Short and long run estimates of the local effects of retirement on health.

Fe, Eduardo and Hollingsworth, Bruce

2015

Short and long run estimates of the local effects of retirement on health.

Eduardo Fé* † [1] and Bruce Hollingsworth‡ [2]

[1] *Blavatnik School of Government, University of Oxford, U.K.*
[2] *Division of Health Research, University of Lancaster, U.K.*

May 15, 2015

Abstract

We explore the existence of short and long term effects of retirement on health. Short term effects are estimated with a regression discontinuity design which is robust to weak instruments and where the underlying assumptions of continuity of potential outcomes are uncontroversial. To identify the long term effects we propose a parametric model which, under strong assumptions, can separate normal deterioration of health from the causal effects of retirement. We apply our framework to the British Household Panel Survey, and find that retirement has little effect on health. However, our estimates suggest that retirement opens the gate to a sedentary life with an impoverished social component and this is a channel through which retirement could indirectly affect health in the long run.

Key Words: *Regression discontinuity, retirement, instrumental variables, health, wild bootstrap.*

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

---

*The authors thank the editor, Linda Sharples, and two anonymous referees for very helpful comments. We also thank the many helpful comments received from Edward Vytlacil, Matias Cattaneo, Rocio Titiunik, Alastair Gray, Thomas Mason, Timothy Powell-Jackson and seminar participants at the Centre for Health Economics at Monash University and the joint HESG-CES at Aix-en-Provence.

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

‡Division of Health Research, Furness College, University of Lancaster, Lancaster LA1 4YT. email: b.hollingsworth@lancaster.ac.uk

1 Introduction.

Over the last few decades OECD countries have witnessed a trend towards early retirement (Krueger and Pischke, 1992; Blundell et al., 2002; Gruber and Wise, 2002, 2004; Banks and Smith, 2006). This has resulted in a rise in the ratio of retirees per person of working age and, consequently, a reduction in savings and additional pressure on the welfare system (an effect exacerbated by a rise in the age at which people leave school; Banks and Smith, 2006). In response, policy makers in developed countries are now debating or implementing changes to the eligibility criteria for retirement, in particular extending the age at which people are eligible for state pensions. However, it has been argued that retirement might affect health, in which case, the impact of retirement on utilisation of health services should be taken into account in the design of policies that delay retirement.

In this paper we estimate the causal effect of retirement on the health and health care utilisation of U.K. male residents. Unlike previous work, which has focused on health outcomes alone (Charles, 2004; Dhaval et al., 2008; Coe and Lindeboom, 2008; Neuman, 2008; Johnston and Lee, 2009; Sickles and Taubman, 1986; Cai, 2010), we also explore whether retirement affects individuals' distribution of time among activities. Any such changes could affect health indirectly depending on the extent to which retirees spend their time in health promoting activities.

Estimation of the causal effect of retirement on health is complicated by the fact that health is a determinant of retirement itself (Bound, 1991; Stern, 1989; Disney et al., 2006). In addition to this source of confounding, retirement happens at a relatively older age when health might have begun to deteriorate. Thus it is difficult to separate trends initiated after retirement from those intrinsic to normal deterioration of health (Stern, 1989; Bound, 1991; Bound and Waidmann, 2007; Bound et al., 2010; Sickles and Taubman, 1986; Cai, 2010). Nonetheless, the institutional setting of retirement in the

U.K. facilitates identification. During the period under consideration (1991-2005), there were not mandatory retirement laws, but men reaching the age of 65 could access their state pension which, in 2013, amounted to £110 (approximately a quarter of the weekly U.K. minimum wage). Because of this incentive, the proportion of retired individuals by age group in any given year exhibited a unique jump of about 20 percentage points at age 65 (the State Pension Age, SPA). If this discontinuity was accompanied by similar significant variation in health outcomes then, under certain assumptions, we can attribute variations in health to retirement. Therefore we estimate, under different sets of assumptions, if there are significant variations in average health status around age 65.

Our analysis begins with an implementation of a Regression Discontinuity Design (Thistlethwaite and Campbell, 1960; Hahn, Todd, and van der Klaauw, 2001; Porter, 2003). The Regression Discontinuity Design (RDD) provides estimates of changes in health upon retirement using data from the years around the SPA. These estimates have a causal interpretation under minimal assumptions but, because of the focus on what happens to health just around the SPA, they can only be interpreted as short-term effects of retirement. Short-term effects are important because the environmental changes that follow retirement are dramatic and often irreversible, therefore they are indicative of future changes in health (for instance, through retirement one loses daily interaction with colleagues at work, which could affect the amount of time spent socialising, and this to one's sense of belonging). However, retirement might induce only smooth changes on health. In this case, one would expect to observe shifts in trends around retirement, rather than discontinuities, and a RDD would be ineffective. To estimate these longer-term effects, we extend recent work by Bound and Waidmann (2007) and estimate a panel data model which directly assess whether reaching the SPA leads to any changes in the slope or level of health outcomes. This second approach makes use of the whole sample available to us -rather than that coming from a neighbourhood around the SPA. As a result, it provides more efficient estimates than the RDD, but a causal interpretation of these estimates requires considerably stronger assumptions.

Our estimation strategy introduces several methodological innovations. In particular, unlike previous work our RDD takes advantage of the longitudinal structure of the data (in order to reduce the strength of the assumptions required for identification) and incorporates recent results by Dong (2013) in order to increases the robustness of our results to the problem of weak instrumental variables -which, in our RDD, translates in a small discontinuity in the distribution of retirements conditional on age. Furthermore, we propose a cluster-robust version of the wild bootstrap schemes in Davidson and MacKinnon (2010) to correct the size of inferential procedures and address several challenges presented by the nature of the data (and, in particular, *the discrete running variable problem* in RDD - discussed in Lee and Card, 2008 and Dong, 2014).

The methods for identification, estimation and inference are summarized in Section 3, after introducing the data and the institutional context of retirement in the U.K. in Section 2 (with much of the technical details deferred to an Appendix). Section 4 of the paper contains the empirical analysis. Overall we find that retirement has little effect on health. However, our estimates suggest that retirement opens the gate to a sedentary life with an impoverished social component and this is a channel through which retirement could indirectly affect health in the long run. We conclude in Section 5 with some remarks on the implications of our results.

2 Data and Institutional Context.

2.1 Data.

Our sample consists of 11,331 male individuals in the first 15 waves (1991-2005) of the BHPS -a total of 62,404 observations. In our analysis, an individual in categorised as being retired if that was his self-reported job market status and he reported not to have undertaken any paid work during the two weeks prior to the interview. Our sample excluded all those individuals who re-enter the job market after retiring or who moved to/from unemployment from/to retirement. With these criteria we could identify 394 retirements.

We study the effects of retirement on a wide range of health indicators. In the BHPS, mental health is approximated with Goldberg's General Health Questionnaire, GHQ-12, (Goldberg, 1978), a twelve domain instrument that measures the inability to carry out normal functions and the appearance of new and distressing psychological phenomena. Respondents score each domain from 1 to 4, with 1 representing the best possible outcome. In addition to the GHQ, interviewees report whether they have any of 12 common health problems and whether, in the previous twelve months, they underwent any of the medical tests in a list.

BHPS has just one variable measuring time use outside work (self-reported weekly hours spent in housework). To get a richer view of the effects of retirement on time use we use an ancillary dataset, the British Household Panel Survey Calibrated Time Use Data, (BHPS-CTU; Kan and Gershuny, 2006). The BHPS-CTU is constructed with time use diary data from the Home On-line (HoL) study, a national representative longitudinal survey containing above 16,000 diary days from 1,000 households which were followed between 1999 and 2001. Exploiting the high correlations between the data in the BHPS and HoL (between 0.74 and 0.94) Kan and Gershuny (2006) use diary data from HoL to generate calibrated estimates of time use for each individual in the BHPS (calibration ensures that the total number of minutes adds to 1,440).

The BHPS-CTU contributes seven variables to our analysis: three primary variables (*routine housework*, including cooking, cleaning, ironing and washing; *sleep, personal care and rest*; *consumption and leisure*) and four derived variables: consumption and leisure time is divided into *social leisure* (spent with other people), *passive leisure* (spent alone), *leisure at home* and *leisure outside home*. Further information regarding the construction of these variables is given in Kan and Gershuny (2006).

2.2 Institutional Context.

The choice of sample and our identification strategy is motivated by the institutional context of retirement in the U.K. In the UK, an individual's income after retirement is determined by the history of contributions to state and private pension schemes. Although occupational and personal pension funds are common among workers, the state pension still plays a crucial role in the system (Pensions Policy Institute, 2013). An individual with a full history of 30 years of National Insurance contributions received, in 2013, a maximum Basic State Pension (BSP) of £110.15 a week. Considering that, according to the UK's Office for National Statistics, in 2013, the average income in the UK was approximately £26,500 (a weekly income, after tax, of about £400), we see that the BSP constitutes an important source of income. Furthermore, besides the BSP there is an 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 (1991-2005), was 65 years for men.

The eligibility criteria present a strong incentive to stay in 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 age 65. This gap can be seen in Figure 1, which describe the distribution of retirement among men in the BHPS. Because during the period under consideration, there were not mandatory retirement laws in the U.K. which could explain this discontinuity (see Appendix A) 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 age 65 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 to identify the causal effects of retirement. To fully quality as an instrument, this indicator variable must be uncorrelated with health. Yet, although SPA is exogenously determined by government, this indicator variable does not qualify as a valid instrument as such. Firstly, one may worry that governments' choice of 65-70 as an appropriate state pension age is motivated by people's health starting to deteriorate more visibly around that age. This does not seem likely. This age range has been the norm since the introduction of Otto von Bismark's state pension in the XIX century -and the motivations behind that age range are unclear, especially if one considers that life expectancy in XIX century Germany was just above 40 years of age (Roser, 2015). More importantly, however, the SPA splits the population into a young and old groups, both of which exhibit dramatically different health and time use profiles. This induces a correlation between the instrument and health and it is because of this reason that further restrictions are required to ensure identification, as we discuss next.

3 Identifying the Effect of Retirement.

Let Y_{is} denote individual i's observed level of an specific health indicator at time s, for $i = 1, \ldots, N$, $s = 1, \ldots, S$. In our empirical application, Y_{is} will represent a time use indicator or summary indices of health (whose construction is detailed in Appendix C). Let $(Y_{1,is}, Y_{0,is}) \in \mathbb{R}^2$ be an individual's potential outcomes (Rubin, 1974) such that $Y_{t,is}$, $t \in \{0, 1\}$ denotes the level of health that individual i would exhibit at time s were he retired ($t = 1$) or employed ($t = 0$). The observed employment status, T_{is} depends, among other factors, on having reached the SPA -which will be indicated by the binary variable D_{is}, a function of age in years, X_{is}, so that $D_{is} = 1$ if $X_{is} \geq x_o$, and 0 otherwise, where the policy threshold is $x_o = 65$. Then, $T_{is} = T_{1,is}D_{is} + (1 - D_{is})T_{0,is}$, where $T_{d,is}$ denotes the employment status that individual i would choose at time s were he eligible ($d = 1$) or not ($d = 0$) for the state pension. We estimate short and long term effects via methods akin to Instrumental Variable (IV) estimation of longitudinal regression models. The key instrument is an indicator of eligibility for SPA. However, because crossing the SPA predicts less than 20% of retirement decisions, the parameter we are estimating is the local treatment effects on compliers (Angrist, Imbens, and Rubin, 1996), that is, the effect of retirement on those individuals whose retirement status changes because eligibility for SPA changes -and thus $T_{1,is} - T_{0,is} = 1$).

3.1 Short Term Effects.

It is possible to identify the short term effects of retirement on health with minimal assumptions by focusing the analysis on the years around the SPA. As revealed by Figure 1, the SPA can explain many retirement decisions in this short space of time, unlike health which, in the absence of any other significant event, should vary little, beyond common trends in the deterioration of health. As a result, once these trends are taken into consideration, any changes in average health can be attributed to those individuals whose retirement status changed because they reached the SPA. Yet it is unclear a priori whether the 20% discontinuity in the distribution of retirees is sufficient to guarantee a strong predictor of retirement (that is, a strong instrumental variable). If not, estimates of the causal effect will be biased and inference will be unreliable (Bound et al., 1995; Staiger and Stock, 1997; Kleibergen, 2002; Davidson and MacKinnon, 2010; Feir et al., 2014). The dynamics of retirement in the U.K. present way of preventing a potential weak instrument problem.

As seen in Figure 1, by age 65 around 60% of men have retired. By age 66, 80% of the individuals in the sample are retired and, beyond that point, the proportion of retirees increases very slowly. As a result there is a dramatic change in the slope (a kink) at the SPA. A fixed effects regression of the retirement indicator on a fourth order polynomial on age, an indicator of SPA-eligibility, its interaction with age and a collection of time dummies yields an estimated coefficient of the interaction term of

-.0384 ($p = 0.000$) -and is largely insensitive to the specification and interval around the SPA considered. As noted in a recent paper by Dong (2013) (see also Card et al., 2012, Calonico et al., 2014), kinks and jumps in the propensity score can be combined a Regression Discontinuity fashion to identify variations in the conditional mean. This expands the set of instrumental variables and prevents biases if the discontinuity in the proportion of retirees is not sufficiently dramatic.

Dong (2013) shows that, if there exists a king or jump at x_o in the propensity score, and potential outcomes and their derivatives are continuous, then

$$\tau = \frac{\lim_{x \to x_o^+} G(x) - \lim_{x \to x_o^-} G(x) + w_n(\lim_{x \to x_o^+} G'(x) - \lim_{x \to x_o^-} G'(x))}{\lim_{x \to x_o^+} P(x) - \lim_{x \to x_o^-} P(x) + w_n(\lim_{x \to x_o^+} P'(x) - \lim_{x \to x_o^-} P'(x))} \tag{3.1}$$

where $\tau = E(Y_1 - Y_0|X = x_o)$ is the local average treatment effect of retirement on compliers, $G(x) = E(Y|X = x)$, $P(x) = E(T|X = x)$, G', P' are first order partial derivatives and $\{w_n\}$ satisfies $\lim_{n\to\infty} w_n = 0$. When $w_n = 0$ we have the RDD parameter in Hahn et al. (2001); otherwise τ is defined by the contribution of kinks and jumps in the mean outcome and propensity score. Nonparametric regressions (Stone, 1977) can be used to estimate each of the above limits. However, if one uses uniform kernels and the same bandwidth throught out, Hahn et al. (2001) and Dong (2013) show that, in a neighbourhood $(x_o - h(n), x_o - h(n))$, estimation of τ amounts to two stage least squares of

$$Y_i = \beta_0 + \beta_1 \tilde{X}_+ \ldots + \beta_p \tilde{X}_i^p + \tau T_i + \varepsilon \tag{3.2}$$

$$T_i = \alpha_o + \alpha_1 \tilde{X}_i + \ldots \alpha_p \tilde{X}_i^p + \pi_1 D_i + \pi_2 D_i * \tilde{X}_i + u. \tag{3.3}$$

where $\tilde{X}_i = X_i - x_o$. In this case, $w_1 = cov(T, D)$, $w_2 = cov(T, X * D)$.

When panel data are available, the RDD literature has pooled the data and averaged the endogenous variables within each wave and across the panel conditional on the running variable (this is referred to as the average cluster approach; Lemieux and Milligan, 2008; Battistin et al., 2009). However,it is possible to exploit the longitudinal structure of the data by estimating the following model instead,

$$
\begin{aligned}
Y_{is} &= \theta_i + \lambda_t + \tau T_{is} + \beta_1 \tilde{X}_{is} + \ldots + \beta_p \tilde{X}_{is}^p + \varepsilon_{is} & (3.4) \\
T_{is} &= \eta_i + \zeta_t + \gamma_1 \tilde{X}_{is} + \ldots + \gamma_p \tilde{X}_{is}^p + \pi_1 D_{is} + \pi_2 D_{is} * \tilde{X}_{is} + \nu_{it}. & (3.5)
\end{aligned}
$$

where θ_i, λ_s, η_i and ζ_s are individual and period specific coefficients. Estimation can be done by transforming the above model into first differences,$(Y_{is} - Y_{i,s-1})$, and applying two stage least squares (Greene, 2004; Wooldridge, 2010). The ensuing estimator is, in essence, the average difference in outcome $(Y_{i,s} - Y_{i,s-1})$ when retirement occurs $(T_{is} - T_{i,s-1} = 1)$ minus the same average when employment/retirement status does not change $(T_{is} - T_{i,s-1} = 0)$. Identification of τ comes from the variation in outcomes experienced by those individuals whose treatment status changes within $\mathcal{W}_o$. Because comparisons are undertaken within individuals right before/right after retirement, comparability of treatment and control groups and continuity of potential outcomes (and their derivatives) are uncontroversial assumptions. Furthermore, we can allow the existence of time invariant unobserved heterogeneity that may affect retirement or other observable traits in unspecified ways, what greatly relaxes the assumptions underlying identification.

To select a bandwidth, $h_n = h(n)$, we follow a Cross-Validation approach similar to those described in Imbens and Lemieux (2008) and Imbens and Kalyanaraman (2012) (see also Calonico et al., 2014; Hart, 1997; Pagan and Ullah, 1999; Li and Racine, 2006). Unlike in standard RDD, here we have the additional complication of having to choose a polynomial order for (3.4) and (3.5) as well as the bandwidth around the SPA. Let the residuals from the parametric model in (3.4) be denoted by e, and let $e_{(i)}$ denoted the i^{th} deleted residual, obtained from estimating (3.4) by IV leaving the i^{th} observation out. Then, our Cross-Validation algorithm selects the parameters (h, p) which minimised the leave-one-out predicted residual sum of squares,

$$PRSS = \sum_{i=1}^{NS} e_{(i)}^2 \tag{3.6}$$

over the combination of (h, p), where $p = 1, \ldots, 5$ and $h = 1, \ldots 10$.

As in Feir et al. (2014) we base inference about τ on the ratio of $\hat{\tau}$ to its standard deviation. To compute p-values we implement the *restricted-efficient* wild bootstrap algorithm in Davidson and MacKinnon (2010) duly adapted to handle correlations at cluster (age) level. The motivation and algorithm underlying this technique are detailed in the Appendix B. Here we just note that this particular method is robust to heteroskedasticity of unknown form and provides accurate inference even in presence of weak instruments. More importantly for our work, it handles the *discrete-running-variable* problem arising from the fact that our assignment variable (age) is a discrete random variable -while standard RDD require continuous assignment variables (see Lee and Card, 2008 and Dong, 2014).

3.2 Retirement in the Long Run.

RDD provides estimates akin to short run causal effects. In our analysis short term effects are important because retirement involves dramatic and permanent environmental changes and, as a result, its short-term consequences are likely to project into the future. However the scope of these effects can be limited. Firstly, short term effects might turn out to be only transitory. Secondly, retirement might affect health only in the long run, in which case one would expect to observe changes in the trends of health around retirement, but not discontinuities. Thus, we complete our analysis estimating the longer term effects of retirement on health.

We follow Bound and Waidmann (2007), and investigate if there are changes in the slope or level of health outcomes at the SPA. Because we are estimating long term effects, we cannot now restrict the analysis to a short interval around the SPA, but this forces us to be explicit about the structure of the correlation between age and health in order to ensure that SPA is a valid exogenous predictor of retirement. To be more precise, we estimate the following panel data extension of the linear model in Bound and Waidmann (2007),

$$y_{is} = \alpha_i + D_{is}\tau + \mathbf{x}_{is}\gamma + \varepsilon_{is} \tag{3.7}$$

where $\mathbf{x}$ includes a fifth polynomial in age and the interaction of D_{is} with age. In (3.7) the coefficient τ can be interpreted as the local causal effect of retirement because the SPA is exogenous determined by government and $\mathbf{x}$ is designed to absorb any correlations between health and age (so that any residual effect of SPA on health must be due to retirement alone). Note, however, that (3.7) is a considerably stronger assumption than those required for identification in our RDD: in RDD we only need to assume that model (3.4) holds only in a neighbourhood of SPA. In particular, (3.4) could be interpreted as a Taylor approximation of the true conditional mean of health outcomes around SPA.

Unlike in Bound and Waidmann (2007), we exploit the longitudinal structure of our data and estimate (3.7) after transforming the model in first differences -and adding a set of time dummies to the first differenced equation. We assess the existence of changes in trends with an F-test for the joint significance of D_{is} and the interaction term and compute p-values via the same cluster-robust wild bootstrap procedure used in the RDD (specialised to the case of no reduced-form equation; see Appendix B.1).

4 Empirical Analysis.

We began our empirical investigation of the causal effects of retirement by undertaking a descriptive analysis of the health and time use outcomes. The full descriptive analysis can be found in the Appendix E, while here we just provide key highlights. As can be inferred from Table 1, retirees exhibited worse patterns of health than non-retirees, but the descriptive analysis suggests that much of this variation is driven by age. In particular, the analysis did not reveal jumps or kinks in health outcomes around the SPA. On the other hand, Table 1 and figure 4) reveal that the additional leisure time available after retirement is primarily invested in passive leisure (which increases by 41%, as opposed to an increase in social leisure of just 1.32%) and leisure at home (which increases by 39%, as opposed to leisure outside home, which decreases by around 1%). Together with an increase of 12% in the amount of time spent sleeping and in personal care, this suggests that the type of leisure that might be associated with a sedentary lifestyle becomes more prevalent with age. Unlike with health outcomes, Figure 4 reveals a clear discontinuity at the SPA for sleep/rest time, leisure and, to some extent, housework what suggests that changes in the distribution of time use are partly driven by retirement.

4.1 Estimates of the Local Causal Effects of Retirement on Health.

We applied the methods in section 3.1 to estimate the short-run causal effects of retirement on health. For our procedures we set the number of Monte Carlo replications to 999, the significance levels at 5% and bandwidth was selected using the Cross-Validation algorithm described in section 3.1. The suitability of the discontinuity design was evaluated by discarding discontinuities in pre-determined variables which might have confounded estimation (see Appendix F). The results are collected in the first two columns of Table 2. They confirm that, as the descriptive analysis had anticipated, in the BHPS retirement does not have an impact on self-reported mental health (p-value, $p = 0.11$), physical health ($p = 0.966$) or health checks ($p = 0.491$).

The methods in 3.2 were subsequently applied to our sample, and the results from the long term analysis are collected in Table 4. We did not find any significant effect of pension eligibility on health (in particular the null hypothesis of no changes in trend around the pension age was accepted for all three health indicators).

4.2 Estimates of the Local Causal Effects of Retirement on time use.

We explored the effect of retirement on the measures of utilisation of time. The estimates of the short term effect of retirement on time use, collected in Table 2, suggest a statistically significant increase in the amount of time retirees devote to leisure activities. Time spent sleeping and in personal care increases around 82 minutes a day, which constitutes a 17% rise with respect to the mean amount of time spent in these activities by the group of 50 to 80 year old employees. Consumption and leisure also receive more attention, increasing by about 140 minutes a day. When we consider the distribution of leisure among activities, the prevalent results are that (1) people invest their time mostly in passive leisure (which increase by about 95 minutes after retirement) as opposed to social leisure (up by around 50 minutes) and (2) people invest more of their spare time in leisure at home (which increases about 100 minutes after retirement) than leisure outside home (which increases by about 46 minutes). Finally, and interestingly, our results seem to suggest that retirement does not induce a significant change in the amount of time devoted to housework.

Aside from their immediate interest, these results also outline the fact that our design is capable of detecting significant departures from the null hypothesis of no-causal effect, thus ruling out the possibility that the lack of significant results in the analysis of health outcomes might be driven by lack of power.

4.3 Robustness check.

We evaluated the sensitivity of our result by varying certain aspects of the estimation process. For the long run analysis, we considered several specifications using various combinations of polynomials in age indicator variables, without obtaining any variation in conclusions.

For the short run analysis, we firstly, recalculated the short-term treatment effects for the intervals $[-4, 3]$ and $[-2, 1]$, but the results did not change substantially. In addition to this, we evaluate the robustness of our estimates by implementing the Randomization Inference framework for RDD in Cattaneo et al. (2014). As detailed in the Appendix D, this method enables researchers to conduct exact inference with very few assumptions (which are a variation of the Exclusion-Independence-SUTVA trinity in Angrist et al., 1996). Furthermore, identification under this scheme is feasible irrespectively of having a continuous or discrete running variable. The results from this analysis are collected in the third and fourth columns of table 2, and confirm that there are not effects of retirement on health outcomes, housework remains largely unaffected and retirees invest more time in passive leisure and leisure at home than active leisure and leisure outside home. There are only unremarkable discrepancies between this second set of estimates and those provided by the asymptotic method in Section 3.

5 Conclusion

The goal of this research was to explore the implications of retirement for health in the short and long run. Using a novel set of techniques, we did not find evidence that retirement has a direct effect on health in the short or long run (in agreement with Coe and Lindeboom, 2008, Neuman, 2008 and Bound and Waidmann, 2007). However we found that, upon retirement, individuals dramatically increase the amount of leisure time spent at home and in activities with a low social component. Unlike in Aguiar and Hurst (2005), we did not find that retirement affects how much housework people do. Together with an increase in the amount of time spent sleeping, our results suggests that retirement can lead to a sedentary, socially isolated life style, and this is a channel through which retirement could have long run effects for health (Cornwell and Waite, 2009; Cohen et al., 1997; Barness et al., 2004; House et al., 1988).

Our estimates rely on a novel Regression Discontinuity Design (which estimates short run effects under minimal assumptions) and a standard panel data model (which estimates long run effects under more rigid assumptions). As in common with all previous work (e.g. Charles, 2004; Dhaval et al., 2008; Coe and Lindeboom, 2008; Neuman, 2008; Mazzonna and Peracchi, 2012) our methods are, in essence, Instrumental Variable methods and thus provide estimates of the local average treatment effect of retirement, that is the effect on those individuals who move into retirement because their eligibility for a state pension changes. Unlike most of the literature, which has based inference on potentially misleading asymptotic normal approximations (Efron, 1979; Hall, 1992; Orme, 1990; Bertrand et al., 2004; Cameron et al., 2008; Davidson and MacKinnon, 2010, etc) - here we computed size-corrected p-values constructing an extension of the wild-bootstrap algorithm in Davidson and MacKinnon (2010) which takes into account correlations at cluster level (Moulton, 1990) as well as the correlations between the structural and reduced from equations in Instrumental Variable methods.

The immediate implication of our results is that recent moves to extend SPA are unlikely to have direct implications for the demand of health services. On the contrary, extending the SPA might turn out to be a mechanism enabling higher levels of health promoting leisure (both social and physical) in the older age. Note, however, that our results are uninformative about the consequences for health of additional active/passive leisure, or the implications of an older workforce for the productivity of an economy. Our results also imply that studies exploring retirement decisions where health appears as a covariate in a regression model are unlikely to be contaminated by endogeneity or reverse causality in health (although other sources of confounding cannot be ruled out -see Bound, 1991).

That retirement has little effect on the health of an overall population does not preclude that certain individuals may experience variations in health upon retirement. Occupation and the level of job satisfaction might be important, albeit ambiguous, predictors of the significance of retirement for health. Individuals unsatisfied with their jobs or in hazardous occupations might experience an improvement in some dimensions of their health when they retire. Yet these same individuals might have accumulated long term health problems which retirement alone could not fix. Alternatively, they might feel the loss of certain dimensions of working life (such as its social dimension). We explored the effects of occupation, job satisfaction and education levels on our results, but found no variation. However, we could not confirm if those results were driven by the lack of effect of retirement or by the loss of statistical power generated by splitting the data into thinner categories - we leave this for future research.

One limitation of this study is that we do not explore the role that private health insurance might have on retirement decisions. This is an important aspect in countries such as the USA where private insurance is pervasive and its availability can affect the profile of retirees as pension age approaches. In countries such as the U.K., however, the existence of universal coverage is likely to mitigate the impact of insurance (indeed, in the BHPS just under 10% of individuals reported having any occupational health insurance in any one wave). Nonetheless, this remains a question of interest that we cannot address here. From a methodological perspective, our approach has emphasised the use of summary indices in situations as in ours where a large number of outcomes were available for study. An alternative approach had been to deal with outcomes directly and to undertake a joint test of hypothesis. Moreover, in common with most empirical work, we only considers average treatment effects. However, the overall distribution of health is of interest for the policy maker because the average health status might remain constant while radical changes might be occurring in the distribution. The implications of these issues are, however, beyond the scope of this paper.

References

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

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

Banks, J. and S. Smith (2006). Retirement in the U.K. *Oxford Review of Economic Policy 22*, 40–56.

Barness, L., C. Mendes de Leon, R. Wilson, J. Bienias, and D. Evans (2004). Social resources and cognotive decline in a population of older African Americans and whites. *Neurology 12*, 2322–2326.

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

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust difference-in-differences estimates? *Quarterly Journal of Economics. 119*, 249–275.

Blundell, R., C. Meghir, and S. Smith (2002). Pension incentives and the pattern of early retirement. *Economic Journal 112*, C153–C170.

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

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

Bound, J., T. Stinebrickner, and T. Waidmann (2010). Health, economic resources and the work decision of older men. *Journal of Econometrics 156*, 106–129.

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.

Cai, L. (2010). The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model. *Labour Economics 17*, 77–90.

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.

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

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

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

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

Cohen, S., W. J. Doyle, D. P. Skoner, B. Rabin, and J. Gwaltney (1997). Social ties and susceptibility to the common cold. *Journal of the American Medical Association 24*, 1940–1944.

Cornwell, E. and L. Waite (2009). Social disconnectedness, perceived isolation and health among older adults. *Journal of Health and Social Behavior 50*, 31–48.

Davidson, R. and J. MacKinnon (2010). Wild bootstrap tests for IV regression. *Journal of Business and Economic Statistics 28*(1), 128–144.

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

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

Efron, B. (1979). Bootstrap methods: another look at the Jacknife. *Annals of Statistics 7*, 1–26.

Feir, D., T. Lemieux, and V. Marmer (2014). Weak identification in Fuzzy Regression Discontinuity desings. Technical report, Micro Theory Working Paper, Microeconomics.ca.

Goldberg, D. (1978). *Manual of the General Health Questionnnaire.* NFER Publishing.

Greene, W. (2004). *Econometric Analysis.* Prentice Hall.

Gruber, J. and D. Wise (2002). *Social Security and retirement around the World.* University of Chicago Press.

Gruber, J. and D. Wise (2004). *Social Security programs and retirement around the World: Micro-estimation.* University of Chicago Press.

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

Hall, P. (1992). *The bootstrap and Edegeworth expansion*. Springer.

Hart, J. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests*. Springer-Verlag New York.

House, J., K. Landis, and D. Umberson (1988). Social relationships and health. *Science 241*, 540–545.

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.

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

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

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

Krueger, A. and J. S. Pischke (1992). The effect of social security on labour supply: a cohort analysis of the notch generation. *Journal of Labor Economics 10*, 412–437.

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

Lemieux, T. and K. Milligan (2008). Incentive effects of social assistance: A regression discontinuity approach. *Journal of Econometrics 142*, 807–828.

Li, Q. and J. Racine (2006). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press.

Lindeboom, M. and M. Kerkhofs (2002). Health and work of the elderly. Subjective health measures, reporting erros and the endogenous relationship between health and work. *IZA Discussion Paper Series 457*.

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

Moulton, B. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics 72*, 334–338.

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

Orme, C. (1990). The small sample performance of the Information Matrix test. *Journal of Econometrics 46*, 309–331.

Pagan, A. and A. Ullah (1999). *Nonparametric Econometrics*. Cambridge University Press.

Pensions Policy Institute (2013). https://www.pensionspolicyinstitute.org.uk/default.asp?p=81.

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

Roser, M. (2015). Life expectancy. http://ourworldindata.org/data/population-growth-vital-statistics/life-expectancy/.

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

Sickles, R. and P. Taubman (1986). An analysis of the health and retirement status of the elderly. *Econometrica 54*, 1339–1356.

Sickles, R. and A. Yazbeck (1998). On the dynamics of demand for leisure and production of health: Evidence from the Retirement History Survey. *Journal of Business and Economic Statistics 16*, 187–197.

Smith, S. (2006). The retirement-consumption puzzle and involuntary early retirement: Evidence from the British Household Panel Survey. *Economic Journal 116*.

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.

Stone, C. J. (1977). Consistent nonparametric regression. *Annals of Statistics 5*, 595–645.

Thistlethwaite, D. and D. Campbell (1960). Regression-Discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology 51*, 309–317.

Wooldridge, J. (2010). *Econometric Analysis of Cross-Section and Panel Data.* MIT University Press.

Table 1: Descriptive statistics. British Household Panel Survey, 1991-2005. Men in the age range 50 to 80. $N = 3,978$, $T = 15$

	Not-retired		Retired	
General Health Questionnaire (GHQ-12). *1 (better than usual) to 4 (much less than usual)*				
	Mean	Std. Dev.	Mean	Std. Dev.
Loss of concentration	2.08	0.44	2.16	0.48
Loss of sleep	1.74	0.70	1.65	0.70
Playing useful role	2.01	0.49	2.14	0.52
Capable of taking decisions	1.98	0.40	2.06	0.40
Constantly under strain	1.95	0.72	1.76	0.72
Problem overcoming difficulties	1.72	0.65	1.68	0.68
Enjoy day-to-day	2.09	0.50	2.17	0.54
Ability to face problems	2.02	0.38	2.07	0.39
Unhappy or depressed	1.75	0.75	1.66	0.73
Losing confidence	1.51	0.67	1.57	0.70
Believe in self-worth	1.30	0.57	1.33	0.59
General happiness	2.00	0.47	2.01	0.41
Summary index	0.00	0.59	0.05	0.60
Health Problem. *1 (Has) or 0 (Has not)*				
	Mean	Std. Dev.	Mean	Std. Dev.
Arms, legs, hands...	0.29	0.45	0.45	0.50
Sight	0.04	0.18	0.10	0.29
Hearing	0.13	0.33	0.26	0.44
Skin/allergies	0.07	0.26	0.08	0.27
Chest/breathing	0.10	0.30	0.21	0.41
Heart/blood pressure	0.22	0.41	0.41	0.49
Stomach or digestion	0.07	0.26	0.11	0.31
Diabetes	0.04	0.21	0.09	0.29
Anxiety or depression	0.04	0.21	0.05	0.22
Alcohol or drugs	0.00	0.06	0.00	0.07
Epilepsy	0.00	0.06	0.01	0.08
Migraine	0.04	0.19	0.04	0.18
Summary Index	0.02	0.38	0.24	0.48
Health Checks *1 (Had) 0 (Had not)*				
	Mean	Std. Dev.	Mean	Std. Dev.
Dental	0.62	0.49	0.44	0.50
Eye	0.41	0.49	0.46	0.50
X-ray	0.13	0.34	0.20	0.40
Blood test	0.49	0.50	0.64	0.48
Cholesterol	0.26	0.44	0.29	0.45
Summary index	0.12	0.56	0.13	0.56
Use of time *Minutes a day, unless otherwise specified*	Mean	Std. Dev.	Mean	Std. Dev.
Housework (hours a week)	5.69	5.78	8.51	8.28
Housework[a]	37.12	17.70	64.20	24.92
Sleep/personal care[a]	570.01	40.10	643.02	29.40
Consumption/leisure activities[a]	422.60	64.72	537.37	26.39
Passive leisure[a]	272.42	50.20	384.13	33.28
Social leisure[a]	150.17	28.85	153.24	28.19
Leisure at home[a]	291.35	52.03	406.71	32.80
Leisure outside home[a]	131.25	26.63	130.66	27.03

[a] Calibrated time. Source: BHPS-CTU.

Table 2: The short term effects of retirement on health. The null hypothesis is that retirement has no effect.

Health outcomes.	Asymptotic Method		Randomization Inference	
	Estimate	P-value	Estimate	P-value
GHQ Summary Index	−0.373	0.112	−0.139	0.269
Health Problem Summary Index	−0.007	0.966	0.228	0.056
Health Check Summary Index	−0.197	0.491	0.151	0.315
N	925		10	
S	14		-	
Retirements	174		Based on 208	
h	[-3,2]		[-5,4]	

* Reject null hypothesis at 5% significance level.

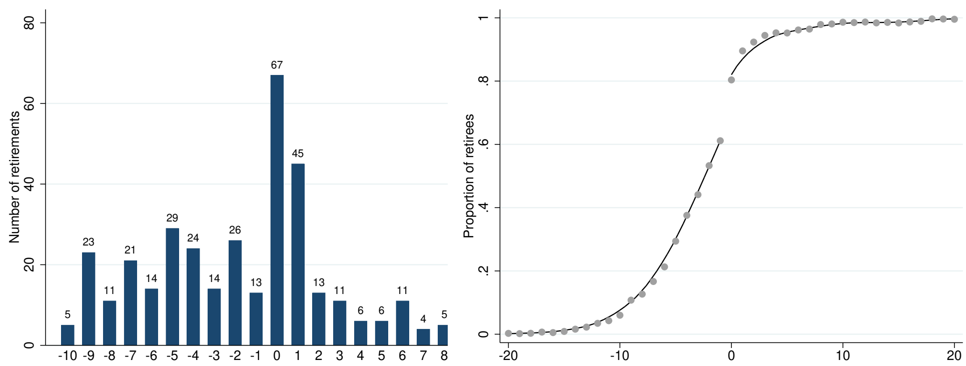
Table 3: The short term effects of retirement on time use

Time use.	Asymptotic Method		Randomization Inference	
	Estimate	P-value	Estimate	P-value
Housework	2.369	0.396	1.995	0.407
Housework[a]	−1.255	0.915	15.464	0.054
Sleep[a]	82.364	0.001*	83.071	0.000*
Consumption/Leisure[a]	146.727	0.000*	138.080	0.000*
Passive Leisure[a]	95.765	0.000*	98.631	0.000*
Social Leisure[a]	50.962	0.000*	39.577	0.011*
Leisure at home[a]	100.428	0.000*	103.400	0.000*
Leisure outside home[a]	46.298	0.000*	34.689	0.007*
N	925		10	
S	14		-	
Retirements	174		Based on 208	
h	[-3,2]		[-5,4]	

[a] Source: BHPS-CTU. Minutes per day.
* Reject null hypothesis at 5% significance level.

Table 4: The long term effects of retirement on health. The null hypothesis is that *Eligible* = *Eligible*Age =0*. Cluster-robust standard errors in parenthesis. Cluster-robust, Wild Bootstrap F-statistic P-value.

Health outcomes.			
	Eligible	Eligible*Age	F-test P-value
GHQ Summary Index	−0.033	0.016	0.634
	(0.008)	(0.010)	-
Health Problem Summary Index	0.006	0.007	0.339
	(0.003)	(0.005)	-
Health Check Summary Index	0.002	0.018	0.606
	(0.009)	(0.011)	-



(a) Retirements by age group. **(b)** Proportion of retirees.

Figure 1: Retirement in the BHPS. The abscissa measures years above/below the State Pension Age. Panel (a) summarizes the count of retirements observed per age group in the sample (in total, we identified 394 retirements between 1991-2005). Plot (b) is the Local Linear Regression (solid line) of the proportion of retirees by age group (grey dots), using bandwidth $h = 0.5\sigma_x n^{-0.2}$ where σ_x is the standard error of the explanatory variable and $n = 41$.

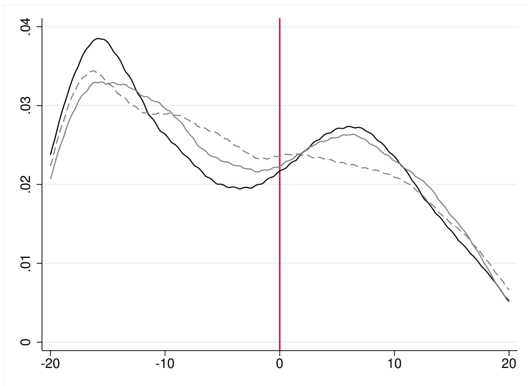
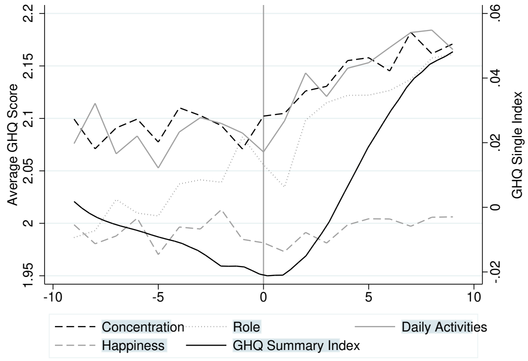
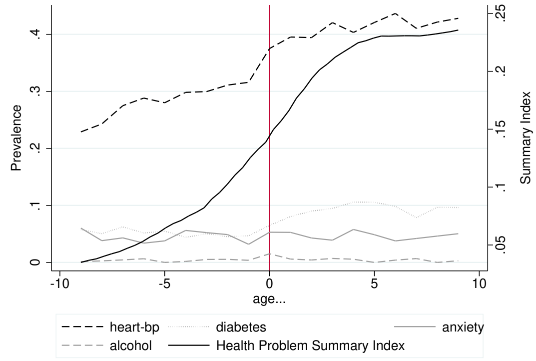


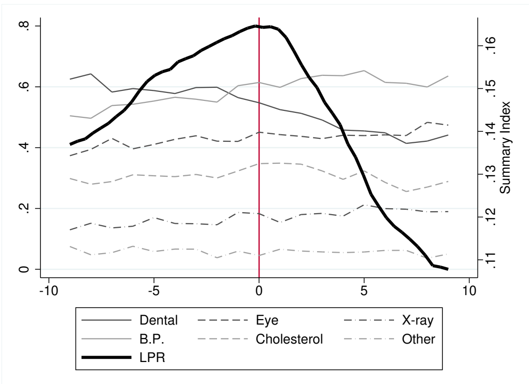
Figure 2: Distribution of age -kernel density estimator- in 1995 (solid black line), 2000 (solid grey line) and 2005 (dashed grey line). Bandwidth $h = 1.06\sigma_x n^{-0.2}$.



(a) Trends in GHQ, by age group.

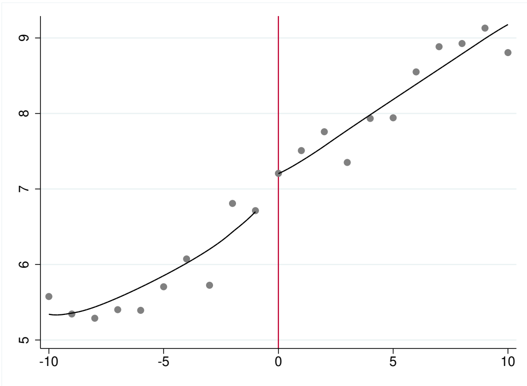


(b) Prevalence of health problems, by age group.

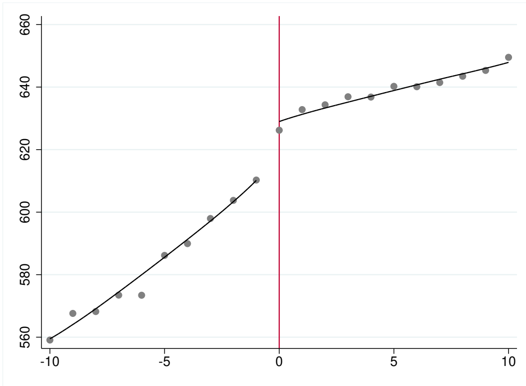


(c) % had health check, by age group.

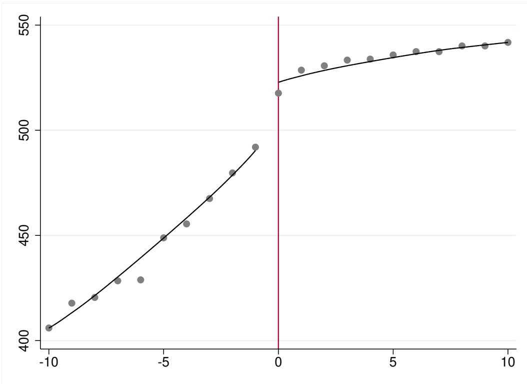
Figure 3: Trends in mental health, physical health and health checks. The abscissa measures years to/from the State Pension Age. In figure (a) he ordinate ranges from 1 (optimal health) to 4 (poor health). In figure (b), the ordinate measures the proportion of individuals in the sample with an specific problem (heart condition/blood pressure, alcohol abuse, diabetes or anxiety). The ordinate in figure (c) measures the proportion of individuals who had a check of the specified type in the previous 12 months.



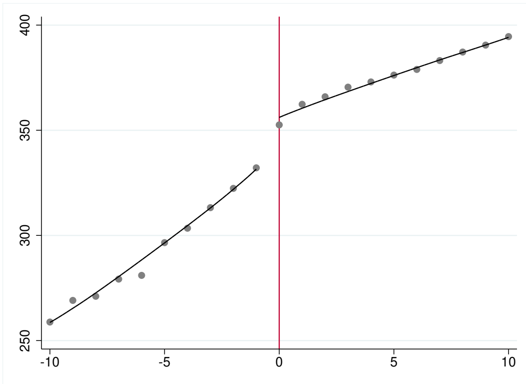
(a) Housework.



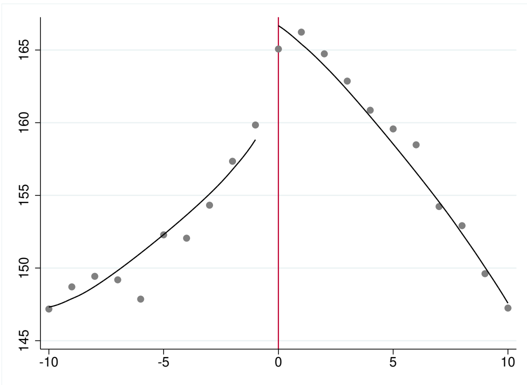
(b) Sleep



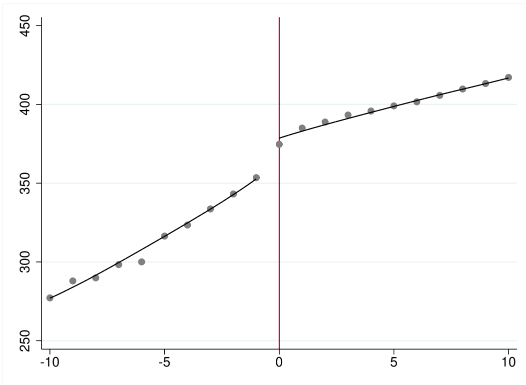
(c) Consumption/Leisure



(d) Passive Leisure



(e) Social leisure



(f) Leisure at home.

Figure 4: Allocation of time, by age group (Local Polynomial Regression -solid line). The abscissa measures years from/to the State Pension Age. In figure (a) the ordinate is measures in hours a week, while in the remaining graphs, the ordinate refers to minutes per day.