

MANCHESTER  
1824

The University  
of Manchester

**Economics**  
**Discussion Paper Series**  
**EDP-1508**

**On The Local Causal Effects of Retirement on  
Human Capital**

Eduardo Fé

Mario Pezzino

July 2015

Economics  
School of Social Sciences  
The University of Manchester  
Manchester M13 9PL

# On The Local Causal Effects of Retirement on Human Capital.

Eduardo F e<sup>\*†</sup> <sup>1</sup> and Mario Pezzino<sup>‡</sup> <sup>2</sup>

<sup>1</sup>*Blavatnik School of Government, University of Oxford, U.K.*

<sup>2</sup>*Economics, University of Manchester, U.K.*

## Abstract

We investigate if retirement has short or long term effects on human capital. As in most previous work, we estimate long term effects under parametric restrictions. In addition we accompany the results with short run estimates produced with a new Randomization Inference framework for Regression Discontinuity Designs when Panel Data are available. This new method rest on very weak assumptions, is robust to weak instruments and identification is unaffected by the discrete-running variable problem. We find that retirement does not significantly affect any of the dimensions considered in the short or long run. To gain some insights on the nature and scope of these results, we present a dynamic programming model which emphasises that job's contributions and individuals' allocation of time before and after retirement are critical determinants of the consequences of retirement on human capital. Our model suggests that (1)leaving aside standard public finance arguments, policy reforms should primarily be aimed at retirees' behaviour and life styles (2) existing estimates are likely to be highly sensitive to sample composition and (3) new detailed data reporting the time allocation of individuals, occupational environment and the contribution of occupations to the development of human capital are critical to provide informative empirical estimates of the consequence of retirement.

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

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

---

\*Blavatnik School of Government, University of Oxford. 10 Merton Street, Oxford, OX1 4JJ. email:eduardo.fe@bsg.ox.ac.uk.

†The authors thank Michael Caputo for very helpful comments.

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

# 1 Introduction.

Over the last decade, a substantial amount of research has been devoted to the estimation of the causal effect of retirement on human capital. This 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 ultimate implications of retirement for a population. For example, if retirement impoverishes mental health or cognitive ability, individuals' increased vulnerability in the older age would be accompanied by an increase 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.

A review of the empirical literature about this topic reveals that studies looking at variations in health after retirement are largely inconclusive<sup>1</sup> while the only two existing studies on cognition<sup>2</sup> find that retirement has a significant negative effect on cognitive functioning. Published work has implicitly focused on the estimation of the long term effects of retirement, but identification in this context requires strong parametric assumptions regarding the path of human capital in the absence of retirement. These assumptions are likely to play an important role when explaining the magnitude and, crucially, significance of results. Furthermore, because human capital itself is an important determinant of retirement (Stern, 1989; Bound, 1991; Disney et al., 2006) researchers have put forward a variety of instrumental variable methods to circumvent reverse causality. Although these strategies are compelling, estimation under different parametric methods decreases the comparability of results. This problem is also exacerbated because studies have analysed retirement in countries with rather significant

---

<sup>1</sup>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, while Johnston and Lee (2009) conclude that retirement improves individuals' sense of well-being and mental health, but not necessarily physical health.

<sup>2</sup>Bonsang et al. (2012) and Mazzonna and Peracchi (2012).

idiosyncrasies. For instance, studies based on UK data need to bear in mind the introduction of a Default Retirement Age between 2006 and 2010, by which employers could effectively dismiss employees reaching age 65 (what introduces a considerable problem of selection on unobservable variables). Studies based on US data need to address the fact that the predominance of private health insurance might affect retirement via credit constraints, again reducing the comparability of retirees and employees at any given age. Finally when it comes to outcomes, received work has been characterised by a focus on separate dimensions of human capital (either health or cognitive functioning) even though one can envisage retirement affecting several, often related, dimensions of life, with effects not necessarily being direct. For example, a direct negative effect of retirement on the frequency of social interactions could have an indirect effect on mental health.

In view of these limitations, the first contribution of this paper is to present a set of new econometric techniques to obtain estimates of the local causal effect of retirement on a series of indicators of human capital (including, among others, health and cognitive functioning). We keep national idiosyncrasies fixed by restricting the study to England, and use the English Longitudinal Study of Ageing (ELSA). As in previous work, we provide estimates from a parametric regression model which intends to capture long term effects. However, we accompany these results with estimates obtained from a new panel data extension of the Randomization Inference (RI) framework for Regression Discontinuity Designs (RDD) devised by Cattaneo, Frandsen, and Titiunik (2014) (see also Fisher, 1935; Rosenbaum, 1996). These estimates are obtained under minimal assumptions, what increases their generality and comparability with future work. Using a RI setting allow us to draw inferences from test with exact size and under very weak restrictions akin to the Independence-Exclusion-SUTVA trinity in Angrist et al. (1996)<sup>3</sup>. By taking a panel data perspective, we can further relax the assumptions in Cattaneo et al. (2014) to allow for endogeneity due to time invariant unobserved heterogeneity. More crucially, the validity of the method does not hinge on whether the running variable (age in our case) is measured discretely (as in our data) or continuously (as is generally expected in RDD)<sup>4</sup>.

With these new methods, we find that retirement does not significantly affect any of the

---

<sup>3</sup>Exact inference is a very important advantage for RDD where data within the identifying window width tend to be scarce, so that asymptotic approximations are likely to be misleading

<sup>4</sup>Thus, the discrete-running variable problem (Lee and Card, 2008) is avoided.

dimensions considered (including cognitive functioning) -although there might be a non-lasting short term effect on self-reported quality of life (which improves a bit). 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. 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. First, our model suggest that heterogeneity is likely to render existing estimates of the causal effects of retirement very sensitive to sample composition. Second, new data illustrating the contribution of a particular job to an individual's human capital, together with detailed time use data, are crucial to provide meaningful empirical estimates. Finally, retirement policies need to take into consideration people's behaviour and life styles upon retirement, as well as rely on an understanding of the role played by different jobs in the dynamics of human capital. Policy reforms that ignore these aspects might have 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 much technical material deferred to the Appendix, including a Monte Carlo experiment). 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 and Institutional Framework.

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, 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 invalidate 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 disabled. 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 6.

As in other popular data sets<sup>5</sup>, 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<sup>6</sup>. Most of these tasks involve memory recall exercises and solving numerical problems and thus are primarily suggestive about an individual's fluid intelligence -which from the point of view of retirement is, potentially, the most vulnerable dimension of cognition<sup>7</sup>. 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

---

<sup>5</sup>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.

<sup>6</sup>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).

<sup>7</sup>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 cognitive 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.

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<sup>8</sup>. Although private 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<sup>9</sup>- 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 measures of human capital outcomes, 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.

---

<sup>8</sup>Much of this section draws from Fé and Hollingsworth (2012).

<sup>9</sup>Pensions Policy Institute <https://www.pensionspolicyinstitute.org.uk/default.asp?p=81>.



Thus, additional identifying assumptions are required. These assumptions are the subject of the next section.

### 3 Estimation and inference.

In this section we outline our identification strategy. In order to improve exposition, technical details are given in the Appendix, and only the key features of the methods are discussed here. We will make use of 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  $Z_{it}$ .

Our study commences by providing estimates of the long run effects of retirement. The identification strategy in this setting is an extension of that in Bound and Waidmann (2007) (see Fé and Hollingsworth, 2012) and investigates if there are changes in the slope or level of outcomes at the SPA. Because we are estimating long term effects, the analysis has to be extended to a wide time interval around the SPA. As a result, successful identification requires us to introduce explicit assumptions about the structure of the correlation between age and human capital in order to ensure that SPA is a valid exogenous predictor of retirement. In particular, we define the model,

$$Y_{it} = \beta_1 R_{it} + \dots \beta_p R_{it}^p + \tau_1 Z_{it} + \tau_2 Z_{it} R_{is} + \alpha_i + \varepsilon_{it} \quad (3.1)$$

where the polynomials in age intend to absorb any correlations between health and age (so that any residual effect of SPA on health must be due to retirement alone). The parameters of interest are  $\tau_1$  and  $\tau_2$  which measure the variations in human capital experienced by those individuals whose retirement status changes due to changes in eligibility for the SPA. We estimate the model in first differences, under three scenarios:  $\tau_1 = 0$ ,  $\tau_2 = 0$  and  $\tau_1 \neq 0$ ,  $\tau_2 \neq 0$ .

Inferences are based on a residual-restricted<sup>10</sup>, cluster-robust wild bootstrap scheme similar to that in Fé and Hollingsworth (2012) (see also Cameron et al., 2008) applied to the  $F$ -test of the null hypothesis  $\tau_1 = 0$ ,  $\tau_2 = 0$  or  $\tau_1 = \tau_2 = 0$ .

### 3.1 A Randomization Inference framework for RDD with Panel Data

The identification assumption (3.1) is along the lines of previous assumptions found in the retirement literature. But, despite allowing for considerable structural flexibility, they are strong and can determine the magnitude and significance of results in spurious ways. So we next explore identification under much weaker conditions. In particular, we extend recent work by Cattaneo, Frandsen, and Titiunik (2014) and present a randomization inference framework for Regression Discontinuity Designs when panel data are available. This section presents the motivation for the method and an outline of its implementation. All the technical details are deferred to the Appendix.

Unlike in the standard program evaluation framework or the above long-term model, 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 in accordance to a certain mechanism (compare Rubin, 1974; Angrist, Imbens, and Rubin, 1996 vs 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 random mechanism 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 con-

---

<sup>10</sup>See Davidson and MacKinnon (2010).

fidence intervals (under stronger regularity conditions; see Hodges and Lehmann, 1963; Rosenbaum, 1996). More importantly, as noted by Rosenbaum and Imbens (2005), the immediate consequence of exact inference is that inferential procedures exhibit a remarkable robustness to the quality of the instrument set, with empirical sizes marginally varying around nominal level even with weak instruments (see, in contrast, Staiger and Stock, 1997; Kleibergen, 2002; Davidson and MacKinnon, 2006).

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<sup>11</sup>-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 ). In RDD, the attractiveness of exact tests cannot be overstated, given the typically small samples sizes available in neighbourhoods of the cut-off point. The technical details of the framework can be found in the Appendix.

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 is the norm in the classic panel data literature, 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)<sup>12</sup>. 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 the Appendix.

---

<sup>11</sup>Stable Unit Treatment Value Assumption.

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

## 4 Empirical Analysis.

Tables 1 and 2 present some descriptive statistics of the sample. Table 1 focuses on pre-determined 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<sup>13</sup> (the Butler Act) which increased compulsory schooling leaving age to 15 in 1947. This had affected to those individuals who were 69 or younger in 2002 (the first wave of ELSA), by increasing their schooling by one year. Thus, the Butler law explains the difference in schooling among groups -because the group of individuals who retire during the sample are younger. The difference in schooling will not be of concern for our short term estimation strategy. However it must be taken into account for the long-term identification strategy (which uses a long time span around the SPA). This also emphasises one of the weaknesses of parametric long term strategies: namely that, at least one section of the population, might not be fully comparable in certain traits. In our long term analysis we will include a binary indicator to identify individuals who were affected by the Butler Act.

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

---

<sup>13</sup>See Oreopoulos (2006).

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

Tables 3 and 4 report the short and long term 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 5. The selected bandwidth was  $[-5, 5]$ <sup>14</sup>.

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 short-run coefficient of self-reported quality of life is negative and significant in the short run (implying an improvement in perceived quality of life), but the significance of this effect disappears in the long 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? To gain insight into this question and better

---

<sup>14</sup>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.

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 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,

$$A_t = \rho A_{t-1} + w(H_{t-1})T_{t-1}^w \quad (5.2)$$

$$H_t = \delta H_{t-1} + \theta T_{t-1}^a + \gamma T_{t-1}^w \quad (5.3)$$

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  $\theta, \gamma \in \mathbb{R}$  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<sup>15</sup>.

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

$$\max_{T_t^h, T_t^a} \left[ \sum_{t=1}^T \beta^t U(T_t^h, T_t^a; H_t) + \beta^T A_T \right] \quad (5.4)$$

subject to (5.2) and (5.3), 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 problem 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  $\theta$  and  $\gamma$  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<sup>16</sup>,

$$U(T_t^h, T_t^a; H_t) = H_t + \alpha_h T_t^h - \frac{(T_t^h)^2}{2} + \alpha_a T_t^a - \frac{(T_t^a)^2}{2}$$

$$w(H_1) = H_1$$

where  $\alpha_i > 0$ ,  $i = h, a$ , represent the marginal benefit that the agent can obtain from investing

---

<sup>15</sup>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.

<sup>16</sup>The fact that  $U(T_t^h, T_t^a; H_t)$  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).

time in household ( $\alpha_h$ ) and leisure activities ( $\alpha_a$ ). 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 (5.4).

**Proposition 1** *The (interior) solution of problem (5.4) is:*

- $t = 1$

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

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

- $t = 2, \dots, T - 1$

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

$$T_{T-k}^a = T - T_{T-k}^h$$

$$k = 1, \dots, T - 2$$

- $t = T$

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

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 ( $\alpha_h$  and  $\alpha_a$ ) and the dynamic effect that these controls have on the system via  $H_t$  (parameters  $\gamma$ ,  $\theta$  and  $\delta$ ) and  $A_t$  (parameter  $\rho$ ), discounted by  $\beta$ . If  $\gamma$  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  $\theta$ ). 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 ( $t = 2, \dots, T$ ) 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  $T_t^h = T_t^a = \frac{1}{2}(\bar{T} + a_h - a_a)$  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  $\gamma$  and  $\theta$ . 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 produce 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  $\gamma$ ,  $\theta$  and agents' decisions affect human capital before and after retirement. From the information in Proposition 1, (5.2) and (5.3), 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:

$$H_2 = \frac{1}{\beta\rho^2} [\beta^T \rho^T H_1 (2\gamma - \theta)] + \delta H_1 - \gamma (a_a + a_h - T) + \theta a_a + \beta\theta \left( 1 + \sum_{i=1}^{T-2} \beta^i \delta^i \right) (2\gamma - \theta) \quad (5.5)$$

$$H_3 = \frac{1}{2}\theta \left( T - a_h + a_a + \theta \sum_i^k \beta^i \delta^{i-1} \right) + \delta H_2 \quad (5.6)$$

The equilibrium expression of  $H_3$  reveals why retirement does not have to necessarily produce a negative effect on human capital. Indeed suppose that  $\delta$  were sufficiently high (e.g. close to 1). Then, if  $\theta$  is positive and sufficiently large, human capital would *increase* after retirement. Of course, if  $\theta$  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  $\gamma$  and  $\theta$  may have on the equilibrium levels of human capital before and after retirement,

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

$$\frac{dH_2}{d\gamma} = -a_h - a_a - 2\beta^{T-1} \rho^{T-2} H_1 + T + 2\beta\theta \left( 1 + \sum_{i=1}^{T-2} \beta^i \delta^i \right) \quad (5.8)$$

$$\begin{aligned} \frac{dH_3}{d\theta} &= \frac{1}{2} \left[ T - a_h + a_a + 2\theta \sum_i^k \beta^i \delta^{i-1} \right. \\ &\quad \left. + 2\delta \left( \frac{-\beta^T \rho^T H_1}{\beta\rho^2} + a_a + 2\beta(\gamma - \theta) \left( 1 + \sum_{i=1}^{T-2} \beta^i \delta^i \right) \right) \right] \end{aligned} \quad (5.9)$$

$$\frac{dH_3}{d\gamma} = \delta \frac{dH_2}{d\gamma} \quad (5.10)$$

These equations reveal that human capital can increase (before and after retirement) with  $\theta$  when  $a_a$  is large, the difference  $(\gamma - \theta)$  is positive and large and  $H_1$  is small. Similarly, human capital increases with  $\gamma$  when both  $a_h$  and  $a_a$  are small,  $\theta$  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  $\gamma$  and  $\theta$ . 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  $\gamma$  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 ( $\theta$ ) is likely to be wide ranging. Therefore, in the absence of sufficient information regarding  $(\theta, \gamma)$ , 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 assist 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<sup>17</sup>. 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 just 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  $\theta$  is positive and large

---

<sup>17</sup>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 Socioeconomic 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.

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,  $\gamma$ , 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)  $\theta$  might still experience an increase in human capital after retirement. This remark calls for sector-based studies and reforms.

## 6 Conclusion.

In this study, we provide estimates of the long-term and short-term causal effect of retirement on a series of indicators capturing different aspects life (quality of life, cognitive ability, relationships, physical, cultural and social activity and qualitative expenditure). All these indicators shape and reveal different dimensions of human capital and are likely to influence retirement decisions to some extent. Long term estimates are obtained (as in previous work -see, for example, Charles, 2004; Bonsang et al., 2012; Mazzonna and Peracchi, 2012) from a parametric model which specifically models human capital dynamics in the absence of retirement. Short term effects were obtained from a new extension of the work in Cattaneo et al. (2014). In particular, we put forward a Randomization Inference framework for Regression Discontinuity Designs which handles within individual correlations due to time invariant heterogeneity. In this new framework identification relies on comparisons of outcomes within individuals (which reduces the contentiousness of the underlying assumption of continuous potential outcomes). Crucially, the method yields test 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. As in most previous work (including Bound and Waidmann, 2007; Coe and Lindeboom, 2008; Charles, 2004; Coe and Zamarro, 2008; Neuman, 2008; Bonsang et al., 2012; Mazzonna and Peracchi, 2012), our methods are designed to estimate the Local Average Treatment Effect of retirement -on those individuals whose retirement status changed coinciding with a change in eligibility for state pension age.

We apply the new estimation framework to the English Longitudinal Study of Ageing

(ELSA), and find that retirement does not have any short-term or long term effect on self-reported quality of life, qualitative expenditure, social, physical or cultural activity or the relationships with friends and partners. Unlike in Mazzonna and Peracchi (2012) and Bonsang et al. (2012) we do not find that retirement affects cognitive functioning. Our results are aligned with those in Kovalchik et al. (2004) who devise experiments on economic decisions with 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 the theory of mind (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. That self-reported qualitative expenditure does not vary also goes along the conclusions in Aguiar and Hurst (2005) (who also find that consumption does change with retirement).

Despite of the accuracy of our econometric methods and the quality of the sample, we believe that our results are enough to bring the question to a closure (although a non-significant treatment effect is not entirely surprising). To explain why, the paper introduces 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. 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.

Overall, the policy implication of this paper is that, if economic arguments are ignored, retirement policies should be primarily aimed at understanding and modifying retirees' behaviour and life styles 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.

## 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.
- 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. 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.
- Cameron, C., J. Gelbach, and D. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*. 90, 414–427.
- 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 J. MacKinnon (2010). Wild bootstrap tests for IV regression. *Journal of Business and Economic Statistics* 28(1), 128–144.
- 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.
- Disney, R., C. Emmerson, and M. Wakefield (2006). Ill health and retirement in Britain: A panel data-based analysis. *Journal of Health Economics*. 25, 621–649.
- 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*.
- Fisher, R. (1935). *The Design of Experiments*. Oliver Boyd.
- Hodges, J. and E. Lehmann (1963). Estimates of location based 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.
- 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 treatment effects of education when compulsory schooling laws really matter. *American Economic Review* 96, 152–175.
- 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.
- 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.

**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 Treatment Effects of Retirement. Results from a Randomization Inference on a Regression Discontinuity Design. The outcomes are Summary Indices transformed into First Differences. The bandwidth was selected as in Calonico et al., 2014.

<b>Outcome</b>	<b>L.A.T.E</b>	<b>P-value</b>	<b>N</b>	<b>T</b>
Cognitive Functioning.	0.037	0.713	2471	2
Cognitive Functioning: Memory	0.012	0.927	2470	2
Social environment	0.176	0.097	2495	2
Relational environment*	0.121	0.325	2055	2
Mental Status*	-0.060	0.539	2447	2
Quality of Life*	-0.169*	0.029	2227	2
Quality of Life: Ladder	0.175	0.378	2178	2
Qualitative Consumption	0.062	0.523	2495	2
Activity Levels: Vigorous	0.047	0.842	2495	2
Activity Levels: Moderate	0.341	0.059	2495	2

---



---

(\*) Higher values denote worse human capital status

**Table 4: Longer Term Effects. OLS in First Differences.**

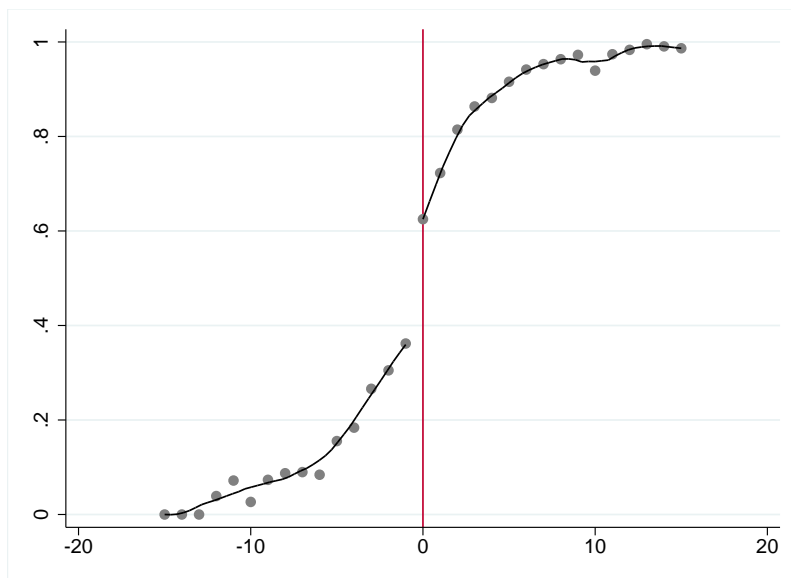
Outcome	Model 1		Model 2		Model 3		NT	
	Eligibility	Interaction	F-Test p-value	Eligibility	F-Test P-value	Interaction		F-Test P-value
Cognitive Functioning.	-0.012	-0.034	0.435	0.008	0.782	-0.032	0.101	8215
Cognitive Functioning: Memory	0.006	-0.007	0.912	0.010	0.717	-0.009	0.770	8214
Social environment	0.034	-0.010	0.124	0.039	0.062	-0.017	0.254	8302
Relational environment*	0.017	-0.016	0.475	0.027	0.441	-0.020	0.388	6514
Mental Status*	0.006	0.032	0.119	-0.013	0.428	0.031	0.061	8135
Quality of Life*	-0.028	0.006	0.195	-0.032	0.123	0.012	0.396	7209
Quality of Life: Ladder	0.038	0.006	0.661	0.034	0.357	-0.003	0.930	6910
Qualitative Consumption	0.014	-0.023	0.140	0.028	0.338	-0.026	0.049	8302
Activity Levels: Vigorous	0.022	0.010	0.925	0.016	0.787	0.005	0.886	8302
Activity Levels: Moderate	-0.005	-0.037	0.414	0.016	0.448	-0.035	0.152	8302

(\*) Higher values denote worse human capital status

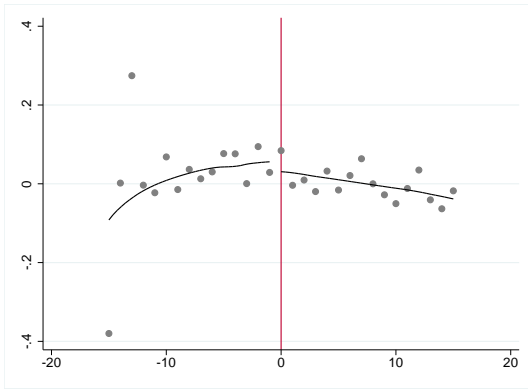
**Table 5:** Bandwidth Selection. We applied the Randomization Inference technique to a collection of pre-determined variables at different bandwidths. The table reveals the estimated coefficients and p-values when  $h = [-5, 5]$ . For  $h > 5$  some of the estimators became significant.

<b>Outcome</b>	<b>L.A.T.E</b>	<b>P-value</b>	<b>N</b>
Finished school at age 14 or less.	0.156	0.054	3215
Finished school at age 16	-0.049	0.137	3215
Parent had cancer	-0.025	0.787	3215
Parent had heart disease	0.021	0.775	3215
Parent manual/clerical occupation	0.023	0.766	3215
Non-white	-0.010	0.746	3215
Has grandchildren	-0.083	0.401	3151

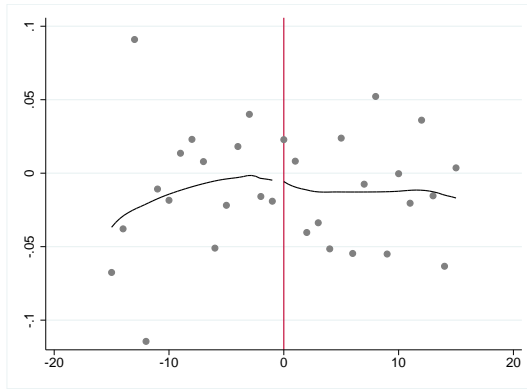
**Figure 1:** Discontinuity in the distribution of retirees by age. The horizontal axis measures years to/from retirement



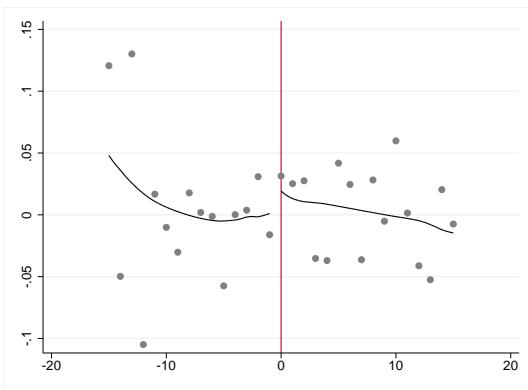
**Figure 2:** Outcomes. The horizontal axis measures years to/from retirement



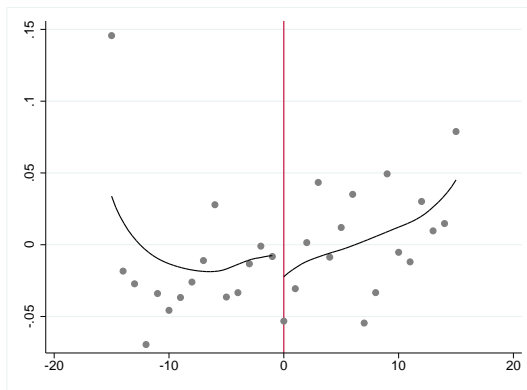
(a) Cognitive Functioning



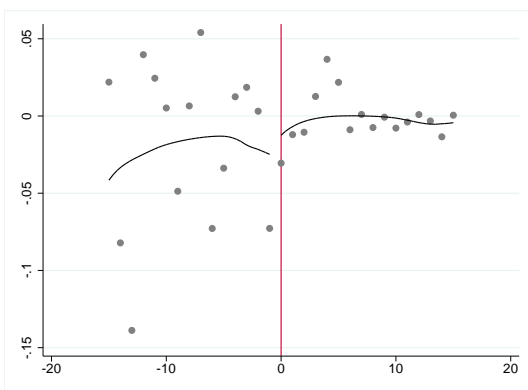
(b) Qualitative Consumption.



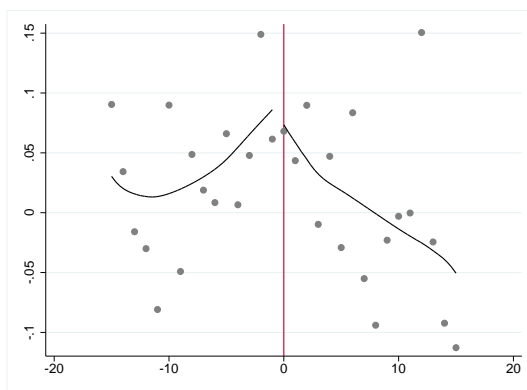
(c) % Social Activity



(d) % Quality of Life



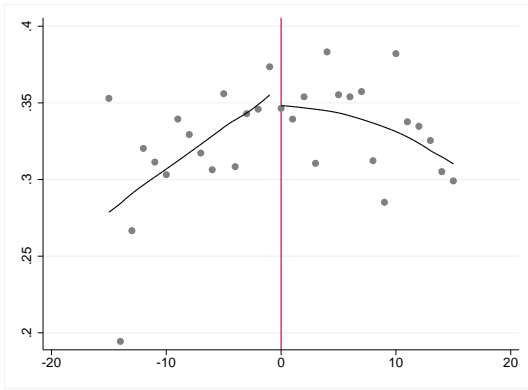
(e) Mental Status



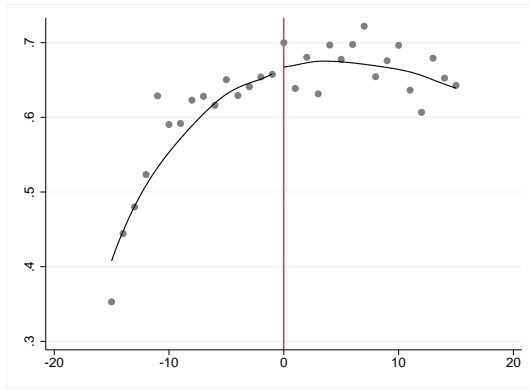
(f) Activity Level (Moderate)



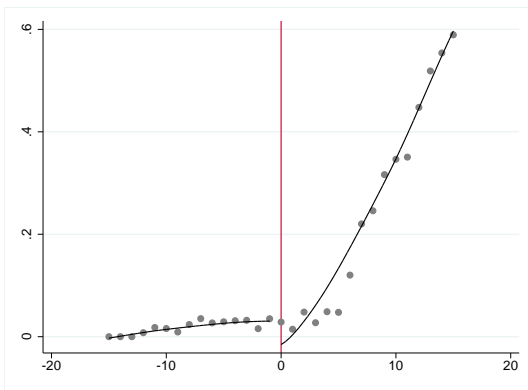
**Figure 3:** Design Checks. The horizontal axis measures years to/from retirement



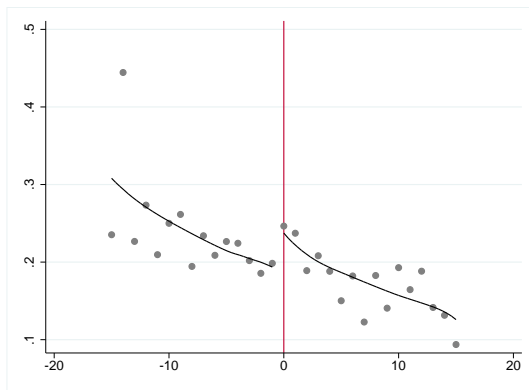
(a) % whose parent had cancer.



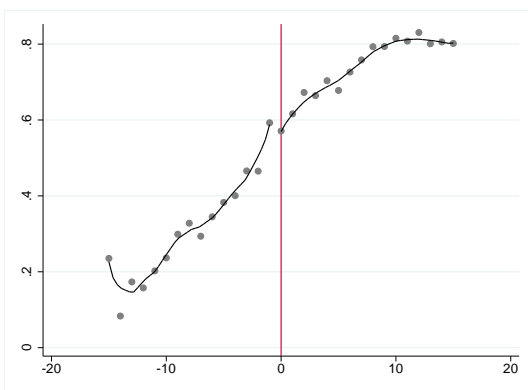
(b) % whose parent had heart disease.



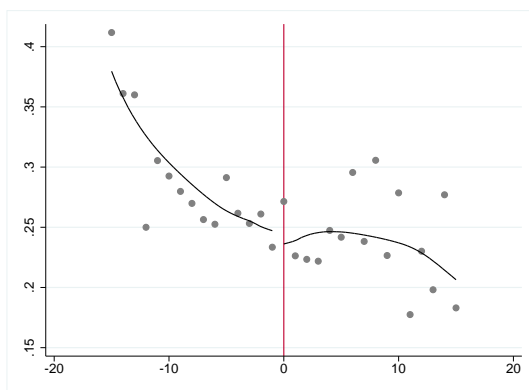
(c) % Finished school age 14



(d) % Finished school age 16



(e) % has grandchildren



(f) % Father in blue collar occupation

**Table 6:** Components of the Summary Indices.

Cognitive Functioning	<ul style="list-style-type: none"> <li>• Recall today’s date</li> <li>• Immediate word recall from a list of 10 words</li> <li>• Animals mentioned in 60 seconds</li> <li>• Letter cancellation (number correct)</li> <li>• Letter cancellation (number missing)</li> <li>• Attention test (signature on clipboard)</li> <li>• Delayed word recall</li> </ul>
Self reported Quality of Life	<ul style="list-style-type: none"> <li>• Based on the CASP-19, a 19 item questionnaire measuring Control, Autonomy, Self-realisation and Pleasure.</li> </ul>
Qualitative Expenditure	<ul style="list-style-type: none"> <li>• Smokes cigarettes</li> <li>• Eat out (at least once every few months)</li> <li>• Cut size of meals in the last 12 months</li> <li>• Sometimes/Often shortage of money stops me doing things I want</li> </ul>
Physical Activity	<ul style="list-style-type: none"> <li>• Practices sport/activities vigorous intensity at least 1/3 times a month</li> <li>• Practices sport/activities moderate intensity at least 1/3 times a month</li> </ul>
Investment in Socio-cultural Activities	<ul style="list-style-type: none"> <li>• Goes to cinema at least every few months</li> <li>• Goes to museum at least every few months</li> <li>• Goes to theatre at least every few months</li> <li>• Meets friends at least weekly</li> </ul>
Quality of Friendships	<ul style="list-style-type: none"> <li>• [...] do your friends understand they way you feel</li> <li>• [...] can rely on your friends if you have a serious problem.</li> <li>• [...] can open up to friends if you need to talk [...]</li> <li>• [...] do your friends criticise you</li> <li>• [...] do your friends let you down</li> <li>• [...] do your friends get to your nerves</li> </ul>

# Appendix to *On The Local Causal Effects of Retirement on Human Capital*.

Eduardo F e\* <sup>1</sup> and Mario Pezzino<sup>† 2</sup>

<sup>1</sup>*Blavatnik School of Government, University of Oxford, U.K.*

<sup>2</sup>*Economics, University of Manchester, U.K.*

## 1 Randomization Inference for the Regression Discontinuity Design.

In this section we present a formal discussion of the work by Cattaneo, Frandsen, and Titiunik (2014). Let  $\mathbf{z}_{W_0}$  and  $\mathbf{d}_{W_0}$  be subvectors of  $\mathbf{Z} = \mathbf{z}$  and  $\mathbf{D} = \mathbf{d}$  correspondign to observations with  $R_{it} \in W_0$ . Randomization Inference can be directly applied in RDD if the following conditions are met.

**Assumption 1** *Cattaneo, Frandsen, and Titiunik (2014). There exists a neighborhood  $W_0 = [\underline{r}; \bar{r}]$  with  $\underline{r} < r_0 < \bar{r}$  such that for all  $i, t$  with  $R_{it} \in W_0$ ,*

### 1. *Local Randomization*

(a)  $F_{R_{it}|R_{it} \in W_0}(r) = F(r)$

(b)  $d_{it}(\mathbf{r}) = d_{it}(\mathbf{z}_{W_0})$  and  $y_{it}(\mathbf{r}, \mathbf{d}) = y_{it}(\mathbf{z}_{W_0}, \mathbf{d}_{W_0})$

### 2. *Local Stable Unit Treatment Value Assumption (L-SUTVA)*.

(a) *If  $z_{it} = \tilde{z}_{it}$ , then  $d_{it}(\mathbf{z}_{W_0}) = d_{it}(\tilde{\mathbf{z}}_{W_0})$*

(b) *If  $z_{it} = \tilde{z}_{it}$  and If  $d_{it} = \tilde{d}_{it}$ , then  $y_{it}(\mathbf{z}_{W_0}, \mathbf{d}_{W_0}) = y_{it}(\tilde{\mathbf{z}}_{W_0}, \tilde{\mathbf{d}}_{W_0})$*

---

\*Blavatnik School of Government, University of Oxford. 10 Merton Street, Oxford, OX1 4JJ. email:eduardo.fe@bsg.ox.ac.uk.

<sup>†</sup>Economics, University of Manchester. Oxford Road, Manchester, M13 9PL. email:mario.pezzino@manchester.ac.uk

3. **Local Exclusion.**  $y_{it}(\mathbf{z}, \mathbf{d}) = y_{it}(\tilde{\mathbf{z}}, \mathbf{d})$  for all  $\mathbf{z}, \tilde{\mathbf{z}}$ .

Assumption A1 allows us to write potential treatment and outcomes as  $d_{it}(z_{it})$  and  $y_{it}(z_{it}, d_{it})$  for every  $i, t$  with  $R_{it} \in W_0$ . In our context, the exclusion restriction implies that potential outcomes are affected by assignment (crossing SPA age) only through its effect on retirement decisions. 1.1.a. implies that age can be considered allocated as random within  $W_0$ . The assumption would be violated if, for example, parenthood decisions would be correlated with potential outcomes (for the cohorts under study, born between 1940 and 1960, 1.1.a. does not appear a too unreasonable assumption). Assumption 1.1.b. implies that potential outcomes are not affected by age other than through its effect on location with respect to  $r_0$ . Furthermore, potential outcomes of individual  $i$  at time  $t$  are not correlated with other potential outcomes outside  $W_0$ . Assumption 1.2. further implies that, within  $W_0$ , potential treatments and potential outcomes are also uncorrelated.

To extend the analysis to panel data, we begin by introducing the following Assumption.

**Assumption 2** *Let  $\theta_i$  and  $\eta_i$  be i.i.d random variables. Then, for all  $i, t$  with  $R_{it} \in W_0$  and any particular value of  $\theta_i, \eta_i$ , the potential treatment,  $d_{it}(\mathbf{z}) + \eta_i$  and potential outcome  $y_{it}(\mathbf{z}, \mathbf{d}) + \theta_i$  satisfy Assumption 1).*

Note that under this assumption, potential outcomes  $d_{it}(\cdot)$  and  $y_{it}(\cdot)$  are not unconditionally fixed -although for given  $\theta_i, \eta_i$  they are, and satisfy Assumption 1. Let  $\Delta\mathbf{Z}$ ,  $\Delta\mathbf{D}$  be the  $N(T-1)$  vectors with elements  $\Delta Z_{it} = Z_{it} - Z_{i,t-1}$  and  $\Delta D_{it} = D_{it} - D_{i,t-1}$  respectively;  $\Delta Y_{it} = Y_{it} - Y_{i,t-1}$ . Define

$$\Delta d_{it}(\Delta\mathbf{z}) = d_{it}(\mathbf{z}) - d_{it-1}(\mathbf{z}) \tag{1.1}$$

$$\Delta y_{it}(\Delta\mathbf{z}, \Delta\mathbf{d}) = y_{it}(\mathbf{z}, \mathbf{d}) - y_{it-1}(\mathbf{z}, \mathbf{d}) \tag{1.2}$$

with  $\Delta Y_{it} = \Delta y_{it}(\Delta\mathbf{Z}, \Delta\mathbf{D})$  and  $\Delta D_{it} = d(\Delta\mathbf{Z})$ . Then, analysis can proceed by condensing Assumptions 1 and 2 within the next (slightly stronger) assumption.

**Assumption 3** *There exists a neighborhood  $W_0 = [\underline{r}; \bar{r}]$  with  $\underline{r} < r_0 < \bar{r}$  such that for all  $i, t$  with  $R_{it} \in W_0$ ,*

### 1. Local Randomization

$$(a) F_{R_{it}|R_{it} \in W_0}(r) = F(r)$$

$$(b) \Delta d_{it}(\mathbf{r}) = \Delta d_{it}(\Delta \mathbf{z}_{W_0}) \text{ and } \Delta y_{it}(\mathbf{r}, \Delta \mathbf{d}) = \Delta y_{it}(\Delta \mathbf{z}_{W_0}, \Delta \mathbf{d}_{W_0})$$

## 2. *Local Stable Unit Treatment Value Assumption (L-SUTVA).*

$$(a) \text{ If } \Delta z_{it} = \Delta \tilde{z}_{it}, \text{ then } \Delta d_{it}(\mathbf{z}_{W_0}) = \Delta d_{it}(\tilde{\mathbf{z}}_{W_0})$$

$$(b) \text{ If } \Delta z_{it} = \Delta \tilde{z}_{it} \text{ and If } \Delta d_{it} = \Delta \tilde{d}_{it},$$

$$\text{then } \Delta y_{it}(\Delta \mathbf{z}_{W_0}, \Delta \mathbf{d}_{W_0}) = \Delta y_{it}(\Delta \tilde{\mathbf{z}}_{W_0}, \Delta \tilde{\mathbf{d}}_{W_0})$$

## 3. *Local Exclusion.* $\Delta y_{it}(\Delta \mathbf{z}, \Delta \mathbf{d}) = \Delta y_{it}(\Delta \tilde{\mathbf{z}}, \Delta \mathbf{d})$ for all $\Delta \mathbf{z}, \Delta \tilde{\mathbf{z}}$ .

Assumption 3 implies that  $\Delta y_{it}(\Delta \mathbf{z}_{W_0}, \Delta \mathbf{d}_{W_0}) = \Delta y_{it}(\Delta d_{it})$  and  $\Delta d_{it}(\Delta \mathbf{z}_{W_0}) = \Delta d_{it}(\Delta z_{it})$ . Furthermore, inference and estimation can now proceed as in Cattaneo et al. (2014), but using the cross-section of first differences instead.

## 1.1 Implementation.

The first step towards point estimation of the local causal effect is the definition of the randomization mechanism determining allocation of  $\Delta Z_{it}$  within  $W_0$ . Let  $N_o$  be the number of observations falling within  $W_0$  and  $M_o$  the number terms with  $\Delta z_{it} = 1$ . As in Rosenbaum and Imbens (2005) and Cattaneo et al. (2014) we assume a fixed margins randomization<sup>1</sup> where  $\mathbb{P}(\Delta \mathbf{Z}_{W_0} = \Delta \mathbf{z}_{W_0}) = \binom{N_o}{M_o}^{-1}$ , for  $\Delta \mathbf{z}_{W_0} \in \Omega$ , the set of all possible permutations of the elements of  $\Delta \mathbf{z}$  in  $W_0$ .

Having indentified a suitable randomization mechanism, we can next test the sharp null hypothesis of no treatment effects. Under this hypothesis,  $\Delta y_{it}(\Delta d_{it}) = \Delta y_{it}$ , and for any suitable test statistic  $T(\Delta \mathbf{Z}_{W_0}, \Delta \mathbf{y}_{W_0})$  with observed value  $\tilde{T}$ , a two sided significance level is

$$\hat{\alpha} = 2 * \min(\mathbb{P}(T(\Delta \mathbf{Z}_{W_0}, \Delta \mathbf{y}_{W_0}) > \tilde{T}); T(\Delta \mathbf{Z}_{W_0}, \Delta \mathbf{y}_{W_0}) < \tilde{T}) \quad (1.3)$$

where

$$\mathbb{P}(T(\Delta \mathbf{Z}_{W_0}, \Delta \mathbf{y}_{W_0}) > \tilde{T}) = \sum_{\Delta \mathbf{z} \in \Omega} \mathbb{I}(T(\Delta \mathbf{Z}_{W_0}, \Delta \mathbf{y}_{W_0}) > \tilde{T}) \mathbb{P}(\Delta \mathbf{Z}_{W_0} = \Delta \mathbf{z}_{W_0}) \quad (1.4)$$

---

<sup>1</sup>The randomization mechanism considered in this article does not allow for clustering within individuals -as would seem necessary for very long panels. Clustering is easily incorporated in the analysis as described in Rosenbaum and Imbens (2005).

and  $\mathbb{P}(\Delta \mathbf{Z}_{W_0} = \Delta \mathbf{z}_{W_0})$  is known, since the randomization mechanism is also known.

In practice, the set of permutations  $\Omega$  can be very large, in which case we can approximate  $\mathbb{P}(T(\Delta \mathbf{Z}_{W_0}, \Delta \mathbf{Y}_{W_0}) > \tilde{T})$  by simulation. Letting  $\{\tilde{T}^*\}_{b=1}^B$  be the sequence of tests obtained by randomly drawing elements from the set  $\Omega$ , then  $\mathbb{P}(T(\Delta \mathbf{Z}_{W_0}, \Delta \mathbf{Y}_{W_0}) > \tilde{T}) \approx B^{-1} \sum_b \mathbb{I}(T^* > \tilde{T})$ .

Note that testing the null hypothesis of no treatment effect does not require Assumptions 3.2 and 3.3. or statements about how the treatment effect interacts with potential outcomes. However, point estimation requires further structural assumptions. As in most of the programme evaluation literature, here we assume a constant, linear treatment effect. Thus, we replace Assumption 2 with the following.

**Assumption 4** *Let  $\theta_i$  and  $\eta_i$  be i.i.d random variables. Then, for all  $i, t$  with  $R_{it} \in W_0$  and any particular value of  $\theta_i, \eta_i$ , the potential treatment,  $d_{it}(z_{it}) + \eta_i$  and potential outcome  $y_{it}(z_{it}, d_{it}) + d_{it}\tau + \theta_i$  satisfy Assumption 1 (for  $\tau \in \mathbb{R}$ ).*

Under Assumptions 2-4 and the null hypothesis  $H_o : \tau = \tau_0$ , the adjusted first differences  $\Delta Y_{it} - \tau_0 \Delta D_{it} = \Delta y_{it}(0)$  are fixed for any value of  $\Delta \mathbf{Z}_{W_0}$ . Therefore, to test the null hypothesis, we define the statistic  $T(\Delta \mathbf{Z}_{W_0}, \Delta \mathbf{Y}_{W_0} - \tau_0 \Delta \mathbf{D}_{W_0})$  and proceed as when testing the sharp null  $H_0 : \tau = 0$ . Once the test is computed, this can be inverted, to obtain the Hodges-Lehmann estimate of  $\tau$  (Hodges and Lehmann, 1963). Let  $\bar{t} = E(T(\Delta \mathbf{Z}_{W_0}, \Delta \mathbf{Y}_{W_0} - \tau_0 \Delta \mathbf{D}_{W_0}))$  be the expected value of the selected test statistics. Then  $\hat{\tau}$  is the value of  $\tau$  such that  $T(\Delta \mathbf{z}_{W_0}, \Delta \mathbf{Y}_{W_0} - \tau_0 \Delta \mathbf{D}_{W_0})$  is as close as possible to  $\bar{t}$ . This amounts to finding the value of  $\tau$  that maximises the p-value of the test. In practice we use the Nelder-Mead simplex algorithm, (Nelder and Mead, 1965; Press et al., 1992).

In principle,  $T(\cdot)$  can be any statistic. Typical choices are differences in mean, tri-mean and the Wilcoxon test. In this article, inspired by the literature on pivot statistics in bootstrap methods (see, for instance, Hall, 1992 or Davidson and Hinkley, 1997), we define  $T$  as the t-ratio associated with  $\Delta Z_{it}$  in the regression model  $\Delta Y_{it} - \tau_0 \Delta D_{it} = c + \beta \Delta Z_{it}$ .

## 1.2 Bandwidth Selection.

In this article we follow the iterative procedure described by Cattaneo et al. (2014) for their cross-sectional framework (for a discussion of bandwidth selection in the standard Regression

Discontinuity Design, see, Imbens and Kalyanaraman, 2012; Calonico et al., 2014; Imbens and Lemieux, 2008; Lee and Lemieux, 2010). The method consist on undertaking randomization inference on a collection of (probably time invariant) pre-determined variables, using bandwidths of decreasing magnitude. The optimal bandwidth is the largest interval within which the Independence-SUTVA-Exclusion trinity in Assumption 1 holds. As explained by Cattaneo et al. (2014) this translates in practice in finding the largest interval within which the causal effect of treatment is statistically insignificant for all the pre-determined characteristics.

### 1.3 Monte Carlo study of the RI-RDD Method.

To facilitate the interpretation of our results, we conducted a limited Monte Carlo experiment to evaluate the capabilities of the inferential framework just discussed. Our data generating process (DGP) was designed to replicate the main characteristics of the data in our empirical application. The running variable was constructed by drawing from a uniform random variable such that  $R_{i0} \sim Uniform$  on  $[-15, 15]$  and, for  $t = 0, 1, 2$ ,

$$D_{it}^* = 0.5R_{it} + 1.4Z_{it} + \mu_i + \varepsilon_{it} \quad (1.5)$$

$$D_{it} = \mathbb{I}(D_{it}^* > 0) \quad (1.6)$$

$$Y_{it}(0) = 0.5Y_{it-1}(0) + \nu_i + v_{it} \quad (1.7)$$

$$Y_{it} = Y_{it}(0) + \tau D_{it} \quad (1.8)$$

where  $\nu_i = \mu_i = \mathcal{N}(0, 1)$ ,  $\varepsilon \sim \mathcal{N}(0, 1)$ ,  $v_{it} \sim (\rho\varepsilon_{it} + \sqrt{1 - \rho^2}\mathcal{N}(0, 1))$  and  $\rho = -0.3$ . Data were drawn from this DGP  $J = 2000$  times, and the number of bootstrap simulations was set at  $b = 599$ . For each part of the study, we computed the average  $\hat{\tau}$ , the Mean Square Error (MSE),  $MSE = J^{-1} \sum_j (\hat{\tau} - \tau_0)^2$  (where  $\tau_0$  was the true parameter value) the Mean Absolute Error,  $MAE = J^{-1} \sum_j |\hat{\tau} - \tau_0|$  and the proportion of rejections of the sharp null hypothesis. We considered samples sizes of 5,000, 10,000 and 20,000 observations and bandwidths of  $\pm 2$ ,  $\pm 4$  and  $\pm 6$ .

The results of the simulation appear in table 1. Randomization Inference produces accurate estimates of  $\tau_0$ , with a very small bias, and the observed MSE and MAE are relatively insensitive to the magnitude of  $\tau_0$ . As can be seen in the final column of the table, the simulation confirms

the *exact* nature of the Randomization Inference, with the % rejections for  $\tau_0 = 0$  departing only slightly from 5%. Nonetheless, unlike MSE and MAE, size seems to be slightly sensitive to the bandwidth selected, specially for the smallest sample size (where size varied between 0.043 and 0.057).

## 2 Construction of a Summary Index.

The construction of Single Index is straightforward. Here we provide an outline of the technique. For excellent discussions on the topic see O'Brien (1984) and Anderson (2008). Consider  $j = 1, \dots, K$  outcomes in a given set of variables (e.g. our 12 mental health measures or our 12 physical health outcomes). Let  $S_{ij}$  be the  $i^{th}$  individual's response for outcome  $j$  in a given domain. Let  $Z_{ij} = S_{ij} - \bar{S}_j / \sigma_j^{(0)}$  be the control-standardized outcome, where  $\sigma_j^{(0)}$  is the standard deviation of  $S_j$  for individuals in full employment, and define vectors  $\mathbf{Z}_j = (Z_{1j}, \dots, Z_{Nj})'$ ,  $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{iJ})'$ . The single mental health index, or class-adjusted score, for individual  $i$  is

$$Y_i = (\iota' \hat{\Sigma}^{-1} \iota)^{-1} \iota' \hat{\Sigma}^{-1} \mathbf{Z}_i \quad (2.9)$$

where  $\hat{\Sigma}$  is the correlation matrix of the standardised scores,

$$\hat{\Sigma} = \begin{pmatrix} \mathbf{Z}'_1 \\ \vdots \\ \mathbf{Z}'_J \end{pmatrix} (\mathbf{Z}_1, \dots, \mathbf{Z}_J) \quad (2.10)$$

As noted by Anderson (2008), summary indices have a number of advantages. Trivially, they reduces the number of tests being carried out. Secondly, the aggregation of outcomes can result in more powerful inference by accumulating several marginal effects. This is important in an analysis of the impact of retirement, given that potentially health status is likely to change only a little, in which case estimated treatment effects of individual health outcomes will approach only marginal significance. Finally, although individual indicators are of some interest, we seek the effect of retirement on overall health, which is only partially revealed by individual measures.



### 3 Two period model of retirement.

In this section we consider a simple two-period version of the model presented in Section 5 of the main text. The objective of this section is to introduce a more general utility function of the individual who has to choose how to optimally allocate time. The model shows that, even in the simplest dynamic setting, results strongly hinge on the characteristics of the individual's utility function, and in particular on the existence of possible forms of complementarity or substitutability between different types of activities. The analysis in this section reinforces that message that (i) understanding the effects of retirement on human capital requires a good understanding of employment and individuals' characteristics; (ii) it is not possible *a priori* to identify how individuals will react to an increase in available time after retirement; (iii) in order to assess the effects of reforms, policy makers should invest particular effort in categorising occupations and clustering individuals' characteristics.

In the two period model, the individual works at time  $t = 0$  and at time  $t = 1$  she retires. At any  $t$  the individual can allocate her total endowment of time,  $\bar{T}$ , between household activities,  $T_t^h$ , leisure,  $T_t^a$  and (for  $t = 0$  only) work  $T_t^w$ . We assume, without much loss of generality, that

$$\bar{T} = T_0^w + T_0^h + T_0^a \quad (3.11)$$

$$\bar{T} = T_1^h + T_1^a \quad (3.12)$$

We will use  $H_t$  to denote a generic measure of stock of human capital. This could be an aggregate measure of the overall stock or an indicator of a single dimension of human capital (such as health or cognitive ability). An individual's stock of capital is the result of three components,

$$H_1 = \delta H_0 + \theta T_0^a + \gamma T_0^w \quad (3.13)$$

Here  $H_0 \in \mathbb{R}_+$  is the initial endowment of cognitive ability which depreciates at rate  $\delta \in [0, 1]$ . Crucially, leisure time and time at work both contribute to a person's human capital, in proportions  $\gamma \in \mathbb{R}$  and  $\theta \in \mathbb{R}$  respectively.

Working in period 0 the individual can earn a salary and invest it in assets. The amount of

assets accumulated at  $t = 1$  are

$$A_1 = \rho A_0 + w(H_0) T_0^w \quad (3.14)$$

where the initial endowment of assets,  $A_0 \in \mathbb{R}_+$  depreciates at rate  $\rho \in [0, 1]$ . The salary that the individual can earn per unit of time depends on her initial endowment of human capital. Specifically  $w'_{H_0} > 0$ ,  $w''_{H_0} \leq 0$ .

The individual chooses the amount of work, leisure and housework so as to maximise the net present value of utility over the life-time, subject to constraints (3.11) to (3.14). That is,

$$\begin{aligned} \max_{T_0^h, T_0^a, T_1^a} \left\{ \right. & U_0(T_0^h, T_0^a, H_0) \\ & + \beta U_1(\bar{T} - T_1^a, T_1^a, (\delta H_0 + \theta T_0^a + \gamma T_0^w)) \\ & \left. + \beta(\rho A_0 + w(H_0) T_0^w) \right\} \end{aligned} \quad (3.15)$$

Suppose that the objective function shows the following reasonable properties.

1.  $U'_{t, T_t^a} \geq 0$ ,  $U'_{t, T_t^h} \geq 0$ ,  $U'_{t, H_t} \geq 0$
2.  $U''_{t, T_t^a} \leq 0$ ,  $U''_{t, T_t^h} \leq 0$ ,  $U''_{t, H_t} \leq 0$
3.  $U''_{t, T_t^a T_t^h} \leq 0$ ,  $U''_{t, T_t^a H_t} \leq 0$ ,  $U''_{t, T_t^h H_t} \leq 0$
4. The Hessian matrix of  $U_t$  is negative semidefinite.

Assumption 1 states that the individual's utility is (weakly) increasing in the time spent in leisure and household activities. Utility may also be increasing in the level of human capital. Assumption 2 states that utility is increasing at a decreasing rate. Assumption 3 allows the possibility that the various dimensions of utility may have forms of complementarity or substitutability. Finally Assumption 4 ensures concavity of the utility function. The first order conditions, FOCs, for the maximisation of (3.15) are:

$$U'_{0, T_0^a} + \beta(\theta - \gamma) U'_{1, H_1} - \beta w = 0$$

$$U'_{0, T_0^h} - \beta \gamma U'_{1, H_1} - \beta w = 0$$

$$\beta \left( U'_{1,T_1^a} - U'_{1,T_1^h} \right) = 0$$

Totally differentiating the FOCs produces:

$$\begin{aligned} & dT_0^a \left[ U''_{0,T_0^a} \right] + dT_0^h \left[ U''_{0,T_0^a T_0^h} \right] + dT_1^a \left[ \beta (\theta - \gamma) U''_{1,H_1 T_1^a} \right] + \\ & + d\theta \left[ \beta U'_{1,H_1} \right] + d\gamma \left[ -\beta U'_{1,H_1} \right] + dH_0 \left[ U''_{0,T_0^a H_0} - \beta w'_{H_0} \right] = 0 \\ & dT_0^a \left[ U''_{0,T_0^a T_0^h} \right] + dT_0^h \left[ U''_{0,T_0^h} \right] + dT_1^a \left[ -\beta \gamma U''_{1,H_1 T_1^a} \right] + \\ & + d\theta [0] + d\gamma \left[ -\beta U'_{1,H_1} \right] + dH_0 \left[ U''_{0,T_0^a H_0} - \beta w'_{H_0} \right] = 0 \end{aligned}$$

$$dT_1^a \left[ \beta \left( U''_{1,T_1^a} - U''_{1,T_1^a T_1^h} \right) \right] = 0$$

Define

$$\Gamma \equiv \begin{bmatrix} U''_{0,T_0^a} & U''_{0,T_0^a T_0^h} & \beta (\theta - \gamma) U''_{1,H_1 T_1^a} \\ U''_{0,T_0^a T_0^h} & U''_{0,T_0^h} & -\beta \gamma U''_{1,H_1 T_1^a} \\ 0 & 0 & \beta \left( U''_{1,T_1^a} - U''_{1,T_1^a T_1^h} \right) \end{bmatrix}$$

Notice that the determinant of  $\Gamma$  is non-negative since the Hessian of  $U_0$  is assumed to be negative semidefinite.

Let us first observe the results of simple comparative statics analysis. Using the Cremer rule for the solution of systems of simultaneous equations when the coefficient matrix is quadratic, comparative statics analysis shows the following.

$$\begin{aligned} \frac{dT_0^a}{d\theta} &= \frac{-\beta^2 U'_{1,H_1} U''_{0,T_0^h} \left( U''_{1,T_1^a} - U''_{1,T_1^a T_1^h} \right)}{|\Gamma|} \\ \frac{dT_0^h}{d\theta} &= \frac{\beta^2 U'_{1,H_1} U''_{0,T_0^a T_0^h} \left( U''_{1,T_1^a} - U''_{1,T_1^a T_1^h} \right)}{|\Gamma|} \\ \frac{dT_0^a}{d\gamma} &= \frac{\beta^2 U'_{1,H_1} \left( U''_{0,T_0^h} - U''_{0,T_0^a T_0^h} \right) \left( U''_{1,T_1^a} - U''_{1,T_1^a T_1^h} \right)}{|\Gamma|} \\ \frac{dT_0^h}{d\gamma} &= \frac{\beta^2 U'_{1,H_1} \left( U''_{0,T_0^a} - U''_{0,T_0^a T_0^h} \right) \left( U''_{1,T_1^a} - U''_{1,T_1^a T_1^h} \right)}{|\Gamma|} \\ \frac{dT_0^a}{dH_0} &= \frac{\beta \left( U''_{1,T_1^a} - U''_{1,T_1^a T_1^h} \right) \left( U''_{0,T_0^a H_0} - \beta w'_{H_0} \right) \left( U''_{0,T_0^a T_0^h} - U''_{0,T_0^h} \right)}{|\Gamma|} \\ \frac{dT_0^h}{dH_0} &= \frac{\beta \left( U''_{1,T_1^a} - U''_{1,T_1^a T_1^h} \right) \left[ -U''_{0,T_0^a} \left( U''_{0,T_0^h H_0} - \beta w'_{H_0} \right) + U''_{0,T_0^a T_0^h} \left( U''_{0,T_0^a H_0} - \beta w'_{H_0} \right) \right]}{|\Gamma|} \end{aligned}$$

Given the very simplified structure of the model, the time allocation after retirement is only defined by the satisfaction of the condition

$$U'_{1,T_1^a} = U'_{1,T_1^h} \text{ and this implies that } \frac{dT_1^a}{d\theta} = \frac{dT_1^a}{d\gamma} = \frac{dT_1^a}{dH_0} = 0$$

Signing the expressions regarding the optimal time allocation at time 0 is not trivial. The way  $\theta$ ,  $\gamma$  and  $H_0$  affect the time allocation is ambiguous and crucially depends on the particular functional form that the utility may take (and in particular of possible forms of complementarity or substitutability between activities in each period).

Even in its very simplified form, the model produces interesting insights. Human capital crucially depends on the way the individual allocates her time. When working, time allocation depends on individual characteristics (i.e. the particular functional form that the utility can take), job characteristics (i.e.  $\gamma$ ) and socio-demographic characteristics (i.e.  $\theta$ ) that in turn define the way leisure choices may affect human capital. After retirement, the individual will have spare time. Whether she will decide to invest her spare time in increasing household activities or leisure (or both) will depend only on the features of the utility function<sup>2</sup>. It follows that the specific effect of retirement (i.e. how each individual spend her time and how this allocation compares to the decision before retirement) will depend on the particular characteristics of the individual (i.e. the utility function), her job and socio-demographic characteristics.

---

<sup>2</sup>The model can be extended to include an additional period after retirement, say time 2 where the individual obtains utility from her human capital. The qualitative results and the message of the model would unchanged. The Appendix provides an extension of the model with  $T > 2$  periods and a linear-quadratic utility function.

**Table 1:** Monte Carlo Simulation. DGP:  $Y_{it}(0) = 0.5Y_{it-1}(0) + \nu_i + v_{it}$ ;  $Y_{it} = Y_{it}(0)$ ;  $D_{it} = \mathbb{I}(0.5R_{it} + 1.4Z_{it} + \mu_i + \varepsilon_{it})$ . Based on 2000 replications.  $Z \sim \text{Uniform}[-15, 15]$ . The number of bootstraps was set to 599. Here  $N$  represents the number of individuals drawn from the DGP. The actual number of individuals used in each estimation depended on the bandwidth.

N	Bandwidth	ATE	MSE	MAE	% Rejections
$T = 2$					
5000	2	0.005	0.353	0.279	0.057
5000	4	0.002	0.297	0.235	0.050
5000	6	0.001	0.266	0.211	0.043
10000	2	0.007	0.236	0.190	0.044
10000	4	0.000	0.205	0.164	0.052
10000	6	0.001	0.185	0.148	0.047
20000	2	0.001	0.170	0.134	0.047
20000	4	0.000	0.146	0.119	0.049
20000	6	0.006	0.132	0.105	0.046
$T = 4$					
5000	2	0.008	0.217	0.172	0.060
5000	4	0.015	0.183	0.145	0.065
5000	6	-0.008	0.161	0.129	0.057
10000	2	0.008	0.152	0.122	0.066
10000	4	0.002	0.126	0.100	0.054
10000	6	0.004	0.115	0.092	0.058
20000	2	0.004	0.107	0.086	0.056
20000	4	0.002	0.090	0.072	0.059
20000	6	-0.001	0.080	0.064	0.046

## References

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

- 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*.
- Davidson, A. and D. V. Hinkley (1997). *Bootstrap Methods and their Applications*. Cambridge University Press.
- Hall, P. (1992). *The bootstrap and Edegeworth expansion*. Springer.
- Hodges, J. and E. Lehmann (1963). Estimates of location based on ranks. *Annals of Mathematical Statistics* 34, 598–611.
- Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the Regression Discontinuity estimator. *Review of Economic Studies*, 1–27.
- Imbens, G. and T. Lemieux (2008). Regression Discontinuity Designs: A guide to practice. *Journal of Econometrics* 142, 615–635.
- Lee, D. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature* 48, 281–355.
- Nelder, J. and R. Mead (1965). A simple method for function minimization. *Computer Journal* 7, 308–313.
- O’Brien, P. (1984). Procedures for comparing samples with multiple endpoints. *Biometrics* 40, 1079–1087.
- Press, W., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992). *Numerical Recipes in C (The art of Scientific Computing), Second Edition*. Cambridge University Press.
- 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.