IFN Working Paper No. 1114, 2016

Retirement Blues

Gabriel Heller-Sahlgren

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

Retirement Blues[*]

Gabriel Heller-Sahlgren[†]
London School of Economics
Research Institute of Industrial Economics

Email: g.heller-sahlgren@lse.ac.uk

Abstract

This paper analyses the short- and long-term effects of retirement on mental health in ten European countries. It exploits thresholds created by regular state pension ages in a fuzzy regression-discontinuity design combined with individual-fixed effects to deal with endogeneity in retirement behaviour. The results display no short-term effects of retirement on mental health, but a large negative longer-term impact. This impact survives a battery of robustness tests, and applies to women and men as well as to people of different educational backgrounds equally. Differences compared to previous research are attributed to the study's differentiation of short- and longer-term effects as well as its utilisation of a cleaner research design. Overall, the paper's findings suggest that reforms inducing people to postpone retirement are not only important for making pension systems solvent, but with time could also pay a mental health dividend among the elderly and reduce public health care costs.

Keywords: Mental Health, Retirement, SHARE, Regression-Discontinuity Design

JEL Classifications: I10, J14, J26

---

[*] This study uses data from SHARE wave 1 and 2, release 2.5.0, as of 24th May 2011, and SHARE wave 4 release 1.0.0, as of 30th November 2012. The SHARE data collection has been primarily funded by the European Commission through the 5th framework programme (project QLK6-CT-2001- 00360 in the thematic programme Quality of Life), through the 6th framework programme (projects SHARE-I3, RII-CT- 2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th framework programme (SHARE-PREP, 211909 and SHARE-LEAP, 227822). Additional funding from the US National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, Y1-AG-4553-01 and OGHA 04- 064, IAG BSR06-11, R21 AG025169) as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions).

[†] The author thanks Henrik Jordahl, Julian Le Grand, and Olmo Silva for useful comments and discussions. This work was supported by the Economic and Social Research Council [grant number ES/J500070/1] and the Jan Wallander and Tom Hedelius Foundation.

1 Introduction

In the post-World War II period, the combination of increasing life expectancy, decreasing fertility rates, and the normalisation of old-age retirement have induced lower overall labour force participation rates in developed countries. The rise of retirement was in turn in large part due to incentives built into public pension systems, which have induced people to exit the labour market voluntarily (e.g. Gruber and Wise 1999, 2004; Hurd, Michaud, and Rohwedder 2012). Decreasing labour force participation rates have put pressure on the sustainability of pension systems, while depressing savings rates and investment levels. Recognising the severity of the situation, politicians have begun to reform state pension systems to incentivise the elderly to postpone retirement, for example by increasing the official retirement age at which state pension benefits may be drawn.

However, these policies may have unintended consequences, which must be taken into account to understand the reforms' total utility. One such issue concerns how retirement affects mental health. In the past decades, public policy has become more concerned with improving people's wellbeing in general. Thus, if retirement has a positive impact on mental health, attempts to increase the effective retirement age may thwart this policy goal, while possibly also leading to higher sick leave rates and rising health expenditures. On the other hand, if retirement has negative effects on mental health, policymakers who seek to incentivise people to postpone retirement could, if they are successful, produce a virtuous circle in which public pensions systems are made sustainable, health expenditures decreased, and mental health among the population improved.

Previous research analysing health effects of retirement yields mixed results, possibly reflecting both methodological choices and a general failure to distinguish between the short- and longer-term effects. This study investigates the impact of retirement on mental health in Europe in both a short- and longer-term perspective. Utilising several waves of panel data from the Survey of Health, Ageing, and Retirement in Europe from ten European countries, it exploits thresholds created by state pension ages in a fuzzy regression-discontinuity design (RDD), combined with individual-fixed effects, to deal with endogeneity in retirement behaviour. The idea is that these thresholds create strong economic incentives to retire once crossed, but should not affect mental health in other ways once smooth age effects are held constant. Previous research has utilised similar strategies in regular instrumental-variable (IV) set-ups – often finding positive or zero short-term effects of retirement on mental health – but tends to ignore common pitfalls that threaten its validity. To the best of our knowledge, this paper is the first to analyse both short- and longer-term effects of retirement on mental health in a full RDD framework, which is only possible because of the longitudinal nature of our study. This also allows us to investigate more thoroughly the validity of the research design utilised.

The results display no short-term impact of retirement on mental health, but strong negative effects after about four years into retirement. The longer-term effect does not differ between men and women, or between people with different educational backgrounds. It also survives a battery of robustness tests, which include analyses of narrower RDD bandwidths and using inverse probability weighting to deal with panel attrition. The differences compared to similar research are attributed to the fact that this study differentiates between short- and long-term effects as well as uses a methodology that provides a cleaner estimate of the impact of retirement per se.

Overall, the findings indicate that politicians do not face a trade-off between increasing state pension ages and improving wellbeing. Inducing people to postpone retirement is not only necessary to make pension systems sustainable, but can also be a way to improve mental health among the elderly. While pension reforms may have immediate negative mental health effects if they postpone eligibility rules late in people's lives, as some research indicates, this paper indicates that they are likely to pay a mental health dividend after some time.

The study proceeds as follows. Section 2 discusses the theoretical mechanisms potentially linking retirement to mental health; Section 3 discusses the previous literature; Section 4 discusses the data utilised; Section 5 outlines the paper's research strategy; Section 6 presents the results; and Section 7 concludes.

2 Theory

Why and how might retirement affect mental health? One way to think about the relationship between the two variables is through an economic lens in which individuals seek to maximise their utility. In Grossman's (2000) human capital model, health acts both as a direct consumption good, since it is important for people's wellbeing, and as an investment, since individuals must be in good health to be able to work and increase their lifetime earnings. And retirement may affect these properties differently: the incentive to be in good health for investment purposes is no longer present in retirement, but since individuals have more free time as retirees, the consumption value of health may increase. In the end, the theoretical net effect then depends on whether the overall marginal utility of health decreases or increases after retirement – which is far from straightforward to predict (Dave, Rashad, and Spasojevic 2006).

A similarly ambiguous story concerns other explanations that do not necessarily rely on rational choice. For example, retirement may affect individuals' stocks of social capital and networks, which research suggests have positive effects on health (e.g. d'Hombres et al. 2010; Folland 2008; Rocco, Fumagalli, and Suhrcke 2014; Ronconi, Brown, and Scheffler 2012). Yet, it is theoretically unclear how retirement affects social interactions: people may lose work colleagues, but they also have more time to establish a new, voluntarily established, social network. Perhaps reflecting this theoretical ambiguity, research finds mixed effects of retirement on the size of individuals' social networks (Börsch-Supan and Schuth 2014; Fletcher 2014). A similarly equivocal story applies to stress. While retirement may decrease work-related stress, it is in itself a life event that may be very stressful. In fact, such ambiguous stories apply to most theoretical mechanisms potentially linking retirement to mental health.

Furthermore, it is important to note that retirement may have different effects in the short- and longer-term perspective. This is partly because the effect of investments in health is likely to operate with a lag, which means that the impact of lower/higher investments does not necessarily bring negative/positive effects until after some time. Similarly, in the beginning, retirement may be perceived more like a holiday rather than as permanent labour market exit. If so, one may also expect people's mental health to improve during this period. In a longer-term perspective, however, the holiday effect may fade out and be replaced by mechanisms generating lower mental health. Alternatively, it may be the case that retirement increases stress and dissatisfaction in the short run, which then subsides in the long run as people acclimatise. All this is related more generally to the 'hedonic treadmill' hypothesis (Brickman and Campbell 1971), which stipulates that life events only affect wellbeing in the short term as individuals adapt with time. Thus, for these reasons, it is clearly important to take into account that short- and longer-term effects may differ, although it is difficult to predict how and in what ways.

3 Previous literature

The impact of retirement on health, mental and physical, has become a topic of increasing interest among researchers in economics and other fields. Correlational studies analysing the association between retirement and mental health find mixed effects (e.g. Dave, Rashad, and Spasojevic 2008; Jokela et al. 2010; Lindeboom, Portrait, and van den Berg 2002; Mein et al. 2003; Mosca and Barrett 2014; Oksanen et al. 2011; Vo et al. 2015; Westerlund et al. 2009). However, the key problem for these studies is that retirement is not random, which means that they cannot tease out causal relationships. While some of them control for individual-fixed effects, which is an improvement, they cannot remove endogeneity stemming from unobserved differences between individuals that vary over time. Overall, therefore, this research is not particularly useful for understanding the causal impact of retirement on mental health.

Improving the methodology, researchers have frequently utilised IV models using eligibility ages at which state pension benefits can be drawn to predict retirement. The idea behind this approach is that reaching these eligibility ages gives rise to significant economic incentives to retire. At the same time, the argument goes, there is no reason why reaching the threshold per se should affect mental health apart from via retirement, once smooth effects of age are held constant. This gives rise to the potential to use these thresholds in a fuzzy RDD framework, although the idea is frequently referred to as a regular IV strategy.[1] Another approach has been to utilise pension reforms in difference-in-

[1] This is probably at least partly because most previous research has not adopted a rigorous RDD framework, as discussed below.

difference set-ups, comparing individuals affected by the reforms with individuals who are not. Overall, studies using either of these research strategies tend to find no or positive effects of retirement on mental health in different settings (e.g. Behncke 2012; Blake and Garrouste 2012; Charles 2004; Coe and Zamarro 2011; Eibich 2015; Fé and Hollingsworth 2012; Fonseca et al. 2014; Johnston and Lee 2009; Latif 2013; Mazzonna and Peracchi 2014; Neuman 2008).[2]

However, there are some problems with this research. First, it tends to include a number of 'bad controls' that are endogenous to retirement, which means that it controls for some of the causal pathways through which retirement may affect mental health (Angrist and Pischke 2009). Such bad controls include consumption, marital status, and income – all of which may both affect mental health and be affected by retirement (e.g. Finnie and Spencer 2013; Haider and Stephens 2007; Stancanelli 2014). Second, most studies ignore potential differences between the short- and longer-term effects of retirement, which, as noted in Section 2, may be quite different. Third, studies evaluating pension reforms in difference-in-difference set-ups ignore that such reforms often impact on behaviour, and mental health, before individuals retire (e.g. de Grip, Lindeboom, and Montizaan 2012; Montizaan, Cörvers, and de Grip 2010). This violates the assumption that the pension reforms used for identification affect mental health solely through retirement, thereby casting doubt on the studies' internal validity.

Another problem in most previous research is that it uses instruments constructed from both regular and early retirement ages. This ignores the potential for self-selection into jobs where individuals are more likely to be able to retire early, which could undermine the validity of the findings. Furthermore, since early retirement ages often differ depending on vocation in European countries, it is difficult to find out which threshold that applies to which segment

[2] Using a regular IV set-up, based on the RDD intuition, Mazzonna and Peracchi (2014) also find weak evidence that longer time spent in retirement increases the likelihood of depression on average. However, the instrument set includes the distance of respondents' actual age from the thresholds. As discussed in Section 5.1, this variable should be included as a control in full RDD set-ups, to allow the age trend to differ on both sides of the threshold and thus decrease the risk that the retirement variable picks up non-linear effects of age. This is not a trivial concern, especially since the authors use a broad age bandwidth and control for a linear age trend only.

of the population. While discontinuities arising at the state pension age should help ameliorate measurement error in retirement status observed in survey data – since the discontinuities are shaped by institutional rules and are therefore uncorrelated with potential measurement error – it is therefore also possible that discontinuities at (alleged) early retirement ages induce new measurement error. This, in turn, is likely to produce attenuation bias in the estimates (Angrist and Pischke 2009). Of course, it is also not clear whether early and regular retirement have the same effects; since countries are currently in the process of increasing regular state pension ages specifically, disentangling the separate impact of retirement at these ages is important.

Moreover, previous multi-country research has generally ignored the possibility that the impact of age on mental health differs across countries. If such differential age effects are correlated with the state pension ages utilised in the analysis, estimates may be biased. It is thus necessary to also control for interactions between age and country dummies in the RDD set-up.

Another general problem with previous RDD-like studies is that they ignore common pitfalls associated with such designs. For example, they do not present results in which the impact of the running variable (age) is allowed to differ on each side of the utilised discontinuities – which should be part of any standard RDD framework (Lee and Lemieux 2010). Similarly, most existing studies do not analyse the sensitivity of the findings by narrowing the range of data analysed around the discontinuity (Angrist and Pischke 2015). Instead, they choose an arbitrary (and wide) range of data, without sufficiently exploring non-linear effects of the running variable or the results' sensitivity to the specific data range. These issues make it rather unsurprising that the authors discuss their methods as straight IV strategies rather than fuzzy RDDs, and thus decrease confidence in the internal validity of the findings.[3]

Overall, therefore, while previous research most often finds no or positive effects of retirement on mental health, it suffers from limitations. Perhaps most important is that previous studies do not generally separate the short- from longer-term effects of retirement. This study aims to remedy the shortcomings

[3] The exception is Eibich's (2015) paper, which finds a short-term positive impact of retirement on mental health in Germany.

highlighted and provide a more rigorous evaluation of the impact of retirement on mental health in Europe, using a fuzzy RDD strategy. Section 5 discusses this strategy in detail.

4 Data

Motivated by the flaws noted in Section 3, this study utilises data from the first, second, and fourth waves of the Survey of Health, Ageing, and Retirement in Europe (SHARE), conducted at different points in 2004-05, 2006-07, and 2011-12 respectively. In these waves, SHARE provides information on a wide range of background and outcome variables from representative samples of individuals who are aged 50 and over in ten European countries (see Börsch-Supan et al. 2013): Austria, Belgium, Denmark, France, Germany, Italy, the Netherlands, Spain, Sweden, and Switzerland.[4] Analysing panel data spanning over several SHARE waves allows the paper to investigate both the short- and longer-term effects of retirement on mental health as well as conduct relatively strong robustness tests on the validity of the design.[5] The data discussed are summarised in Table 1 at the end of this section.

4.1 Sample

The main sample includes individuals who were interviewed in the first, second, and fourth wave of SHARE and who were 50–75 years old at the second wave interview. The state pension ages for the ten European countries utilised in the RDD, which take into account pension reforms in recent decades, are displayed in Table A1 in the Appendix. The threshold varies between 55 and 67 years, depending on country, gender, and cohort. However, individuals reaching their state pension age between the first and the second SHARE waves – which is the relevant threshold for the RDD outlined in Section 5 – were between 60 and 65.[6]

[4] The SHARE dataset has been used widely in related economic research (e.g. Coe and Zamarro 2011; Godard 2016; Mazzonna and Peracchi 2012; Rohwedder and Willis 2010).
[5] The third wave was a special survey that did not enquire respondents about their mental health.
[6] The exception is in Denmark, where individuals born before July 1939 instead faced a retirement age of 67. To ensure that individuals of the same age are analysed in all countries, the right-hand side of the RDD bandwidth is calculated from age 65 also for the Danish sample. This makes the bandwidth even narrower for Danes born before July 1939. However, all results are essentially identical if the Danish bandwidth is instead extended by two years on the right-hand side. This is not too surprising since it only increases the sample size by 25-33 individuals.

This gives a bandwidth of about ten years for the RDD, which is in line with the one used by Moreau and Stancanelli (2013), and Stancanelli and Van Soest (2012).[7] The total sample thus consists of maximum 8,552 individuals across the ten countries, all observed three times, with the exact sample size depending on the restrictions discussed in Section 4.2. However, a narrow bandwidth of about three years is also utilised in robustness tests, which means that people who were 57-68 years old at the second wave are analysed. This decreases the sample size to maximum 4,697 individuals, again all observed three times.

4.2 Retirement

The study employs three different definitions of retirement. First, it classifies individuals as retired if they claim to be retired or give a date at which they retired that precedes the interview date. Given the methodology outlined in Section 5, this means that we analyse the impact of retirement compared to the status of employed/self-employed, homemaker, unemployed, being engaged in other activities, as well as the permanently ill or disabled. This definition is useful since all non-retirees, not only those who currently are engaged in paid work, are likely to respond to retirement incentives at the state pension age – and, for policy purposes, the mental health impact of retirement due to those incentives is relevant regardless from which category respondents officially retire. Like Eibich (2015), we thus include individuals who were not working prior to retirement in the sample.

However, to ensure that the above definition of retirement does not drive the results, the study also employs an alternative definition based instead on retirement from the labour force only. In this definition, which is similar to Coe and Zamarro's (2011), homemakers, respondents who report being permanently ill or disabled, and those who are engaged in other activities are instead included in the retirement category, as long as they at the same time do not report to have done any paid work in the past four weeks. Respondents in these categories who report having done paid work are still included as non-retirees.[8] Finally, in the third definition, the sample is instead simply restricted to individuals who claim

[7] Since rules vary across countries, and depend on gender as well as cohort within countries, the bandwidth varies slightly in accordance with the state pension ages noted.

[8] Results are almost identical if the latter respondents instead are excluded from the sample.

to be retired or to be working, which include employed/self-employed respondents and other respondents who report having done paid work in the past four weeks.

4.3 Mental health

Following previous research in the field, the primary measure of mental health is the Euro-D scale. The scale was developed specifically as a common depression gauge in the European Union and has shown to be valid for research purposes (see Prince et al. 1999). The scale ranges from 0 to 12, with higher values indicating stronger depressive tendencies, counting whether or not respondents reported having had problems with the following in the past month: appetite, concentration, depression, enjoyment, fatigue, guilt, interest, irritability, pessimism, sleep, suicidality, and tearfulness. Furthermore, like previous research, the paper also analyses the likelihood that respondents crosses the conventional threshold for clinical depression, which is defined as a score of 4 or more on the Euro-D scale. In both cases, therefore, higher values indicate worse mental health.

4.4 Control variables

If the design discussed in Section 5 produces random variation in respondents' retirement statuses, the only necessary control to ensure a causal interpretation of the estimates is respondents' age (and in some cases also its polynomials to control for non-linear effects of the running variable). This is especially true since individual-fixed effects as well as interactions between country-fixed effects and age are included. However, to test robustness of the estimates, it is also useful to include lagged mental health status to ensure that mean reversion does not bias the findings. Doing so may also increase precision in the estimates.[9] It is also possible to interact individuals' gender and educational levels with the indicator for retirement in order to investigate potential heterogeneous effects. These issues are discussed more formally in Section 5.1.

[9] Similarly, lagged dependent variables are often included to increase precision in analyses of randomised experiments (Angrist and Pischke 2009).

4.5 Attrition

It is important to consider potential selection problems due to non-random panel attrition, which is entirely ignored by most studies exploiting panel surveys to evaluate various outcomes of retirement (e.g. Bonsang, Adam, and Perelman 2012; Eibich 2015). Attrition does not affect the study's internal validity, but could pose a threat to its external validity if it makes the sample unrepresentative of the underlying population from which it is drawn. Like in most panel surveys, and especially in those focusing on the elderly, attrition in SHARE is substantial: about 50 per cent of people interviewed in the first SHARE wave disappear by the fourth wave. Of course, this has no bearing on the results if attrition is unrelated to how retirement affects mental health, but this cannot be established a priori.[10]

However, it is possible to include individual-fixed effects, which control for time-invariant variables that affect both respondents' propensity to remain in the panel's fourth wave as well as their retirement behaviour and mental health. In relatively short panels, like the one analysed here, it appears reasonable to assume that attrition is mainly caused by such time-invariant variables. As discussed in Section 5, this solution is therefore adopted in the paper's standard methodology.

Nevertheless, as a further robustness check, the study also exploits inverse probability weighting, which allows attrition to be non-random as long as its causes are captured by individuals' observable characteristics at the time before they drop out of the panel (Moffit, Fitzgerald, and Gottschalk 1999). This means that we estimate the probability that individuals who were interviewed in the first wave remain in the fourth wave, from variables observed in the final wave in which they participated prior to the fourth wave.[11] These variables include age, employment status, marital status, education level, and a battery of self-assessed, mental, and physical health variables as well as indicators for cognitive

[10] Interestingly, however, research indicates that mental health and employment status are barely related at all to the probability of attrition after the first wave of SHARE (De Luca 2009).

[11] This means that the probability of remaining in the fourth wave among individuals dropping out after the first wave is predicted based on their observable characteristics in the first wave. Similarly, the probability among those who dropped out after the second wave and those who remained in the fourth wave are predicted based on their observable characteristics in the second wave.

achievement.[12] In addition, country-fixed effects are included to ensure that differential attrition across countries does not bias the findings. The inverse of the probability of remaining in the panel, as predicted by this model, is then used as weights in the robustness regressions analysing the impact of retirement on mental health.[13] If attrition poses a serious problem for the study's external validity, one would expect the results from the weighted models to differ significantly compared with the ones that exclude weights. If the results are very similar, on the other hand, it is unlikely that attrition poses a serious problem. Although it is of course impossible to demonstrate conclusively a complete absence of attrition bias, the combination of individual-level fixed effects and inverse probability weighting based on a battery of baseline controls simply leave little room for any remaining bias.

Table 1: Descriptive statistics

	50-75 year bandwidth				57-68 year bandwidth			
Demographics	Average	Standard deviation	Min	Max	Average	Standard deviation	Min	Max
Age	62.95	6.65	50	75.92	62.84	3.47	57	68.92
Reached state pension age	0.46	0.50	0	1	0.43	0.50	0	1
Retired (1)	0.51	0.50	0	1	0.56	0.50	0	1
Retired (2)	0.65	0.48	0	1	0.71	0.46	0	1
Retired (3)	0.60	0.49	0	1	0.66	0.48	0	1
Mental health	Average	Standard deviation	Min	Max	Average	Standard deviation	Min	Max
Euro-D	2.27	2.11	0	12	2.18	2.08	0	12
Clinical depression	0.24	0.43	0	1	0.23	0.42	0	1

Note: All demographic variables are measured at the second interview while all outcome variables are measured at the third interview. Both mental health variables are scaled so that higher values indicate worse mental health. The descriptive statistics are almost identical when restricting the sample based on the third definition of retirement, even though the sample size decreases.

[12] More specifically, the health and cognitive indicators include: self-assessed health, the Euro-D scale, the number of drugs taken, the number of diagnosed conditions, the number of limitations with activities of daily living, the number of mobility limitations, as well as memory and numeracy scores.

[13] The regression predicting the probability to remain in the panel is estimated using a probit model, but all weighted results are essentially identical when using a linear probability model instead.

5 Research design

As noted in Section 3, any valid research design used to evaluate the causal effect of retirement on mental health must take into account that the former is likely endogenous to the latter. This section discusses the research design employed in the study to deal with endogeneity in retirement behaviour.

5.1 Obtaining exogenous variation in retirement behaviour

When analysing the impact of retirement on mental health, the easiest strategy is to estimate:

$$mh_i = \alpha + \beta_1 r_i + \beta_3 x_i + \varepsilon_i \tag{1}$$

where mh_i is the mental health outcome analysed; r_i is a dummy variable taking either the value 1 (retired) or 0 (not retired); and x_i is a vector of control variables assumed to affect both retirement and mental health.

The critical assumption is that $Cov(r_i, \varepsilon_i | x_i) = 0$. But if x_i does not include all factors that affect both r_i and mh_i, or if mh_i has an independent impact on r_i, the results will be plagued by endogeneity. In addition, measurement error in r_i may generate attenuation bias. Any of these issues would ensure that $Cov(r_i, \varepsilon_i | x_i) \neq 0$ (Angrist and Pischke 2009). And as noted in Section 3, the critical assumption is not likely to hold for this paper's purposes. To ensure a causal interpretation of the paper's findings, it is thus key to obtain variation in retirement behaviour that is exogenous to mental health.

To do so, the paper proposes a fuzzy RDD based on retirement ages in European pension systems, which create age thresholds at which economic incentives to retire increase substantially. In this set-up, the discontinuities act as instruments for individuals' employment status in a 2SLS model, with age as the running variable (Angrist and Piscke 2009). As noted in Section 3, this idea has been exploited in previous research, but this generally does not use a full RDD framework. Furthermore, the set-up has never before been used to investigate the longer-term effects of retirement. Because of the potential problems that may arise by using the early retirement age, as discussed in Section 3, the paper focuses solely on the regular state pension age.

Because of the multi-country context analysed, the discontinuity in the RDD varies across individuals, since eligibility ages are different across countries and within countries. For example, women and men are often not affected by the same retirement age, and in some cases the threshold also depends on when individuals faced their retirement decisions. This, however, is irrelevant for the design's validity, although it is important to allow the effect of age to differ across countries. The idea behind the design is formalised as follows:

$$P(r_i = 1|age_i) = \begin{cases} f_1(age_i) \text{ if } age_i \geq sp_i \\ f_0(age_i) \text{ if } age_i < sp_i \end{cases}, \text{where } f_1(sp_i) \neq f_0(sp_i)$$

where sp_i is the applicable eligibility age. For this paper's purposes, the assumption is that $f_1(sp_i) > f_0(sp_i)$, since economic incentives raise the likelihood that individuals retire when they reach the eligibility age. Thus, the probability of $r_i = 1$ as a function of age_i can be written:

$$P(r_i = 1|age_i) = f_0(age_i) + [f_1(age_i) - f_0(age_i)]\, \overline{sp}_i$$

where $\overline{sp}_i$ is a dummy variable with the value of 1 if $age_i \geq sp_i$ and 0 if $age_i < sp_i$. In this case, individuals cannot manipulate the running variable to end up on either side of the discontinuity, which would pose a problem for the RDD (Imbens and Wooldridge 2009). Yet, the strategy is dependent on the ability to separate smooth effects of age_i from the impact at $\overline{sp}_i$. The paper follows most of the literature and assumes a quadratic age trend in the baseline estimates. Recent research shows that estimates including higher order polynomials are often misleading and that they should not be used. Instead, it is preferable to also display estimates with narrower bandwidths (Gelman and Imbens 2014). The paper thus also displays results from models with a narrower bandwidth, as discussed in Section 4.1. We then include a linear age trend instead, but also report estimates with the quadratic specification in the Appendix.

A straightforward way to analyse the longer-term impact of retirement on mental health with the above framework would then be to estimate the following 2SLS model:

$$\begin{aligned} r_{it-1} &= \alpha + \beta_1 \overline{sp}_{it-1} + \beta_2 age_{it-1} + \beta_3 {age_{it-1}}^2 + \gamma_c(age_{it-1}) \\ &\quad + \gamma_c({age_{it-1}}^2) + \delta_i + \mu_t + \varrho_t + \varepsilon_{it} \qquad (2) \end{aligned}$$

$$mh_{it} = \alpha + \beta_1\widehat{r_{it-1}} + \beta_2 age_{it-1} + \beta_3 {age_{it-1}}^2 + \gamma_c(age_{it-1}) + \gamma_c({age_{it-1}}^2) + \delta_i + \mu_t + \varrho_t + \varepsilon_{it} \quad (3)$$

where $\widehat{r_{it-1}}$ is the predicted values of r_{it-1} from the first stage with $\overline{sp}_{it-1}$ as the excluded instrument; γ_c denotes country dummies, which are interacted with age_{it-1} and ${age_{it-1}}^2$; δ_i denotes individual-fixed effects; and μ_t and ϱ_t represent separate year- and month-fixed effects respectively.[14] The interaction between γ_c and the age variables are included to account for cross-country differences in the impact of the running variable. Including δ_i in turn means that the model controls for all time-invariant individual characteristics that predict retirement, mental health, and panel attrition.[15]

However, like most previous research, the above model assumes that the running variable's impact is the same on both sides of the threshold, which is not necessarily the case. To take this into account, it is necessary to replace age_{it-1} and ${age_{it-1}}^2$ with $\widetilde{age}_{it-1}$ and ${\widetilde{age}_{it-1}}^2$, which in the estimation denote $(age_{it-1} - sp_{it-1})$ and $(age_{it-1} - sp_{it-1})^2$ respectively, while also including their interactions with $\overline{sp}_{it-1}$. In addition, it is then necessary to interact $\widetilde{age}_{it-1}$ and ${\widetilde{age}_{it-1}}^2$ with γ_c to account for cross-country differences in age effects below and above the discontinuities. This means that the running variable and its polynomials are centred and interacted with the discontinuity, which ensures that the coefficient of $\widehat{r_{it-1}}$ still measures the jump in the dependent variable at the threshold (see Angrist and Pischke 2009, 2015). The estimation then reads:

$$r_{it-1} = \alpha + \beta_1\overline{sp}_{it-1} + \beta_2\widetilde{age}_{it-1} + \beta_3{\widetilde{age}_{it-1}}^2 + \beta_4\overline{sp}_{it-1}(\widetilde{age}_{it-1}) + \beta_5\overline{sp}_{it-1}({\widetilde{age}_{it-1}}^2) + \gamma_c(\widetilde{age}_{it-1}) + \gamma_c({\widetilde{age}_{it-1}}^2) + \gamma_c[\overline{sp}_{it-1}(\widetilde{age}_{it-1})] + \gamma_c[\overline{sp}_{it-1}({\widetilde{age}_{it-1}}^2)] + \delta_i + \mu_t + \varrho_t + \varepsilon_{it} \quad (4)$$

$$mh_{it} = \alpha + \beta_1\widehat{r_{it-1}} + \beta_2\widetilde{age}_{it-1} + \beta_3{\widetilde{age}_{it-1}}^2 + \beta_4\overline{sp}_{it-1}(\widetilde{age}_{it-1}) + \beta_5\overline{sp}_{it-1}({\widetilde{age}_{it-1}}^2) + \gamma_c(\widetilde{age}_{it-1}) + \gamma_c({\widetilde{age}_{it-1}}^2) + \gamma_c[\overline{sp}_{it-1}(\widetilde{age}_{it-1})] + \gamma_c[\overline{sp}_{it-1}({\widetilde{age}_{it-1}}^2)] + \delta_i + \mu_t + \varrho_t + \varepsilon_{it} \quad (5)$$

[14] Including interactions between μ_t and ϱ_t generates essentially identical estimates. However, because of the few observations interviewed in the same year-month combination, the number of clusters is insufficient to reliably calculate standard errors and test statistics.
[15] When analysing the binary indicator of depression, a non-linear model could also be an alternative. However, regular 2SLS is preferable even in these cases and especially when fixed effects are included (Angrist and Pischke 2009).

Since δ_i is included, the model effectively analyses the effect of a change in retirement status between the first and second waves (about two years) on the change in mental health between the second and fourth waves (about four years). Because the shift in retirement status may occur at any point between the two interviews, but mental health is measured at the exact point of the interviews, this means that the model focuses only on the change in mental health that occurs after the shift in retirement status has taken place. This is useful since we otherwise risk picking up changes in mental health that occurred between the first and second interviews but before the shift in retirement status occurred. The set-up also makes it possible to test the model's assumptions by controlling for the change in mental health between the first and second waves, when the shift in retirement status occurs, as discussed in Section 5.2.

However, the set-up also risks ignoring potential immediate mental health effects of a shift in retirement status between the first and the second waves. This is not a problem for the validity of the analysis per se, but it may affect the interpretation of the coefficients. For example, if the immediate effect of retirement is negative, and the above set-up shows positive effects, this may not necessarily indicate a positive longer-term impact but could also mean that mental health is reverting following the negative immediate impact. If so, the long-term impact may instead best be interpreted as zero. Yet, as discussed in Section 5.2, if we also control for mh_{it-1}, the change in mental health between the first and the second waves is held constant. This ensures that any difference between the treatment and control groups due to differential mental health trends as retirement is taking place is ignored, which means that any detected long-term impact does not pick up reversion from short-term effects. However, to further ensure that our interpretation is accurate, we also analyse the change in mental health between the first and fourth waves in robustness tests.

Of course, another way to investigate these issues is to analyse the short-term effects directly. Indeed, an important rationale behind the study is to investigate whether the short-term effects of retirement differ from the longer-term effects. Thus, in models analysing the short-term impact of retirement, r_{it} is included as an endogenous predictor with $\overline{sp}_{it}$ as its instrument, giving $\widehat{r_{it}}$ in the second-stage equations. In these estimations, all variables in equations (4) and (5) are included but measured at t instead of $t-1$.[16] Effectively, the models then analyse the impact of retiring between the first and the second waves on the change in mental health over the same period. In this way, it is possible to compare and contrast the short- and longer-term effects of retirement on mental health.[17] However, it is not possible to conduct the same robustness tests to ensure we do not pick up changes in mental health prior to retirement.

The RDD discussed analyses the average impact of retirement on mental health, but it is also important to investigate potential heterogeneous effects depending on gender – represented by dummy variable g_i, which takes either the value 1 (female) or 0 (male) – and ISCED-97 educational levels (ed_i), which range from 0 (pre-primary education) to 6 (research degree). We collapse the variable into three levels: low (0-2), medium (3-4), and high (5-6). We also create dummy variable ed_l_i, which takes the value 1 for education levels below the mean/median (upper-secondary school) and zero otherwise. By interacting these variables with r_{it-1} (or r_{it}), and in turn instrumenting the interactions with the variables' interactions with $\overline{sp}_{it-1}$ (or $\overline{sp}_{it}$), it is possible to analyse such potential heterogeneous effects within the above framework.

5.2 Assumptions and potential issues

As discussed, a fuzzy RDD is fundamentally an IV model. A useful instrument must first of all be relevant, which in this case means that it should correlate with retirement. It must also be valid, in this case meaning that it must be exogenous to mental health, and further satisfy the monotonicity requirement, which here means that there cannot be people who are induced not to retire because they reach the state pension age (Imbens and Angrist 1994). All these requirements also apply to fuzzy RDDs (Hahn, Todd, and Van der Klaauw 2001).

Does the instrument satisfy the above requirements? The short answer is: most likely. First, public pension systems, and incentives in them, have been

---

[16] In estimations analysing the short-term effects, the sample is restricted to observations in the first and the second waves. This ensures that the short- and longer-term effects are compared among the same individuals. However, results for the direct impact over the second and fourth waves are presented in Table A2 in the Appendix.

[17] The main results are very similar if separate indicators for $\widetilde{age}_{it}$ and $\widetilde{age}_{it}^{\,2}$, as well as their interactions with $\overline{sp}_{it}$ and γ_c, are also included when analysing longer-term effects. This is expected if the strategy induces random variation in past retirement behaviour.

found to be important for retirement behaviour (e.g. Börsch-Supan, Brugiavini, and Croda 2009; Gruber and Wise 1999, 2004; Hurd, Michaud, and Rohwedder 2012). The instrument is thus certainly relevant, as displayed by previous research analysing the impact of retirement on health and cognitive functioning in the same context (e.g. Coe and Zamarro 2011; Mazzonna and Peracchi 2012). Section 6 displays that this holds true also in this paper. Since there are considerable financial incentives to retire once an individual reaches the retirement age, it is also unlikely that it induces some people not to retire because they reach it, in which case the monotonicity requirement would not be satisfied. For these reasons, the first two requirements are probably satisfied.

In regard to validity, $\overline{sp}_{it-1}$ is not allowed to affect mh_{it} apart from via r_{it-1} once smooth age effects and individual-fixed effects are controlled for. It is highly likely that this requirement is satisfied, although it is certainly possible to come up with stories why it would not be. For example, as Behncke (2012) notes, reaching the official eligibility age may have a direct impact – positive or negative – on mental health if it is considered a milestone in a person's life. While this possibility cannot be ruled out, it is unlikely to pose a problem for this study. For example, if the state pension age is considered a milestone, it is likely because it increases the probability of retirement. Such effects could thus be seen as part of the total impact of retirement on mental health. Furthermore, an important focus of the paper is on the longer-term effects for which any possible short-term milestone effects are less relevant.

Two other potential issues that may threaten the study's validity, which are generally ignored by previous research, are that (1) individuals may select into jobs where they can retire early, and (2) retirement planning may affect local estimates around the state pension age (see Behncke 2012). Nevertheless, these potential issues are unlikely to pose any problems for this paper. First, as noted above, the design circumvents the problem of self-selection by solely using the regular retirement age in the RDD. Second, including δ_i captures all period-invariant differences between individuals, which means that only changes in retirement planning between the waves analysed could potentially bias the findings. Thus, the powerful combination of RDD and individual-level fixed effects makes bias occurring from retirement planning unlikely. Furthermore, retirement planning should be seen as part of the effects on mental health of retirement induced by the official state pension age. Indeed, people are always likely to plan for voluntary retirement, which is now the norm in Europe and thus key for external validity purposes. Indeed, the impact of planned retirement is of crucial importance to policymakers, given their current efforts to change state pension systems in order to induce higher labour force participation.

In addition, having access to three waves of panel data means that it is possible to investigate more thoroughly whether the RDD produces random variation in retirement status and thus whether the instrument is valid. Indeed, if this is the case, the coefficient of $\widehat{r_{it-1}}$ should not differ much when including mh_{it-1} as independent variable, albeit precision may increase since mh_{it-1} is likely to be a good predictor of mh_{it} (Lee and Lemieux 2010). Including mh_{it-1} also makes it possible to test and control for potential mean reversion, which may affect the findings (e.g. Angrist and Pischke 2009). Similarly, as noted in Section 5.1, it also makes it possible to test whether potential short-term retirement effects between the first and the second waves affect the interpretation of the longer-term effects between the second and fourth waves. Because individual-fixed effects are included, mh_{it-1} is mechanically correlated with ε_{it} (Nickell 1981), but this is not a problem for this study's purposes, as long as $\widehat{r_{it-1}}$ is orthogonal to mh_{it-1}, which the latter is supposed to test.

Finally, as noted, the bandwidth utilised may impact the findings in the RDD. Choosing the bandwidth involves a trade-off between consistency and efficiency: a smaller bandwidth decreases the likelihood of bias, but fewer observations simultaneously increase the variance (Lee and Lemieux 2010). Unlike most previous research, this paper restricts the bandwidth in robustness tests to ensure that the results do not hinge on the one utilised in the main set-up.

Similar to IV models, the fuzzy RDD estimates a local average treatment effect (LATE) of retirement on mental health (Imbens and Angrist 1994). In this study, the LATE is relevant for individuals who are more likely to retire once they reach the relevant state pension eligibility age. While we cannot analyse the average treatment effect of retirement through all venues, the estimates are still highly relevant for policymakers. This is because the state pension system is a key tool with which they may influence people's retirement behaviour.

6 Results

Table 2 displays naïve OLS estimates that include individual-fixed effects, but do not correct for endogeneity in retirement behaviour. It shows highly insignificant coefficients in models analysing the Euro-D scale and the threshold for clinical depression. This holds true when analysing the short-term association between the first and the second wave, and when analysing the longer-term association between the second and fourth waves. Meanwhile, the results in Table A2 show a negative direct association between the second and fourth waves, indicating a positive relationship between retirement and mental health over that period. Overall, the naïve estimates thus support previous findings of a zero or positive impact of retirement on health.

Table 2: Estimates from individual-fixed effects OLS models

	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Short-term associations						
r_{it}	-0.04	-0.03	-0.04	0.00	-0.00	0.00
Waves 1-2	(0.07)	(0.07)	(0.07)	(0.01)	(0.02)	(0.01)
Long-term associations						
r_{it-1}	-0.02	0.02	-0.11	0.01	0.01	-0.01
	(0.07)	(0.07)	(0.08)	(0.02)	(0.02)	(0.02)
n	8,552	8,552	7,096	8,552	8,552	7,096

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level are in parentheses. All regressions include individual-, year-, and month-fixed effects, a quadric age trend and interactions with country dummies. They analyse observations within the 10-year bandwidth.

Turning to our main research strategy to deal with endogeneity, Table 3 displays the estimates from the fuzzy RDD model in equations (4) and (5). The first stage results show that the instrument is strong, with the F statistics always displaying values considerably higher than 23.1, which is the relevant threshold when using cluster-robust standard errors (Olea and Pflueger 2013). The unreported coefficients indicate that reaching the state pension age threshold increases the likelihood of retirement by 11-18 percentage points, depending on the definition of retirement, and are always statistically significant at the 1 per cent level. It is thus clear that there is sufficient variation in retirement behaviour in the data, and that it can be predicted well by the state pension age threshold. In other words, the instrument in our fuzzy RDD is certainly relevant.

The second-stage results, in turn, display a different picture in regard to the impact of retirement on mental health compared with the naïve OLS estimates. First, there is no evidence that retirement has any short-term effects, with all coefficients being highly non-significant. As displayed in Table A2, this also holds true for the impact of retirement over the second and fourth waves. Second, however, the estimates uniformly indicate that retirement has a large, negative longer-term impact. The coefficients in the second panel display that retirement increases the overall Euro-D score by 0.71–1.11 standard deviations, depending on which retirement definition that is utilised. Similarly, retirement increases the likelihood of being clinically depressed in the longer-term by 28-43 percentage points. Since we do not detect any short-term effects on mental health between the first and second waves, this indicates that the longer-term coefficients pick up a negative effect of retirement on mental health rather than merely a reversion following positive short-term effects.

Table 3: Main RDD estimates (10-year bandwidth)

	RDD	RDD	RDD	RDD	RDD	RDD
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
			Short-term effects			
$\widehat{r_{it}}$	-0.11	-0.17	-0.33	-0.03	-0.04	-0.07
Waves 1-2	(0.46)	(0.72)	(0.59)	(0.10)	(0.16)	(0.14)
F statistic	98.03	44.41	57.09	98.03	44.41	57.09
Hausman	0.88	0.84	0.64	0.80	0.81	0.57
			Long-term effects			
$\widehat{r_{it-1}}$	1.55***	2.44***	1.68**	0.29**	0.45**	0.38**
	(0.52)	(0.86)	(0.67)	(0.11)	(0.18)	(0.15)
F statistic	98.75	44.95	57.66	98.75	44.95	57.66
Hausman	<0.01	<0.01	<0.01	0.01	0.01	<0.01
$\widehat{r_{it-1}}$	1.49***	2.35***	1.52**	0.28***	0.43**	0.35**
	(0.47)	(0.78)	(0.61)	(0.10)	(0.17)	(0.14)
mh_{it-1}	YES	YES	YES	YES	YES	YES
F statistic	98.72	44.93	57.62	98.75	44.94	57.70
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	8,552	8,552	7,096	8,552	8,552	7,096

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level are in parentheses. All models include the variables in equations (4) and (5).

This interpretation receives further support from the models including mh_{it-1} in the third panel. These display almost identical longer-term estimates, while precision increases somewhat. This is expected if the research strategy captures the causal impact of retirement on mental health, as discussed in Section 5. Since these models also control directly for the change in mental health that occurred simultaneously as the change in retirement status, they further indicate that the long-term impact of retirement is negative in absolute terms rather than just inducing a reversion from short-term positive effects. Overall, therefore, all estimates from the main RDD point in one direction: retirement has no impact on mental health in the short run, but a large negative effect in the longer term.

Meanwhile, in the models evaluating the longer-term effects, the Hausman tests always reject the null hypothesis of no endogeneity, which indicates that OLS estimates are biased downward. This is unsurprising given the results in Table 2, but may be unexpected since potential reverse causality – with poor mental health increasing the probability of retirement – is most likely to bias the OLS estimates in favour of finding a negative effect of retirement on mental health. Yet, the differences are certainly plausible since omitted variables and measurement error may simply bias estimates downward more than potential reverse causality biases estimates upward.[18]

6.1 Robustness tests

6.1.1 Three-year bandwidth

How sensitive are the results to the specific bandwidth? Table 4 displays estimates from models with the sample restricted to individuals aged within approximately three years of the state pension threshold and instead including a linear age trend. Again, there is little evidence of any short-term effects. However, the longer-term effects remain statistically significant in all cases, despite the fact that 45 per cent of the main sample is dropped. And although the size of the coefficients decreases slightly, they are not statistically different from the baseline estimates. Estimates in Table A3 also show that the long-term effect remains just as conspicuous when including a quadratic age trend, despite the

[18] Indeed, research analysing the impact of retirement on cognitive ability also finds that OLS results are biased in the same way (Bonsang, Adam, and Perelman 2012).

narrow bandwidth, with the absolute effect size then instead growing somewhat compared to the baseline estimates. Overall, therefore, the results are robust to using a narrower bandwidth, which further supports the idea that our research design does capture the causal impact of retirement on mental health.

Table 4: RDD estimates (3-year bandwidth with a linear age trend)

	RDD	RDD	RDD	RDD	RDD	RDD
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
			Short-term effects			
$\widehat{r_{it}}$	-0.04	-0.07	-0.23	0.01	0.01	-0.01
Waves 1-2	(0.39)	(0.60)	(0.48)	(0.09)	(0.14)	(0.11)
F statistic	125.79	60.73	79.58	125.79	60.73	79.58
Hausman	0.99	0.99	0.79	0.97	0.91	0.87
			Long-term effects			
$\widehat{r_{it-1}}$	1.28***	1.96***	1.21**	0.25**	0.38**	0.27**
	(0.43)	(0.69)	(0.52)	(0.10)	(0.15)	(0.12)
F statistic	127.25	61.61	80.54	127.25	61.61	80.54
Hausman	<0.01	<0.01	0.01	<0.01	<0.01	0.02
$\widehat{r_{it-1}}$	1.26***	1.93***	1.10**	0.26***	0.39***	0.27**
	(0.39)	(0.63)	(0.48)	(0.09)	(0.14)	(0.11)
mh_{it-1}	YES	YES	YES	YES	YES	YES
F statistic	127.21	61.59	80.40	127.25	61.62	80.55
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	4,697	4,697	3,807	4,697	4,697	3,807

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level in parentheses.

6.1.2 Longer-term effects between the first and fourth waves

The study has so far found strong negative effects of retirement between the first and second SHARE waves on the change in mental health between the second and fourth waves. Since we find no short-term effects, and since the models analysing longer-term effects with mh_{it-1} included display almost identical results, retirement thus appears to have a long-term negative impact on mental health in an absolute sense. To further investigate whether this interpretation is correct, Table 5 displays results from models that effectively analyse the impact of a change in retirement status between the first and second waves on the change in mental health between the first and fourth waves. The models thus incorporate any short-term effects directly. All results are almost identical, although some estimates become slightly less precise. This is not surprising since the models also pick up potential changes in mental health that occurred before the change in retirement status and any (insignificant) short-term effects. Overall, we thus conclude that the results lend further support to the interpretation that retirement has long-term negative effects on mental health.

Table 5: Longer-term effects between the first and fourth waves

	RDD	RDD	RDD	RDD	RDD	RDD
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
			10-year bandwidth			
$\widehat{r_{it-1}}$	1.43***	2.25***	1.35**	0.26**	0.42**	0.31**
	(0.52)	(0.86)	(0.68)	(0.11)	(0.19)	(0.15)
F statistic	98.75	44.95	57.66	98.75	44.95	57.66
Hausman	<0.01	<0.01	0.02	0.02	0.02	0.03
n	8,552	8,552	7,096	8,552	8,552	7,096
			3-year bandwidth			
$\widehat{r_{it-1}}$	1.23***	1.89***	0.98*	0.26***	0.40***	0.27**
	(0.44)	(0.70)	(0.55)	(0.10)	(0.15)	(0.12)
F statistic	127.25	61.61	80.54	127.25	61.61	80.54
Hausman	<0.01	<0.01	0.03	<0.01	<0.01	0.02
n	4,697	4,697	3,807	4,697	4,697	3,807

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level in parentheses. The first panel includes a quadratic age trend and the second panel includes a linear age trend.

6.1.3 Inverse probability weighting

As noted in Section 4.5, it is important to investigate whether panel attrition threatens the external validity of the findings. Table 6 displays results from exactly the same specifications as those estimated previously, with the exception that respondents are weighted by their inverse probability to remain in the panel, as predicted by the baseline characteristics discussed in Section 4.5.[19] All estimates are almost identical to the ones obtained without weights, irrespective of outcome analysed and bandwidth utilised. In fact, the coefficients become slightly larger and a little more precise, even though a few observations are lost because of missing values on the variables used to predict the probability that respondents remain in the panel. The conclusion from this exercise is thus that it is highly unlikely that selective attrition poses a threat to the study's findings.

[19] For space consideration, Table 6 only reports weighted estimates from Tables 3 and 4 with mh_{it-1} included. All results are again almost identical when excluding mh_{it-1}.

Table 6: Estimates from models using inverse probability weighting

	RDD	RDD	RDD	RDD	RDD	RDD
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Weighted estimates from the 3rd panel in Table 3						
$\widehat{r_{it-1}}$	1.82***	2.72***	2.04***	0.34***	0.51***	0.45***
	(0.51)	(0.82)	(0.66)	(0.11)	(0.18)	(0.15)
F statistic	89.86	45.57	55.42	89.91	45.60	55.50
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Weighted estimates from the 1st panel in Table 5						
$\widehat{r_{it-1}}$	1.61***	2.41***	1.73**	0.31**	0.43**	0.42**
	(0.56)	(0.87)	(0.71)	(0.12)	(0.19)	(0.17)
F statistic	89.91	45.63	55.48	89.91	45.63	55.48
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	8,486	8,486	7,038	8,486	8,486	7,038
Weighted estimates from the 3rd panel in Table 4						
$\widehat{r_{it-1}}$	1.54***	2.25***	1.50***	0.31***	0.45***	0.34***
	(0.43)	(0.65)	(0.52)	(0.09)	(0.14)	(0.12)
F statistic	116.21	63.46	76.99	116.27	63.54	77.09
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Weighted estimates from the 2nd panel in Table 5						
$\widehat{r_{it-1}}$	1.44***	2.11***	1.37**	0.32***	0.46***	0.36***
	(0.48)	(0.71)	(0.58)	(0.11)	(0.16)	(0.13)
F statistic	116.29	63.54	77.13	116.29	63.54	77.13
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	4,657	4,657	3,774	4,657	4,657	3,774

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level are in parentheses.

6.2 *Heterogeneous effects*

While all results so far indicate that retirement has a negative average impact on mental health in the longer-term perspective, this does not necessarily mean that the effect is the same among different groups of people. Table 7 shows results from models that allow the impact of retirement to differ depending on gender and educational background, as discussed in Section 5.1. Overall, there is little evidence of heterogeneous effects.[20] This indicates that the negative longer-term impact of retirement on mental health applies to men and women as well as individuals of different socio-economic backgrounds similarly.[21]

[20] For presentational purposes, the coefficient for $\widehat{r_{it-1}}$ is suppressed in this table. All models include mh_{it-1}, but results are again very similar if it is excluded.

[21] In unreported regressions, we also estimated separate models for men and women and for individuals with high and low educational backgrounds, but again found little evidence of heterogeneity.

Table 7: Heterogeneous effects

	RDD	RDD	RDD	RDD	RDD	RDD
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
10-year bandwidth						
$\widehat{r_{it-1}}*g_i$	0.17	1.64	0.18	-0.05	0.18	-0.01
	(0.49)	(1.03)	(0.63)	(0.11)	(0.21)	(0.14)
F statistics	111.22	60.42	81.07	111.22	60.25	81.23
Hausman	<0.01	<0.01	0.01	0.04	0.03	0.04
n	8,552	8,552	7,096	8,552	8,552	7,096
$\widehat{r_{it-1}}*ed_i$	0.05	-0.23	-0.04	-0.00	-0.05	-0.04
	(0.30)	(0.42)	(0.34)	(0.07)	(0.09)	(0.08)
F statistics	49.52	44.56	45.82	49.47	44.54	45.71
Hausman	<0.01	<0.01	<0.01	0.01	<0.01	0.02
$\widehat{r_{it-1}}*l\_ed_i$	-0.38	-0.08	-0.41	-0.08	-0.03	-0.08
	(0.52)	(0.74)	(0.61)	(0.11)	(0.16)	(0.14)
F statistics	94.06	53.15	69.46	94.02	53.14	69.33
Hausman	<0.01	<0.01	<0.00	0.05	0.03	0.03
n	8,489	8,489	7,041	8,489	8,489	7,041
3-year bandwidth						
$\widehat{r_{it-1}}*g_i$	0.11	1.16	-0.00	-0.06	0.12	-0.05
	(0.48)	(0.83)	(0.58)	(0.11)	(0.18)	(0.14)
F statistics	96.09	54.46	70.91	96.11	54.11	71.10
Hausman	<0.01	<0.01	0.03	0.02	0.02	0.06
n	4,697	4,697	3,807	4,697	4,697	3,807
$\widehat{r_{it-1}}*ed_i$	0.01	-0.24	-0.01	0.00	-0.05	-0.02
	(0.30)	(0.41)	(0.33)	(0.06)	(0.09)	(0.07)
F statistics	45.22	37.46	38.61	45.18	37.43	38.45
Hausman	<0.01	<0.01	0.02	<0.00	<0.01	0.02
$\widehat{r_{it-1}}*l\_ed_i$	-0.29	-0.00	-0.41	-0.08	-0.03	-0.10
	(0.50)	(0.70)	(0.58)	(0.11)	(0.15)	(0.13)
F statistics	86.71	49.61	65.06	86.68	49.58	64.96
Hausman	<0.01	<0.01	0.04	0.04	0.04	0.09
n	4,659	4,659	3,776	4,659	4,659	3,776

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level are in parentheses.

7 Conclusion

As policymakers worldwide have begun to reform state pension systems to induce higher labour force participation among the elderly, research investigating the causal impact of retirement on health and wellbeing has become increasingly important. While previous studies analysing mental health have generally found positive or no effects, it suffers from limitations. Perhaps most conspicuous is that nobody thus far has separated short- from longer-term effects of retirement in a proper RDD framework. This is a significant shortcoming since there are good theoretical reasons to believe that the short- and longer-term effects of retirement differ.

This study has aimed to remedy these issues by investigating the short- and long-term effects of retirement on mental health in ten European countries. Analysing panel data from the Survey of Health, Ageing, and Retirement in Europe, it has utilised a fuzzy RDD approach that exploits discontinuities in age-dependent retirement incentives within state pension systems. Although the results show no impact in the short run, there is strong evidence of a considerable negative longer-term effect. This effect is apparent both when analysing the Euro-D scale as well as the cut-off point for clinical depression, and survives a range of robustness tests. It applies to women and men equally and also operates independently of individuals' educational background. While this indicates that people of different socio-economic backgrounds are equally hurt by retirement, future research should nevertheless analyse in more detail how effects may differ between individuals retiring from different professions.

Also, although the study has found a negative long-term impact of retirement on mental health, it is silent on the mechanisms through which this effect operates. Policymakers and practitioners would certainly benefit from understanding these mechanisms when attempting to counter the negative long-run impact; identifying the specific mechanisms linking retirement to declining mental health in the long run remains an important topic for future research to investigate.

Nevertheless, overall, this study's findings indicate that policymakers do not face a trade-off between making state pension systems solvent and improving mental health among the elderly. Certainly, as displayed by other research, reforms affecting eligibility to state pensions may have direct negative mental health effects operating independently of retirement, at least if these reforms affect people late in their lives. With time, however, this paper's findings indicate that such reforms not only are necessary to make pension systems sustainable, but may also be an efficient way to improve mental health among the elderly.

References

Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Angrist, Joshua D. and Jörn-Steffen Pischke. 2015. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton, NJ: Princeton University Press.

Börsch-Supan, Axel, Martina Brandt, Christian Hunkler, Thorsten Kneip, Julie Korbmacher, Frederic Malter, Barbara Schaan, and Stephanie Stuck. 2013. "Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)." *International Journal of Epidemiology* 42(4):992-1001.

Börsch-Supan, Axel, Agar Brugiavini, and Enrica Croda. 2009. "The Role of Institutions and Health in European Patterns of Work and Retirement." *Journal of European Social Policy* 19(4):341-358.

Börsch-Supan, Axel and Morten Schuth 2014. "Early Retirement, Mental Health and Social Networks." Pp. 225-250 in *Discoveries in the Economics of Aging*. Chicago: University of Chicago Press.

Börsch-Supan, Axel and Christina B. Wilke 2006. "The German Public Pension System: How It Will Become an NDC System Look-Alike." Pp. 573-610 in *Pension Reform: Issues and Prospects for Non-Financial Defined Contribution (NDC) Schemes*, edited by Robert Holzmann and Edward Palmer. Washington, DC: World Bank.

Behncke, Stefanie. 2012. "Does Retirement Trigger Ill Health?" *Health Economics* 21:282-300.

Blake, Hélène and Clémentine Garrouste. 2012. "Collateral Effects of a Pension Reform in France." Working Paper No. 2012-25, Paris School of Economics, Paris.

Bonsang, Eric, Stéphane Adam, and Sergio Perelman. 2012. "Does retirement affect cognitive functioning?" *Journal of Health Economics* 31(3):490-501.

Brickman, Philip and Donald Campbell 1971. "Hedonic Relativism and Planning the Good Society." Pp. 287-302 in *Adaptation-level Theory: A Symposium*. New York: Academic Press.

Charles, Kerwin K. 2004. "Is Retirement Depressing? Labor Force Inactivity and Psychological Well-Being in Later Life." *Research in Labor Economics* 23:269-299.

Coe, Norma B. and Gema Zamarro. 2011. "Retirement Effects on Health in Europe." *Journal of Health Economics* 30(1):77-86.

Dave, Dhaval, Inas Rashad, and Jasmina Spasojevic. 2006. "The Effects of Retirement on Physical and Mental Health Outcomes." NBER Working Paper No. 12123, National Bureau of Economic Research, Cambridge, MA.

Dave, Dhaval, Inas Rashad, and Jasmina Spasojevic. 2008. "The Effects of Retirement on Physical and Mental Health Outcomes." *Southern Economic Journal* 75(2):497-523.

de Grip, Andries, Maarten Lindeboom, and Raymond Montizaan. 2012. "Shattered Dreams: The Effects of Changing the Pension System Late in the Game." *Economic Journal* 122(559):1-25.

De Luca, Giuseppe. 2009. "Determinants of the Attrition Process in the First Two Waves of SHARE." Unpublished Paper, The Royal Society, London.

d'Hombres, Beatrice, Lorenzon Rocco, Marc Suhrcke, and Martin McKee. 2010. “Does Social Capital Determine Health? Evidence from Eight Transition Countries.” *Health Economics* 19(1):56-74.

Eduardo, Fé. and Bruce Hollingsworth. 2012. “Estimating the Effect of Retirement on Mental Health via Panel Discontinuity Designs.” Working Paper, Munich Personal RePEc Archive.

Eibich, Peter. 2015. “Understanding the Effect of Retirement on Health: Mechanisms and Heterogeneity.” *Journal of Health Economics* 43:1-12.

Finnie, Ross and Bryon G. Spencer. 2013. “How do the Level and Composition of Income Change after Retirement? Evidence from the LAD.” Working Paper No. 114, Canadian Labour Market and Skills Researcher Network, Vancouver.

Fletcher, Jason M. 2014. “Late Life Transitions and Social Networks: The Case of Retirement.” *Economics Letters* 125:459-462.

Folland, Sherman. 2008. “An Economic Model of Social Capital and Health.” *Health Economics, Policy and Law* 3:333-348.

Fonseca, Raquel, Arie Kapteyn, Jinkook Lee, Gema Zamarro, and Kevin Feeney. 2014. “A Longitudinal Study of Well-Being of Older Europeans: Does Retirement Matter?” *Journal of Population Ageing* 7:21-41.

Gelman, Andrew and Guido Imbens. 2014. “Why High-order Polynomials Should not be Used in Regression Discontinuity Designs.” NBER Working Paper No. 20405, National Bureau of Economic Research, Cambridge, MA.

Godard, Mathilde. 2016. “Gaining Weight Through Retirement? Results from the SHARE Survey.” *Journal of Health Economics* 45:27-46.

Grossman, Michael 2000. “The Human Capital Model.” Pp. 347-408 in *Handbook of Health Economics*, edited by Joseph P Newhouse and Anthony J Culyer. New York: Elsevier.

Gruber, Jonathan and David A. Wise, eds. 1999. *Social Security and Retirement around the World*. Chicago, IL: University of Chicago Press.

Gruber, Jonathan and David A. Wise, eds. 2004. *Social Security Programs and Retirement around the World: Micro-Estimation*. Chicago, IL: University of Chicago Press.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design.” *Econometrica* 69(1):201-209.

Haider, Steven J. and Melvin Stephens. 2007. “Is There a Retirement-Consumption Puzzle? Evidence Using Subjective Retirement Expectations.” *Review of Economics and Statistics* 89(2):247-264.

Hurd, Michael, Pierre-Carl Michaud, and Susann Rohwedder. 2012. “The Displacement Effect of Public Pensions on the Accumulation of Financial Assets.” *Fiscal Studies* 33(1):107-128.

Imbens, Guido W. and Joshua D. Angrist. 1994. “Identification and Estimation of Local Average Treatment Effects.” *Econometrica* 62(2):467-475.

Imbens, Guido W. and Jeffrey W. Wooldridge. 2009. “Recent Developments in the Econometrics of Program Evaluation.” *Journal of Economic Literature* 47:5-86.

Johnston, David W. and Wang-Sheng Lee. 2009. “Retiring to the Good Life? The Short-Term Effects of Retirement on Health.” *Economics Letters* 103(1):8-11.

Jokela, M, JE Ferrie, D Gimeno, T Chandola, MJ Shipley, J Head, J Vahtera, H Westerlund, MG Marmot, and M Kivimäki. 2010. "From Midlife to Early Old Age: Health Trajectories Associated with Retirement." *Epidemiology* 21(3):284-290.

Latif, Ehsan. 2013. "The Impact of Retirement on Mental Health in Canada." *Journal of Mental Health Policy and Economics* 16(1):36-46.

Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48:281-355.

Lindeboom, Maarten, France Portrait, and Gerard J. van den Berg. 2002. "An Econometric Analysis of the Mental-Health Effects of Major Events in the Life of Older Individuals." *Health Economics* 11:505-520.

Litwin, Howard. 2007. "Does Early Retirement Lead to Longer Life?" *Ageing & Society* 27:739-754.

Mazzonna, Fabrizio and Franco Peracchi. 2012. "Ageing, Cognitive Abilities and Retirement." *European Economic Review* 56(4):691-710.

Mazzonna, Fabrizio and Franco Peracchi. 2014. "Unhealthy Retirement?" EIEF Working Paper 09/14, Einaudi Institute for Economics and Finance, Rome.

Mein, G, P Martikainen, H Hemingway, S Stansfeld, and M Marmot. 2003. "Is Retirement Good or Bad for Mental and Physical Health Functioning? Whitehall II Longitudinal Study of Civil Servants." *Journal of Epidemiology & Community Health* 57:46-49.

Moffit, Robert, John Fitzgerald, and Peter Gottschalk. 1999. "Sample Attrition in Panel Data: The Role of Selection on Observables." *Annales d'Économie et de Statistique* 55-56:129-152.

Montizaan, Raymond, Frank Cörvers, and Andries de Grip. 2010. "The Effects of Pension Rights and Retirement Age on Training Participation: Evidence From a Natural Experiment." *Labour Economics* 17(1):240-247.

Moreau, Nicolas and Elena Stancanelli. 2013. "Household Consumption at Retirement: A Regression Discontinuity Study on French Data." IZA Discussion Paper No. 7709, Institute for the Study of Labor (IZA), Bonn.

Mosca, Irene and Alan Barrett. 2014. "The Impact of Voluntary and Involuntary Retirement on Mental Health: Evidence from Older Irish Adults." Discussion Paper No. 8723, Institute for the Study of Labor, Bonn.

Neuman, Kevin. 2008. "Quit Your Job and Get Healthier? The Effect of Retirement on Health." *Journal of Labor Research* 29:177-201.

Nickell, Stephen J. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49(6):1417-1426.

Oksanen, Tuula, Jussi Vahtera, Hugo Westerlund, Jaana Pentti, Noora Sjösten, Marianna Virtanen, Ichiro Kawachi, and Mika Kivimäki. 2011. "Is Retirement Beneficial for Mental Health? Antidepressant Use Before and After Retirement." *Epidemiology* 22(4):553-559.

Olea, José L. M. and Carolin E. Pflueger. 2013. "A Robust Test for Weak Instruments." *Journal of Business & Economic Statistics* 31(3):358-369.

Prince, M J., F Reischies, A T. Beekman, R Fuhrer, C Jonker, S L. Kivela, B A. Lawlor, A Lobo, H Magnusson, M Fichter, H van Oyen, M Roelands, I Skoog, C Turrina, and J R. Copeland. 1999. "Development of the EURO-D Scale - A

European Union Initiative to Compare Symptoms of Depression in 14 European Centres." *British Journal of Psychiatry* 174(4):339-345.

Rocco, Lorenzo, Elena Fumagalli, and Marc Suhrcke. 2014. "From Social Capital to Health - and Back." *Health Economics* 23(5):586-605.

Rohwedder, Susann and Robert J. Willis. 2010. "Mental Retirement." *Journal of Economic Perspectives* 24(1):119-138.

Ronconi, Lucas, Timothy T. Brown, and Richard M. Scheffler. 2012. "Social Capital and Self-Rated Health in Argentina." *Health Economics* 21(2):201-208.

SSA. 2006. "Social Security Programs Throughout the World: Europe, 2006." SSA Publication No. 13-11801, Social Security Administration, Washington, DC.

Stancanelli, Elena. 2014. "Divorcing Upon Retirement: A Regression Discontinuity Study." Discussion Paper No. 8117, Institute for the Study of Labor, Bonn.

Stancanelli, Elena and Arthur Van Soest. 2012. "Joint Leisure Before and After Retirement: A Double Regression Discontinuity Approach." IZA Discussion Paper No. 6698, Institute for the Study of Labor (IZA), Bonn.

Westerlund, Hugo, Mika Kivimäki, Archana Singh-Manoux, Maria Melchoir, Jane E. Ferrie, Jaana Pentti, Markus Jokela, Constanze Leineweber, Marcel Goldberg, Marie Zins, and Jussi Vahtera. 2009. "Self-Rated Health Before and After Retirement in France (GAZEL): A Cohort Study." *The Lancet* 374:1889-1896.

Vo, Kha, Peta M. Forder, Meredith Tavener, Bryan Rodgers, Emily Banks, Adrian Bauman, and Julie E. Byles. 2015. "Retirement, Age, Gender and Mental Health: Findings from the 45 and Up Study." *Aging & Mental Health* 19(7):647-657.

Appendix

Table A1: State pension ages

Country	*Men*	*Women*
Austria	65	60
Belgium	65	60-63
Denmark	65-67	65-67
France	60	60
Germany	65	60-62
Italy	60-65	55-60
Netherlands	65	65
Spain	65	65
Sweden	65	65
Switzerland	65	62-63

Note: The state pension ages are based on those provided by Mazzonna and Perrachi (2012), with slight adjustments based on data from other sources (see Börsch-Supan and Wilke 2006; SSA 2006).

Table A2: Associations and effects between the second and fourth waves

Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS	FE-OLS
r_{it}	-0.14**	-0.15**	-0.16**	-0.03**	-0.03**	-0.04**
Waves 2-4	(0.06)	(0.06)	(0.07)	(0.01)	(0.01)	(0.01)
	RDD	RDD	RDD	RDD	RDD	RDD
10-year bandwidth						
$\widehat{r_{it}}$	0.07	0.10	0.09	0.09	0.11	0.10
Waves 2-4	(0.33)	(0.45)	(0.40)	(0.08)	(0.10)	(0.09)
F statistics	181.87	98.39	115.90	181.85	98.37	115.80
Hausman	0.69	0.73	0.72	0.16	0.20	0.21
n	8,546	8,546	7,226	8,546	8,546	7,226
3-year bandwidth						
$\widehat{r_{it}}$	-0.43	-0.60	-0.51	0.02	0.02	0.01
Waves 2-4	(0.28)	(0.39)	(0.33)	(0.06)	(0.09)	(0.07)
F statistics	220.12	117.51	144.90	220.42	117.58	145.36
Hausman	0.22	0.18	0.18	0.55	0.67	0.78
n	4,698	4,698	3,954	4,698	4,698	3,954

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level are in parentheses. The model specifications correspond to those in the first panels in Table 2, Table 3, and Table 4, but the RDD estimates also include mh_{it-1}. The number of individuals differs slightly compared to the models in the paper. This is because of a few instances of missing data, and, in the case of the models using the third retirement definition, because more individuals reported themselves to be either working or retired between the second and fourth waves.

Table A3: Combining a quadratic age trend with the three-year bandwidth

	RDD	RDD	RDD	RDD	RDD	RDD
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
$\widehat{r_{it-1}}$	2.46***	3.51***	2.53***	0.46***	0.66**	0.57***
	(0.77)	(1.20)	(0.95)	(0.17)	(0.26)	(0.21)
F statistics	52.29	27.43	34.70	52.29	27.43	34.70
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
$\widehat{r_{it-1}}$	2.16***	3.09***	2.07**	0.37**	0.53**	0.42**
	(0.69)	(1.07)	(0.84)	(0.15)	(0.22)	(0.18)
mh_{it-1}	YES	YES	YES	YES	YES	YES
F statistics	52.20	27.32	34.47	52.37	27.36	34.76
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	4,697	4,697	3,807	4,697	4,697	3,807

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level are in parentheses.