of life. In addition to this, individuals were asked to place themselves in a 10 step ladder, in accordance to their satisfaction with life as a whole (higher values revealing greater satisfaction). We construct separate indices with CASP-19 and an individual's self-reported position in the ladder.

Our fifth summary index measures qualitative variations in expenditure. It is constructed from binary indicators of whether the respondent smokes, eats out at least once every few months, cut the size of a meal in the last year or feels that shortage of money often prevents him from doing what he wants. Physical activity is measured with two independent indices of whether the individual engages, at least 1-3 times a month, in vigorous or moderate intensity sports/activities. Socio-cultural activity measures attendance to public venues (cinemas, theatres and museums) and social gatherings (meetings with friends). We further construct a summary index of an individual's affective relationship with his friends from six four-point ordinal variables measuring the frequency with which a person feels in a particular way in relation to her friends. The scales are coded so that higher scores denote worse outcomes. Finally, a summary index for mental health is constructed from 8 dichotomous variables measuring whether, during the previous week, an individual felt a range of negative emotions such as isolation, restlessness or depressed.

2.1 Institutional Framework.

For most of the recent British history, an individual's income after retirement has been determined by the history of contributions to private and public pension schemes[6]. Although private

[5]Note that although this type of tasks has become common place in large surveys and are simple and convenient to administer, to the best of our knowledge, their ability to reveal anything about cogntive functioning has not been tested. In particular, we have not found any study correlating these measures to more formal tests, such as such as the Wechsler Adult Intelligence Scale or the Raven Progressive Matrix Test.

[6]Part of this section draws from **?**.

pensions are common in the U.K. -for example, in 2012, about 41%-55% of 30-59 year old adults were contributing to a private pension scheme[7]- public pensions still play a crucial role in the system. An individual with a full history of 30 years of National Insurance contributions received, in 2013, a maximum Basic State Pension (BSP) of £110 ($172), that is approximately a quarter of the UK weekly minimum wage. Furthermore, there is a State Second Pension (S2P) which provides additional income based on an individual's earnings over their entire working life. Eligibility for BSP and S2P requires enough qualifying years (currently 30) and, importantly for our identification strategy, reaching a pension age (the state pension age, SPA) which, for the period under consideration in this study was 65 years for men and 60 for women.

The eligibility criteria present a strong incentive to stay at work until the SPA and, as a result, the proportion of retired individuals by age group in any give year exhibits a unique jump of around 20 percentage points at the SPA. This gap can be seen in Figure 1 which describe the distribution of retirement among individuals in ELSA. As a result, the SPA arises as a predictor of retirement decisions, (although its predictive power is limited by the fact that, as revealed by Figure 1, by SPA around 60% of individuals have already retired).

Given SPA's ability to predict retirement decisions and the fact that it is exogenously determined by Government, one can argue that a dummy variable indicating if a person has reached the SPA is a valid instrumental variable for retirement. In particular, if the discontinuity in the distribution of retirees is accompanied by a significant discontinuity in the conditional mean of measures of human capital outcomes (a jump) or its first derivatives (a kink), then, under certain assumptions, this variation in outcomes could be attributed to retirement. Yet, although SPA is exogenously determined by government, it splits the population into a young and old groups, both of which exhibit dramatically different human capital profiles. This induces a correlation between the instrument and human capital. Thus, additional identifying assumptions are required. These assumptions are the subject of the next section.

2.2 Previous studies.

A number of studies have explore the effect of retirement on some dimensions of human capital (health and cognitive ability in particular) by putting forward compelling instrumental variable

[7]Pensions Policy Institute https://www.pensionspolicyinstitute.org.uk/default.asp?p=81.

strategies, typically based on peculiarities of retirement laws determining the eligibility criteria for (state) pensions. Most of these studies have relied on linear regression models designed to estimate permanent causal effects through discontinuities in conditional means. In these studies, the interpretation of any estimated effect as permanent is imposed by the structure of the parametric model itself, which explicitly specifies the path of human capital in the absence of retirement[8].

Parametric linear regression can be restrictive and the underlying assumptions might decisively explain the magnitude and significance of estimated causal effects (e.g. White, 1982). Furthermore, this parametric approach implicitly assumes that, just as a communicable disease, retirement has an immediate manifestation. Although the changes induced by retirement are indeed dramatic, they mostly affect environmental factors -such as leisure time. Environmental changes might subsequently affect human capital, thus establishing an indirect line of causation between retirement and human capital, but in this case the path of human capital is more likely to reveal smooth changes in trends around retirement than breaks.

A few number of empirical studies have tried to go beyond the linear regression models to estimate the causal effect of retirement on human capital. The earliest attempts in this direction are Coe and Zamarro (2008), Johnston and Lee (2009) and Fé and Hollingsworth (2012, 2015) who, building on previous work by Battistin et al. (2009), consider Regression Discontinuity Designs (RDD) to estimate the effects of retirement on health. RDD relax parametric restrictions by focusing on variations in outcomes within a narrow interval around retirement[9] (Hahn et al., 2001; Porter, 2003; Card et al., 2012). However the results in Coe and Zamarro (2008) and Johnston and Lee (2009) are vulnerable to the so called discrete-running variable problem[10] (which raises concerns about the statistical significance of their results -Lee and Card,

[8]Published results from parametric studies are often contradictory. Although the only two existing studies on cognition (Bonsang et al., 2012 and Mazzonna and Peracchi, 2012) find that retirement has a significant negative effect on cognitive functioning, studies on the effect of retirement on health are largely inconclusive. For example, seminal work by Charles (2004) concludes that the "... *direct effect of retirement on well-being is positive once the fact that retirement and well being are simultaneously determined is accounted for...*". Dhaval et al. (2008) finding that retirement leads to a 6-9% decrease in mental health, and a 5-6% increase in illness conditions. Coe and Lindeboom (2008) conclude that there are no negative health effects of retirement. Neuman (2008) finds that there is strong evidence dismissing the idea that retirement harms health

[9]Furthermore, unlike parametric approaches, discontinuity designs put explicit emphasis on the comparability of control and treatment groups, forcing researchers to undertake a thorough exploratory analysis of pre-determined characteristics of the individuals in the sample -which should be homogeneous in a neighbourhood of the policy threshold. This helps to detect important features of the sample that would go unnoticed in a parametric study. For example, we shall see that a set of the individuals in our sample were affected by a change in the school leaving age in England (the Butler Act of 1944) which systematically changed the prospects of affected individuals (Oreopoulos (2006)). Standard parametric models neglecting this feature of the data would draw conclusions on the basis of non-comparable groups.

[10]Lee and Card (2008) note that if the variable determining eligibility for treatment is discrete valued, the assumptions

2008) and while Fé and Hollingsworth (2015) present a novel RDD robust to discrete running variables as well as weak instruments (and which takes into account small sample distortion in inferences as well as individual heterogeneity), this is achieved at the expense of methodological complexity and ancillary parametric restrictions which reduce the generality of their results. On the conceptual side, only two studies contemplate the possibility of a smooth change in trends upon retirement. The seminal contribution is the paper by Bound and Waidmann (2007) who estimate the local effects of retirement on health using a parametric regression model designed to capture discontinuities in the first derivatives of a conditional means. Fé and Hollingsworth (2015) further extend the model to panel data. Neither study finds any significant effect of retirement on health however, as with most previous literature, these results rely on parametric restrictions which decrease the generality of the results.

3 Estimation and inference.

In this section we present an non-technical outline of our identification strategy leaving, for the purpose of clarity, the technical details to the Appendix. We will use the following notation. The key policy variable in retirement legislation is age, which we denote by R_{it} for individual $i = 1, \dots, N$ at time $t = 1, \dots, T$. R_{it} is a random variable, and could be continuous or -crucially for later developments- discrete. If an individual reaches the SPA, then $Z_{it} = \mathbb{I}(R_{it} \geq r_o) = 1$, where $\mathbb{I}$ is an indicator taking value 1 if the statement in brackets is true (0 otherwise). The policy cut-off r_o is 60 for women and 65 for men. We consider situations where compliance with assignment is imperfect, so that actual treatment, D_{it}, may not coincide with Z_{it} (that is, a fuzzy discontinuity design).

Our identification strategy is a randomization inference (RI) framework for discontinuity designs in conditional means and their first order partial derivatives which exploit the panel nature of the data. Unlike in the standard program evaluation framework, in RI potential outcomes are fixed characteristics of an individual. Variation in the data comes from the policy variable determining allocation into treatment, whose values are allocated at random accordance to a distribution (compare Rubin, 1974; Angrist, Imbens, and Rubin, 1996 vs

underlying RDD are violated and nonparametric identification is infeasible. Further parametric assumptions can help to recover identification but then standard errors must be corrected in ways determined by the specific assumptions made about the nature of the measurement error induced by the parametric approximation -see Lee and Card (2008) and Dong (2014).

Fisher, 1935; Rosenbaum, 1996). Formally, let $\mathbf{R}$ and $\mathbf{Z}$ be the $NT \times 1$ vectors of scores of the running variable and assignment indicators. The potential treatment status of individual i at time t when $\mathbf{R} = \mathbf{r}$ is denoted by $d_{it}(\mathbf{r}) \equiv d_{it}(\mathbf{z})$. This notation emphasises that potential outcomes are fixed, non-random objects. Prior to the determination of $\mathbf{R}$, however, potential outcomes are random, in which case $\mathbf{D}$ denotes the $NT \times 1$ vector of random treatment status for the whole sample. Then, $y_{it}(\mathbf{r}, \mathbf{d}) \equiv y_{it}(\mathbf{z}, \mathbf{d})$ is the potential outcome of individual i at time t when $\mathbf{D} = \mathbf{d}$ and $\mathbf{R} = \mathbf{r}$. Note that, in principle, potential outcomes depend on the whole history of R (and D).

Unlike in asymptotic settings, the distribution underlying R can generally be observed, inferred or simulated from the data, implying that researchers can then construct *exact* tests of hypothesis. These tests can be subsequently inverted to obtain point estimates and confidence intervals (under stronger regularity conditions; see Hodges and Lehmann, 1963; Rosenbaum, 1996). The attractiveness of exact tests in a discontinuity context cannot be overstated. Given the focus of these method on a short interval around the policy threshold, only a small proportion of the observations available are effectively used, what tends to result in small samples sizes under which asymptotic results are likely to provide rough approximations. Equally important, is the fact that, as noted by Rosenbaum and Imbens (2005), exact inference confers a remarkable robustness of RI methods to the quality of the instrumental variable set, with empirical sizes marginally varying around nominal level even with weak instruments. This is in contrast with asymptotic methods where the quality of inferences greatly depends on the strength of the instruments (Staiger and Stock, 1997; Kleibergen, 2002; Davidson and MacKinnon, 2006).

3.1 Randomization Inference in RDD with Panel Data.

Cattaneo, Frandsen, and Titiunik (2014) devise the randomization inference framework for cross-sectional RDD, by noting that randomization inference can be applied in RDD if it is possible to identify a neighbourhood, W_0, of the cut-off point where the familiar Independence-SUTVA[11]-Exclusion trinity holds (Angrist, Imbens, and Rubin, 1996). Crucially, identification in this context does not hinge on whether the running variable in the analysis is continuous or discrete (see, in contrast, Lee and Card, 2008 or Dong, 2014).

[11]Stable Unit Treatment Value Assumption.

When panel data is available, the assumptions in Cattaneo et al. (2014) are rather restrictive. In particular, Assumptions 1.1.b and 1.2. in the Appendix (corresponding to the local randomization and SUTVA assumptions in their article) rule out correlations within individuals -a key characteristic of panels. The problem can be solved if, as discussed in the Appendix, we assume that cross-sectional correlations are due to additive, time invariant unobserved heterogeneity in potential outcomes (in W_0). First-differencing the data would eliminate correlations due to unobserved heterogeneity and, then we only need to assume that the differenced data satisfies Assumptions A1', A2' and A4 in Cattaneo et al. (2014) (summarized in Assumption 1 in the Appendix; in particular we ensure that variation in potential outcomes outside and inside W_0 are uncorrelated, which is a milder assumption once heterogeneity has been differenced out)[12]. Once these assumptions are in place, the RI framework in Cattaneo et al. (2014) can be applied to the first-differenced data with minor variations -as described in our Appendix.

3.2 Randomization Inference in RKD with Panel Data.

We are further interested in the estimation of kinks in the conditional mean of human capital. Nonparametric and semiparametric estimation of kinks with a known policy threshold has been discussed by Simonsen et al. (2010) and Card et al. (2012) who introduced the Regression Kink Design which, in a fashion similar to the RDD, identifies discontinuities in the first order partial derivative of conditional means -under the additional assumptions of continuous differentiable potential outcomes. The method is, however highly sensitive to curvature in the underlying relationship between the outcome and the assignment variable. Thus, in a recent paper, Ganong and Jager (2014) present a RI framework for RKD which enjoys equivalent properties to the methods in Cattaneo et al. (2014) (exact tests, robustness to weak instruments and discrete running variables) and produces accurate inference base on tests with little size distortion. Unlike in Cattaneo et al. (2014) the randomization mechanism in Ganong and Jager (2014) relies on the assumption that the location of the policy kink can be considered as randomly drawn from a known interval and, in this interval, one can reassign the location of the kink and calculate RK estimates at these placebo kinks. In our setting, however, the discrete running variable considerably limits the number of points at which the placebo kinks can be computed

[12]The assumptions required for identification, as well as the inferential and estimation procedures are formally developed in our Appendix, where we also present a Monte Carlo simulation evaluating the merits of our procedure.

and, therefore, our ability to construct an accurate p-value.

In this article we combine early work by Bound and Waidmann (2007) and Fé and Hollingsworth (2015) with the traditional RI approach to produce a panel data variation of the methods in Ganong and Jager (2014). Unlike in their paper, the randomization mechanism relabels the treatment status of observations, while holding fixed their observed outcomes. As in Bound and Waidmann (2007) we investigate if there are changes in the slope or level of outcomes at the SPA. Because the SPA is exogenously determined by government -and tradition- and predicts, to a great extent, retirement behaviour, any change in trend at SPA can be attributed to those individuals whose labour force status changes due to reaching the SPA. This approach thus differs from the RDD described in the previous section, where variation in retirement status was instrumented by variation in state pension eligibility. In the terminology of discontinuity designs, we now opt for a Sharp design -as opposed to the Fuzzy design introduced in the previous section. An additional reason to pursue a sharp design is that existing Kink Regression literature (in particular, results in Card et al., 2012 and Dong, 2013) seem to suggest that a kink in the propensity score is necessary to identify kinks in the structural equation. However, neither Figure 1 nor a fixed effects regression in first differences suggested the existence of kinks at the policy cut-off in the distribution of retirees by age group.

As discussed in the Appendix, identification of kinks is possible under the local randomization and local SUTVA introduced for the RDD, together with additive time invariant unobserved heterogeneity. Unlike in the RDD, however, we must impose homogeneous treatment effects, which is relatively restrictive.

4 Empirical Analysis.

Tables 1 and 2 present some descriptive statistics of the sample. Table 1 focuses on predetermined variables and Table 2 focuses on outcomes. In each of these tables, columns 2 and 3 compare the profiles of those individuals who are observed to retire in the sample with the profiles of those other individuals whose job status remains unchanged (because they remain employed or retired during the years of the sample). In total 366 individuals are observed to move from employment into retirement. These individuals are younger on average (62.45 vs 65.53 years of age), but otherwise exhibit a similar demographic profile than the remaining individuals in the sample.

It is worth noting, however, the difference in the proportion of individuals who finished school at age 14 -considerably smaller in the group of individuals who retire in the sample. The reason for this variation has to do with the introduction of the 1944 Education Act[13] (the Butler Act) which increased compulsory schooling leaving age to 15 in 1947. This Act affects to those individuals in the sample who were 69 or younger in 2002 (the first wave of ELSA) and its effect is patent in panel c in Figure 3 which plots the distribution of individuals by age group who finished school at age 14 or earlier and exhibits a dramatic change in the slope at x= 6. The Butler Act has been showed to have a considerable effect on the socio-economic prospects of the individuals affected (Oreopoulos, 2006) and thus by not taking this effect into account, causal inferences would be based on considerably different populations. A discontinuity setting, by forcing researchers to empirically investigate the homogeneity of predetermined characteristics, mitigates the risk of gross confounders, such as the Butler law. In regression analyses, on the contrary, considerations such as this tend to go unnoticed.

Columns 4 and 5 in the tables compare the mean outcomes of observations with job status equal to retired against the rest of the sample. This is a comparison of the older echelon in the sample to the younger one, with allocation to either cohort in terms of retirement status. As a result, the group of retirees is older, has more grand-children, and the proportion of individuals who left school at age 14 is higher (for the reasons mentioned above).

In terms of outcomes Table 2 reveals that, as expected, average scores are typically 0, whilst the standard errors vary slightly around 0.5. There are some differences between the average scores of those individuals who are observed to retire and those who are not. The latter group exhibits worse mental health, cognitive functioning, and quality of life (but we know this group is older in age and includes the oldest individuals in the sample). When comparing the outcomes of retired and non-retired individuals in columns 4 and 5 of table 2, the differences are fairly small. In line with these findings, Figure 2, which plots local linear regressions of the first differenced outcomes ($Y_{it} - Y_{it-1}$) by age group, does not suggest any dramatic changes in trends around the SPA. Rather, first differenced outcomes tend to exhibit great variability across all age ranges.

Overall, the descriptive analysis does not suggest major differences in the pre-determined

[13]See Oreopoulos (2006).

characteristics or outcomes of retirees and non-retirees in our sample (beyond the difference in schooling introduced by the Butler Act of 1944).

Table 3 reports the RDD and RKD estimates of the causal effect of retirement on each of the summary indices considered. In both procedures, the computation of p-values relied on 999 bootstraps, and the significance threshold were set at 5%. In the short term procedure the bandwidth was selected on the basis of the procedure in Cattaneo et al. (2014), using the variables in 4. The selected bandwidth was $[-5, 5]$[14].

The results in the table confirm the conclusions suggested by the descriptive analysis. Namely, there are no effects of retirement on any of the domains considered. These conclusions were insensitive to small variations in the bandwidth. In particular note that, the coefficient of self-reported quality of life is negative and significant in the RDD analysis (implying an improvement in perceived quality of life), but the significance of this effect disappears in the RKD. This is indicative that, although retirement might affect individuals' quality of life in the short run, the effect is irrelevant for the longer run.

5 Discussion: An elusive causal effect.

Despite the quality of data and the strength of the identification strategy, our results do not reveal any causal effect of retirement on the dimensions of human capital considered. Should we then conclude that retirement is innocuous? For example, a number of published studies have reported variation in treatment effects across different occupational groups. To gain insight into this question and better understand the scope of our results (and those in previous studies), we introduce a parsimonious, albeit highly illustrative, dynamic model of retirement. At the core of the model sit the defining features of retirement, namely the cessation of an occupation and the sudden increase in the availability of leisure time. As a result, the mechanisms through which retirement could affect human capital are the contribution of a particular job to human capital (as opposed to occupation *per se*) and an individual's optimal allocation of time between leisure and other household activities (before and after retirement).

Thus, consider a representative agent who lives for $T \geq 2$ periods. In each period, the agent

[14] Prior to implementing the method, we undertook a graphic exploration of any potential discontinuities in the distribution of pre-determined individual's characteristics (what would suggest that events other than retirement might affect variations in human capital -thus confounding our estimates). As can be seen in Figure 3, we do not observe major breaks in trend in predetermined characteristics around the SPA.

has to decide the amount of time spent in leisure, T^a, home production, T^h and at work, T^w, so that the total amount of time available to the individual satisfies $\bar{T} = T^w + T^a + T^h$. Period $t = 1$ represents the working life of the individual. At the beginning of $t = 1$ the agent is endowed with assets A_1 and human capital H_1 (we will use H to denote a generic measure of stock of human capital which could be an aggregate measure of the overall stock or an indicator of a single dimension -such as health or cognitive ability). The individual can then access the job market and decide how she divides her time between work, T_1^w, leisure, T_1^a and household activities, T_1^h. At the end of period $t = 1$ she receives a payment for her work. The magnitude of the wage depends on the individual's level of human capital at the beginning of the period, $w(H_1)$. At the beginning of period $t = 2$ the agent retires, so that $T_t^w = 0$ for $t \geq 2$.

The dynamic equations characterising the accumulation of assets and human capital over life are given by,

$$A_t = \rho A_{t-1} + w(H_{t-1})T_{t-1}^w \tag{5.1}$$

$$H_t = \delta H_{t-1} + \theta T_{t-1}^a + \gamma T_{t-1}^w \tag{5.2}$$

The financial rate of return of assets is denoted by $\rho > 1$, while $\delta \in (0, 1)$ represents the rate of obsolescence of human capital. The first important feature of the model is the parameter $\theta \in \mathbb{R}$, which captures the effects that leisure might have on human capital. The sign of this parameter could be positive (e.g. studying, exercising, attending cultural events, socializing, etc.) or negative (e.g. excessive eating and drinking, watching TV, etc.) and which effect prevails may depend on a myriad of factors such as income, age, and social and cultural background. The second crucial aspect of the model is $\gamma \in \mathbb{R}$, which measures the marginal contribution of time at work to overall human capital. As with θ, $\gamma \in \mathbb{R}$ implies that labour supply can affect the stock of human capital, but the overall effect is ambiguous. For example, some occupations may be physically burdensome, speeding up normal deterioration of health. Other occupations may promote the accumulation of human capital through, for instance, continuous intellectual development[15].

The individual derives utility from her stock of human capital, leisure and, possibly from the

[15] Note, that the model assumed that T^h does not affect human capital. This assumption could be relaxed, but doing so would leave unchanged the main argument of the section.

time spent in home production. She maximises her lifetime utility, weighted by the subjective discount rate $\beta \in (0,1)$, for which she solves the following optimization problem,

$$\max_{T_t^h, T_t^a} \left[\sum_{t=1}^{T} \beta^t U\left(T_t^h, T_t^a; H_t\right) + \beta^T A_T \right] \tag{5.3}$$

subject to (5.1) and (5.2), where A_T is the stock of asset remaining in the last period of life.

As we illustrate in the Appendix with a two period version of this model, the solution of this probleml strongly hinges on the particular features of the individual's utility function and, in particular, the sign of the cross-partial second derivatives with respect to the time dimensions. However, the sign of these derivatives is in general ambiguous. As a result, it is not feasible to conclude much in terms of the causal effects of retirement on human capital without reliable information regarding individuals' preferences. In addition the contributions of θ and γ are essential to understand the effect of retirement on H. In order to provide further insights into the role of retirement for human capital, let us further characterise individual's preferences and wage equations with the following simple and commonly used linear-quadratic functional forms[16],

$$\begin{aligned} U\left(T_t^h, T_t^a; H_t\right) &= H_t + \alpha_h T_t^h - \frac{\left(T_t^h\right)^2}{2} + \alpha_a T_t^a - \frac{\left(T_t^a\right)^2}{2} \\ w\left(H_1\right) &= H_1 \end{aligned}$$

where $\alpha_i > 0$, $i = h, a$, represent the marginal benefit that the agent can obtain from investing time in household (α_h) and leisure activities (α_a). Both activities come at a (quadratic) cost, ensuring that function U is a well-behaved strictly concave function in (T_t^h, T_t^a). Proposition 1 describes the solution of dynamic programming problem (5.3).

Proposition 1 *The (interior) solution of problem (5.3) is:*

[16]The fact that $U\left(T_t^h, T_t^a; H_t\right)$ is additive separable in T_t^h, T_t^a and H_t is clearly a simplification. Nonetheless the assumption allows us to obtain tractable analytical solutions to the dynamic programming problem that do not depend (at least in the case of retirement) on the state variable H_t. The additive separability of the utility function therefore allows us to clearly identify the way the state variable H_t affects the solution of the problem during employment (via the utility and the salary) and during retirement (via utility only).

- $t = 1$

$$T_1^h = \alpha_h - \beta\gamma \left(1 + \sum_{i=1}^{T-1} \beta^i \delta^i\right) - \beta^{T-1}\rho^{T-2}\gamma H_1$$
$$T_1^a = \alpha_l + \beta(\theta - \gamma)\left(1 + \sum_{i=1}^{T-2} \beta^i \delta^i\right) - \beta^{T-1}\rho^{T-2}\gamma H_1$$

- $t = 2, ..., T - 1$

$$T_{T-k}^h = \frac{1}{2}\left(T + a_h - a_a - \theta \sum_{i}^{k} \beta^i \delta^{i-1}\right)$$
$$T_{T-k}^a = T - T_{T-k}^h$$
$$k = 1, ..., T - 2$$

- $t = T$

$$T_T^h = \frac{1}{2}(T + a_h - a_a) = T_T^a$$

At $t = 1$ the agent chooses T_1^h and T_1^a optimally by considering both their direct contributions to utility at that time (α_h and α_a) and the dynamic effect that these controls have on the system via H_t (parameters γ, θ and δ) and A_t (parameter ρ), discounted by β. If γ is positive (i.e. working contributes to the improvement of human capital), then the agent prefers to work more and reduce other activities at time $t = 1$. This effect may be mitigated or exacerbated depending on the choice of leisure activities, since T_1^a may also contribute to the enhancement or reduction of human capital (depending on the sign and level of θ). Not surprisingly, if $\gamma > 0$ then an agent with a high initial endowment of human capital (H_1) may decide to work more in order to exploit the future advantages in terms of H_t and A_t. These results will be reversed if $\gamma < 0$.

After retirement ($t = 2, ..., T$) if $\theta = 0$ (that is, leisure activities do not affect the accumulation of human capital) then the agent will choose a constant allocation of time $T_t^h = T_t^a = \frac{1}{2}\left(\bar{T} + a_h - a_a\right)$ that depends on the direct benefit that each activity brings in each period. If $\theta > 0$ ($\theta < 0$), the allocation of time tends to be skewed in favor of T_t^a (T_t^h). In fact T_t^a contributes both towards the agent's utility and the development of human capital, while T_t^h only has a direct effect on utility. Nonetheless, the dynamic effect of T_t^a tends to be weaker over time and disappears at $t = T - 1$.

To identify the effect that retirement might have on the time allocation, we can compare the optimal values of T_t^h and T_t^a at times $t = 1$ (i.e. employment) and 2 (retirement). Once again, the net effect of retirement on time use depends on the specific values of γ and θ. However, we can observe the following particular cases:

1. $\gamma > 0$, i.e. the time spent at work contributes to the intellectual development of the individual. In this case retirement is likely to produces a significant increase in household and leisure activities, so that $T_1^h \leq T_2^h$ and $T_1^a \leq T_2^a$. In period $t = 1$ it was advantageous for the agent to work and reduce other activities. After retirement the agent can only choose between household and leisure time, with a choice skewed in favour of the former (if $\theta < 0$) or the latter (if $\theta > 0$).

2. $\gamma < 0$, the time spent at work has a negative effect on the human capital of the individual. This would be the case of jobs particularly taxing on the body and mind of the worker. In this case retirement is likely to produce minor changes in household and promoting activities. Work has negative effect on the accumulation of human capital, with negative dynamic effects for the agent. The individual therefore decides to limit the labour supply in favour of T_1^h and T_1^a (in particular if $\theta > 0$). After retirement, where labour supply is not an option, we should observe a small increase in the levels of T^h and T^a with a preference for leisure activities if $\theta > 0$.

We are ultimately interested in understanding the way that γ, θ and agents' decisions affect human capital before and after retirement. From the information in Proposition 1, (5.1) and (5.2), we know that the equilibrium levels of human capital at the beginning of time $t = 2$ (pre-retirement) and time $t = 3$ (post-retirement) are:

$$\begin{aligned} H_2 &= \frac{1}{\beta\rho^2}\left[\beta^T\rho^T H_1 (2\gamma - \theta)\right] + \delta H_1 - \gamma (a_a + a_h - T) \\ &+ \theta a_a + \beta\theta\left(1 + \sum_{i=1}^{T-2}\beta^i\delta^i\right)(2\gamma - \theta) \qquad (5.4) \\ H_3 &= \frac{1}{2}\theta\left(T - a_h + a_a + \theta\sum_{i}^{k}\beta^i\delta^{i-1}\right) + \delta H_2 \qquad (5.5) \end{aligned}$$

The equilibrium expression of H_3 reveals why retirement does not have to necessarily produce a negative effect on human capital. Indeed suppose that δ were sufficiently high (e.g. close to 1). Then, if θ is positive and sufficiently large, human capital would *increase* after retirement. Of course, if θ is negative, leisure activities would produce a negative effect on human capital after retirement and $H_3 < H_2$.

While the sign of $(H_3 - H_2)$ depends on the particular values that the parameters can take, it is also interesting to analyse the effect that γ and θ may have on the equilibrium levels of human capital before and after retirement,

$$\frac{dH_2}{d\theta} = a_a - \beta^{T-1}\rho^{T-2}H_1 + 2\beta(\gamma - \theta)\left(1 + \sum_{i=1}^{T-2}\beta^i\delta^i\right) \tag{5.6}$$

$$\frac{dH_2}{d\gamma} = -a_h - a_a - 2\beta^{T-1}\rho^{T-2}H_1 + T + 2\beta\theta\left(1 + \sum_{i=1}^{T-2}\beta^i\delta^i\right) \tag{5.7}$$

$$\begin{aligned}\frac{dH_3}{d\theta} &= \frac{1}{2}\Bigg[T - a_h + a_a + 2\theta\sum_i^k\beta^i\delta^{i-1} \\ &+ 2\delta\left(\frac{-\beta^T\rho^T H_1}{\beta\rho^2} + a_a + 2\beta(\gamma-\theta)\left(1+\sum_{i=1}^{T-2}\beta^i\delta^i\right)\right)\Bigg]\end{aligned} \tag{5.8}$$

$$\frac{dH_3}{d\gamma} = \delta\frac{dH_2}{d\gamma} \tag{5.9}$$

These equations reveal that human capital can increase (before and after retirement) with θ when a_a is large, the difference $(\gamma - \theta)$ is positive and large and H_1 is small. Similarly, human capital increases with γ when both a_h and a_a are small, θ is positive and large and H_1 is sufficiently large.

The key empirical implication of the model is that the overall effect of retirement on human capital depends on the interplay between γ and θ. These parameters, however, are likely to vary widely across individuals, so that heterogeneity in population-wide samples is likely to be substantial. Individuals in identical occupations can experience significantly different γ depending on their life expectations, colleagues, management etc. But even if this parameter could be fixed, what people do in their spare time and how this impacts in their human capital (θ) is likely to be wide ranging. Therefore, in the absence of sufficient information regarding (θ, γ), empirical estimates of the causal effect of retirement are likely to be heavily dependent on the $(\theta - \gamma)$ composition of a specific sample.

The question is then to what extent existing data can assists in taming this heterogeneity. Time use data is typically scarce or imprecisely measured in longitudinal studies (for example, the British Household Panel Survey -BHPS- a large study running since 1991, only includes subjects' estimated weekly hours spent in housework, widely defined). Longitudinal studies containing time use diaries are scarce for adult populations[17]. Occupational and labour supply data is routinely collected, however this data is uninformative about the ways a specific post contributes to one's human capital (e.g. intellectual demands, training, social networks, human relations) and is often missing for individuals who retire or are retired at the point when collecting the sample. Thus, a naive separation of samples on the basis of white/blue collar, manual/non-manual worker or formal international classification systems are unlikely to solve the $\theta - \gamma$ problem. Finally, it seems that the sample sizes required to estimate the causal effect of retirement on human capital are also likely to be very large. For instance, in our study power comes from 366 events of retirement and above 8000 observations. Creating sub-samples on the basis of blue/white collar occupations or other international classification system is unlikely to circumvent the $\theta - \gamma$ problem but will lead to a quick loss of statistical power. It appears clear that more data, possibly based on time diaries recording individuals' activities before and after retirement, is essential to allow meaningful estimation.

From a policy perspective, the key insight from the model is that, as far as human capital is concerned, what really matters is to educate individuals to ensure θ is positive and large rather than presuming a detrimental effect of retirement on people's lives and worrying about modifying the retirement age (of course, the overall debate around extending the retirement age also has an important public finance dimension -productivity of older workers and long term sustainability of pension schemes- which we are not considering here). In addition, the contribution of work to human capital, γ, has repercussions for human capital after retirement (H_2). If this contribution is sufficiently significant as to ensure a large enough H_2, then some individuals with low (but positive) θ might still experience an increase in human capital after retirement. This remark calls for sector-based studies and reforms.

[17]Neither ELSA, BHPS, the Study of Health and Retirement in Europe (SHARE), the Health and Retirement Survey, Panel Study of Income Dynamics or the German Socieconomic Panel contain time use diaries. Of course, time use diaries are not a panacea. A promising approach between the imprecise questions routinely contained in the mentioned studies and time use diaries are the instruments developed in Browning and Gørtz (2012) which provide informative data without incurring in the burden of sampling detail logs of activities.

6 Conclusion.

The goals of this article were to empirically investigate the magnitude and sign of the causal effects of retirement on human capital and provide a theoretical framework within which to interpret the scope and limitations of results and those find in previous studies. We provide estimates of the causal effect of retirement on a series of indicators of human capital. Identification relied on a new Randomization Inference framework for discontinuity designs which is applicable when panel data are available. The framework relaxes the assumptions in Cattaneo et al. (2014) to allow additive time invariant heterogeneity. It further enables a panel-data variation of the RI framework for kink discontinuity designs in Ganong and Jager (2014). In this new framework identification relies on comparisons of outcomes within individuals (which reduces the contentiousness of the underlying assumption of continuous/continuous-differentiable potential outcomes). Crucially, we can construct test of hypothesis with exact size (even with very small samples) and the interpretation of the estimated parameter is independent on whether the assignment variable is continuous or discrete. The attractiveness of exact tests in a discontinuity context cannot be overstated. Give the focus of these method on an short interval around the policy threshold, only a small proportion of the observations available are effectively used, what tends to result in small samples sizes under which asymptotic methods are likely to provide rough approximations and misleading inferences. Equally important is the fact that, as noted by Rosenbaum and Imbens (2005), exact inference confers a remarkable robustness of RI methods to the quality of the instrumental variable set, with empirical sizes marginally varying around nominal level even with weak instruments. This is in contrast with asymptotic methods where the quality of inferences greatly depends on the strength of the instruments (Staiger and Stock, 1997; Kleibergen, 2002; Davidson and MacKinnon, 2006).

We apply the new estimation framework to the English Longitudinal Study of Ageing (ELSA) and find that retirement does not induce any jumps or kinks on self-reported quality of life, qualitative expenditure, social, physical or cultural activity or the relationships with friends and parters. Our results are aligned with those in Coe and Lindeboom (2008), Neuman (2008), Bound and Waidmann (2007) and Fé and Hollingsworth (2015). That self-reported qualitative expenditure does not vary upon retirement was also found by Aguiar and Hurst (2005) (who also find that consumption does change with retirement). Unlike in Mazzonna and

Peracchi (2012) and Bonsang et al. (2012) we do not find that retirement affects cognitive functioning. This particular finding is also related to Kovalchik et al. (2004) who study responses to a series of laboratory experiments in two populations, one of healthy elderly individuals (average age 82) and one of younger students (average age 20). They examine confidence, decisions under uncertainty, differences between willingness to pay and willingness to accept and strategic thinking. They find that the older adults' behaviour is similar to that of young adults, contrary to the notion that economic decision making is impaired with age.

To contextualise our results, the paper further contributes a parsimonious, albeit powerful, dynamic model of retirement. At a superficial level, the model suggest that retirement could affect certain groups of the population more dramatically than others -in which case specific sectoral retirement policies might be advisable. However, the model also emphasises that the dynamics of human capital after retirement are determined by how time use and jobs themselves (as opposed to occupations) contribute to human capital. The problem is that these two parameters are wide-ranging across populations and existing longitudinal studies do not provide enough information or sample sizes to tame the ensuing heterogeneity. Time use data are typically scarce or imprecisely measured in longitudinal studies (for example, the British Household Panel Survey -BHPS- a large study running since 1991, only includes subjects' estimated weekly hours spent in housework, widely defined). Longitudinal studies containing time use diaries from adult populations are also scarce. Occupational and labour supply data is routinely collected, however this data is uninformative about the ways a specific post contributes to one's human capital (e.g. intellectual demands, training, social networks, human relations, development) and is often missing for individuals who retire or are retired at the point when collecting the sample. Job satisfaction, which is generally available in surveys, seems like a crude proxy for the relevant information. In the absence of sufficient information regarding these parameters, empirical estimates of the causal effect of retirement are likely to be heavily dependent on the composition of the sample. Furthermore, a naive separation of samples on the basis of white/blue collar, manual/non-manual worker or formal international classification systems are unlikely to solve the problem. Thus, overall, we believe that available results are not enough to bring the question to a closure.

The policy implication derived from the above discussion is, however, clear: retirement policies should be designed on an understanding of the behavioural implications of retirement and their focus should be on promoting positive behaviours and life styles among retirees, rather than the incentives to stay or not at work for longer. From an empirical point of view, new and more detailed data on time use together with the contribution of specific jobs (as opposed to occupations) to individuals' human capital, is essential to understand and reliably estimate the causal effect of retirement on human capital.

We recognise that our study might be subject to some potential limitations. The main objection relates to the extent to which discontinuity designs are capturing or not permanent effects. Parametric studies simply assume that any estimated effect is permanent -an assumption implicit in the functional form of the postulated models. Discontinuity designs swap parametric restrictions for an explicit statement. However, the bandwidth utilised in our application was of plus-minus five years around retirement which is more indicative of mid-term effects rather than short term effects. Secondly, studies of this type can always benefit from larger sample sizes. In the case of existing UK data, however, the introduction of the Default Retirement Age between 2006 and 2010 and recent increases in the State Pension Age for women and men considerably restrict the amount of data available for analysis of the causal effect of retirement. Nonetheless, our methods are designed to produce tests with exact size in small samples, what endows our results of great robustness. The final limitation of the study is that, in common with previous work, no exploration is made of distributional effects even though it is conceptually feasible that retirement affects inequality in wealth or behaviours. This is left, however, for future research.

References

Aguiar, M. and E. Hurst (2005). Consumption versus expenditure. *Journal of Political Economy 113*, 919–948.

Anderson, M. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool and Early Training projects. *Journal of the American Statistical Association 103*, 1481–1495.

Angrist, J., G. Imbens, and D. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*, 444–455.

Battistin, E., A. Brugiavini, E. Rettore, and G. Weber (2009). The retirement consumption puzzle: Evidence from a Regression Discontinuity approach. *American Economic Review 99*, 2209–2226.

Bonsang, E., S. Adam, and S. Perelman (2012). Does retirement affect cognitive functioning? *Journal of Health Economics 31*, 490–501.

Bound, J. (1991). Self-reported versus objective measures of health in retirement models. *Journal of Human Resources 26*, 106–138.

Bound, J., D. Jaeger, and R. Baker (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association 90*, 443–450.

Bound, J. and T. Waidmann (2007). Estimating the health effects of retirement. Technical report, Michigan Retirement Research Center Research Paper No. UM WP 2007-168.

Browning, M. and M. Gørtz (2012). Spending time and money within the household. *Scandinavian Journal of Economics*.

Calonico, S., M. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for Regression Discontinuity Designs. *Econometrica 82*, 2295 2326.

Card, D., D. Lee, Z. Pei, and Weber (2012). Nonlinear policy rules and the identification and estimation of causal effects in a Generalized Regression Kink Design. *NBER Working Paper 18564*.

Cattaneo, M., B. Frandsen, and R. Titiunik (2014). Randomization inference in the Regression Discontinuity Design: An application to party advantages in the U.S. Senate. *Journal of Causal Inference*.

Charles, K. (2004). Is retirement depressing?: Labor force inactivity and psychological well-being in later life. *Research in Labor Economics 23*, 269–299.

Coe, N. and M. Lindeboom (2008). Does retirement kill you? Evidence from early retirement windows. Technical report, Tilburg University, Center for Economic Research, Discussion Paper 2008-93.

Coe, N. and G. Zamarro (2008). Retirement effects on health in Europe. Technical report, RAND Labor and Population Working Paper W-588.

Davidson, R. and R. MacKinnon (2006). The power of asymptotic and bootstrap tests. *Journal of Econometrics 133*, 421–441.

Dhaval, D., R. I. Rashad, and J. Spasojevic (2008). The effects of retirement on physical mental health outcomes. *Southern Economic Journal 75*, 497–523.

Dong, Y. (2013). Regression Discontinuity without the discontinuity. Technical report, MPRA Paper 25427.

Dong, Y. Y. (2014). Regression Discontinuity Applications with rounding errors in the running variable. *Journal of Applied Econometrics*.

Fé, E. and B. Hollingsworth (2012). Estimating the effect of retirement on mental health via panel discontinuity designs. *MPRA Paper 38162*.

Fé, E. and B. Hollingsworth (2015). Short and long run estimates of the local effects of retirement on health. *Journal of the Royal Statistical Society, A*, Forthcoming.

Fisher, R. (1935). *The Design of Experiments*. Oliver Boyd.

Ganong, P. and S. Jager (2014). A permutatoin test and estimation alternatives for the regression king desing. Technical report.

Hahn, J., P. Todd, and W. van der Klaauw (2001). Identification and estimation of treatment effects with a regression discontinuity design. *Econometrica 69*, 201–209.

Hodges, J. and E. Lehmann (1963). Estimates of location basd on ranks. *Annals of Mathematical Statistics 34*, 598–611.

Johnston, D. and W. Lee (2009). Retiring to the good life? The short term effects of retirement on health. *Economic Letters 103*, 8–11.

Kan, M. Y. and J. Gershuny (2006). Infusing time diary evidence into panel data: An exercise in calibrating time-use estimates for the BHPS. Technical report, Institute for Social and Economic Research, University of Essex.

Kleibergen, F. (2002). Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica 70*, 1781–1803.

Kovalchik, S., C. Camerer, D. Grether, C. Plott, and J. Allman (2004). Aging and decision making: a comparison between neurologically health elderly and young individuals. *Journal of Economic Behaviour and Organization 58*, 79–94.

Lee, D. and D. Card (2008). Regression discontinuity inference with specification error. *Journal of Econometrics 142*, 655–674.

Mazzonna, F. and F. Peracchi (2012). Ageing, cognitive abilities and retirement. *European Economic Review 56*, 691–710.

Neuman, K. (2008). Quit your job and get healthier? The effect of retirement on health. *Journal of Labor Research 29*, 177–201.

O'Brien, P. (1984). Procedures for comparing samples with multiple endpoints. *Biometrics 40*, 1079–1087.

Oreopoulos, P. (2006). Estimating average and local average treatement effects of education when compulsory schooling laws reall matter. *American Economic Review 96*, 152–175.

Porter, J. (2003). Estimation in the Regression Discontinuity model. Technical report, Harvard University.

Rosenbaum, P. (1996). Identification of causal effects using instrumental variables: Comment. *Journal of the American Statistical Association 91*, 465–468.

Rosenbaum, P. and G. Imbens (2005). Robust, accurate confidence intervals with a weak instrument: Quarter of birth and education. *Journal of the Royal Statistical Society, A 168*, 109–126.

Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology 66*, 688–701.

Simonsen, M., L. Skipper, and N. Skipper (2010). Price sensitivity of demand for prescription drugs: Exploiting a Regresion Kink design. *Working Papper 10-01*.

Staiger, D. and J. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*, 557–586.

Stern, S. (1989). Measuring the effect of disability on labour force participation. *Journal of Human Resources 24*, 361–395.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica 53*, 1–16.

Table 1: Descriptive Statistics. Columns 2 and 3 compare means and proportions of individuals who are observed to retire in the sample with those of individuals whose market status remains invariant in the sample. Columns 5 and 6 compare the descriptive statistics of those individuals whose labour market status is retired vs those whose labour status is employed.

	Retires in the Sample			Reports to be retired		
	Total	Yes	No	Total	Yes	No
Age	65.35	62.47	65.53	65.50	70.73	56.99
	9.01	4.44	9.20	8.95	6.61	4.71
%Female	0.50	0.53	0.50	0.50	0.47	0.54
	0.50	0.50	0.50	0.50	0.50	0.50
% Non-white	0.02	0.02	0.02	0.02	0.02	0.02
	0.15	0.14	0.15	0.15	0.14	0.15
Number grandchildren	3.04	2.67	3.06	3.05	3.90	1.67
	4.11	3.45	4.15	4.07	4.44	2.89
Finished school at 14	0.18	0.04	0.19	0.18	0.27	0.03
	0.39	0.19	0.39	0.38	0.45	0.17
Finished school at 16	0.19	0.22	0.19	0.19	0.17	0.22
	0.39	0.42	0.39	0.39	0.37	0.42
Parent had heart disease	0.65	0.67	0.65	0.65	0.66	0.62
	0.48	0.47	0.48	0.48	0.47	0.49
Parent had cancer	0.34	0.37	0.33	0.33	0.33	0.33
	0.47	0.48	0.47	0.47	0.47	0.47
Father had blue collar jobs	0.25	0.28	0.25	0.25	0.25	0.26
	0.44	0.45	0.43	0.43	0.43	0.44
Sample size	6138	366	5772	8302	5138	3164

Table 2: Descriptive Statistics. Columns 2 and 3 compare means and proportions of individuals who are observed to retire in the sample with those of individuals whose market status remains invariant in the sample. Columns 5 and 6 compare the descriptive statistics of those individuals whose labour market status is retired vs those whose labour status is employed.

	Retires in the Sample			Reports to be retired		
	Total	Yes	No	Total	Yes	No
Mental Health*	0.01	-0.11	0.02	0.00	-0.00	0.00
	0.61	0.50	0.62	0.63	0.62	0.64
Cognitive Functioning	-0.02	0.08	-0.03	-0.02	-0.09	0.10
	0.59	0.51	0.59	0.60	0.62	0.53
Social environment	-0.01	0.07	-0.01	0.00	-0.03	0.04
	0.65	0.59	0.65	0.66	0.65	0.67
Relational environment*	0.01	-0.01	0.01	0.00	-0.02	0.04
	0.59	0.59	0.59	0.61	0.60	0.62
Quality of Life*	-0.00	-0.11	0.00	0.00	-0.01	0.02
	0.49	0.42	0.50	0.50	0.50	0.51
Subjective consumption	-0.01	0.03	-0.01	-0.01	-0.03	0.02
	0.50	0.42	0.50	0.51	0.50	0.54
Level of physical activity (Vigorous)	-0.00	0.15	-0.01	-0.00	-0.07	0.11
	0.95	0.97	0.94	1.00	0.97	1.05
Level of physical activity (Moderate)	-0.02	0.18	-0.03	-0.00	-0.04	0.06
	0.94	0.83	0.95	1.00	1.01	0.97
Sample size	6138	366	5772	8302	5138	3164

(*) Higher values denote worse human capital status

Table 3: Local Average Treament Effects of Retirement. Results from a Randomization Inference on the Regression Discontinuity and Kink Regression Designs. The outcomes are Summary Indices transformed into First Differences. The bandwidth was selected as in Calonico et al., 2014.

	RDD		**RKD**		
Outcome	**L.A.T.E**	**P-value**	**L.A.T.E**	**P-value**	**N**
Cognitive Functioning.	0.037	0.713	−0.017	0.431	2471
Cognitive Functioning: Memory	0.012	0.927	0.025	0.386	2470
Social environment	0.176	0.097	−0.044	0.058	2495
Relational environment*	0.121	0.325	−0.046	0.066	2055
Mental Status*	−0.060	0.539	0.030	0.148	2447
Quality of Life*	−0.169*	0.029	0.021	0.165	2227
Quality of Life: Ladder	0.175	0.378	−0.010	0.792	2178
Qualitative Consumption	0.062	0.523	−0.031	0.147	2495
Activity Levels: Vigorous	0.047	0.842	0.001	0.985	2495
Activity Levels: Moderate	0.341	0.059	−0.024	0.519	2495

(*) Higher values denote worse human capital status

Table 4: Bandwidth Selection. We applied the kink and discontinuity RI techniques to a collection of pre-determined variables at different bandwidths. For both techniques, the suggested bandwidth was $h = [-5, 5]$, for which we obtained the estimates and P-values in the table. For $h > 5$ some of the estimators became significant.

	RDD		**RKD**	
Outcome	**L.A.T.E**	**P-value**	**L.A.T.E**	**P-value**
Finished school at age 14 or less.	0.156	0.054	0.003	0.864
Finished school at age 16	−0.049	0.137	−0.078	0.055
Parent had cancer	−0.025	0.787	−0.017	0.738
Parent had heart disease	0.021	0.775	−0.089	0.052
Parent manual/clerical occupation	0.023	0.766	−0.063	0.150
Non-white	−0.010	0.746	0.001	0.939
Has grandchildren	−0.083	0.401	0.001	0.976

Figure 1: Discontinuity in the distribution of retirees by age. The horizontal axis measures years to/from retirement

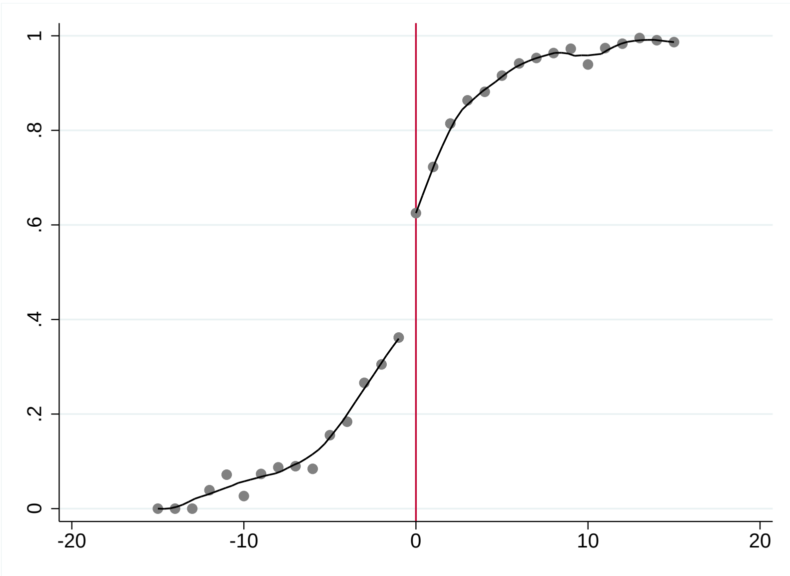
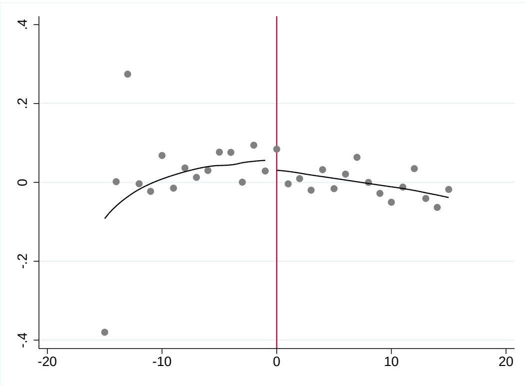
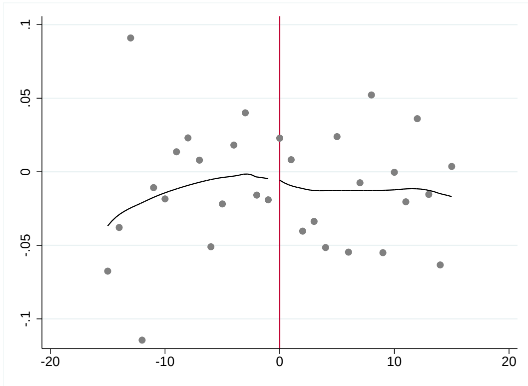


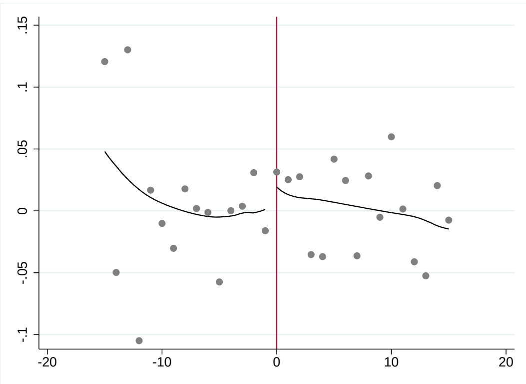
Figure 2: Outcomes. The horizontal axis measures years to/from retirement



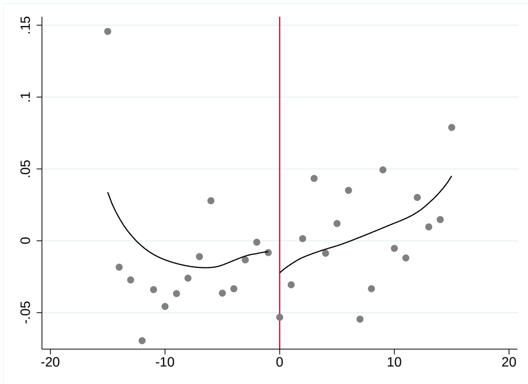
(a) Cognitive Functioning



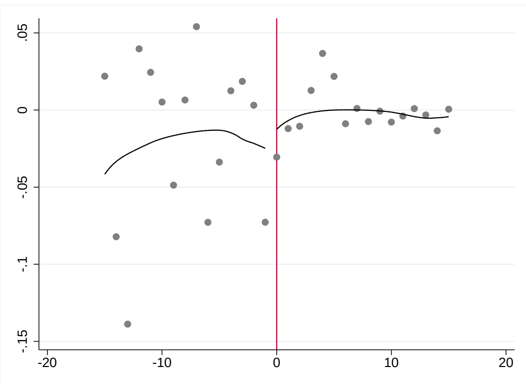
(b) Qualitative Consumption.



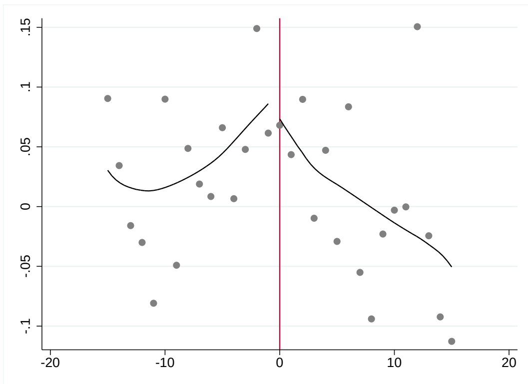
(c) % Social Activity



(d) % Quality of Life

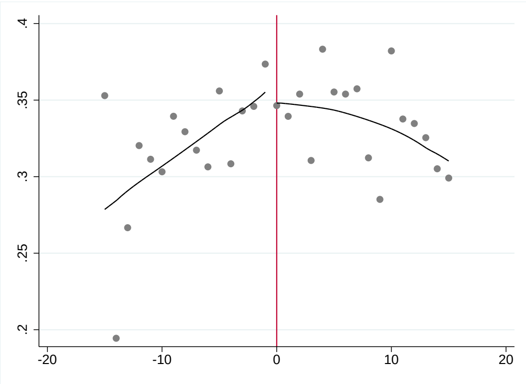


(e) Mental Status

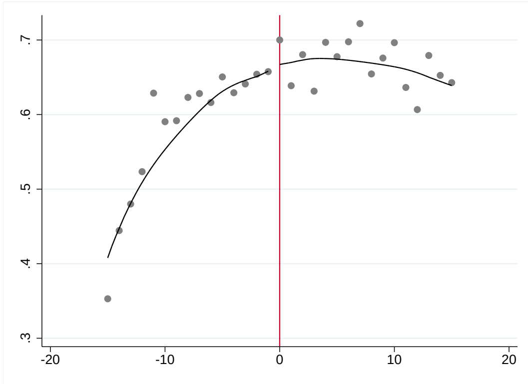


(f) Activity Level (Moderate)

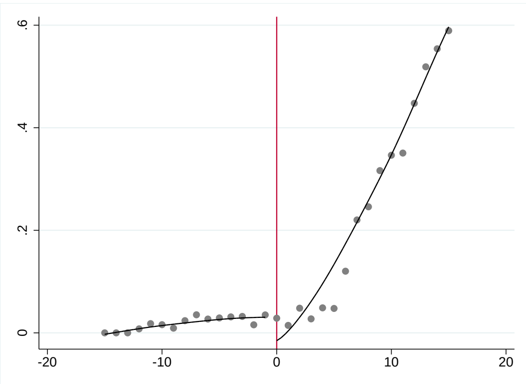
Figure 3: Design Checks. The horizontal axis measures years to/from retirement



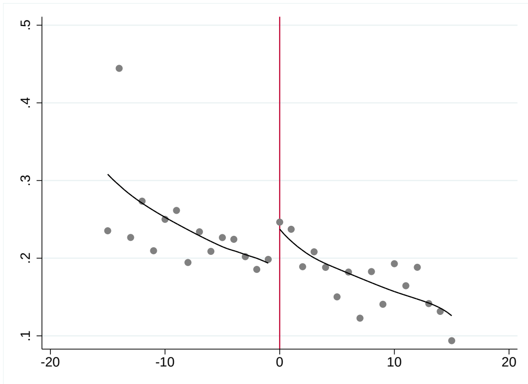
(a) % whose parent had cancer.



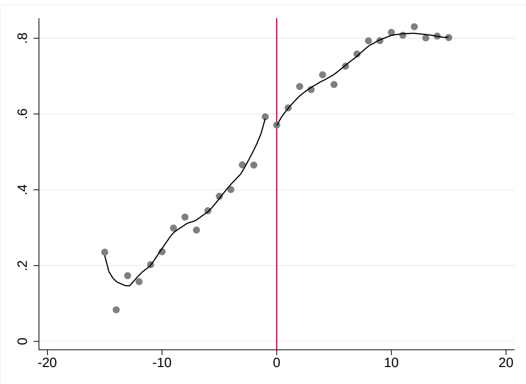
(b) % whose parent had heart disease.



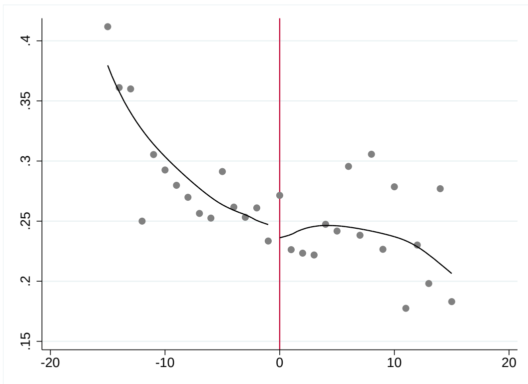
(c) % Finished school age 14



(d) % Finished school age 16



(e) % has grandchildren



(f) % Father in blue collar occupation

Table 5: Components of the Summary Indices.

Cognitive Functioning	• Recall today's date • Immediate word recall from a list of 10 words • Animals mentioned in 60 seconds • Letter cancellation (number correct) • Letter cancellation (number missing) • Attention test (signature on clipboard) • Delayed word recall
Self reported Quality of Life	• Based on the CASP-19, a 19 item questionnaire measuring Control, Autonomy, Self-realisation and Pleasure.
Qualitative Expenditure	• Smokes cigarettes • East out (at least once every few months) • Cut size of meals in the last 12 months • Sometimes/Often shortage of money stops me doing things I want
Physical Activity	• Practices sport/activities vigorous intensity at least 1/3 times a month • Practices sport/activities moderate intensity at least 1/3 times a month
Investment in Socio-cultural Activities	• Goes to cinema at least every few months • Goes to museum at least every few months • Goes to theatre at least every few months • Meets friends at least weekly
Quality of Friendships	• [...] do your friends understand they way you feel • [...] can rely on your friends if you have a serious problem. • [...] can open up to friends if you need to talk [...] • [...] do your friends criticise you • [...] do your friends let you down • [...] do your friends get to your nerves