Retirement Blues

Gabriel Heller-Sahlgren

DOI/Link: https://doi.org/10.1016/j.jhealeco.2017.03.007

Reference: Heller Sahlgren, Gabriel (2017). "Retirement Blues". *Journal of Health Economics*, 54(July), 66–78.

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†,‡,§

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

Forthcoming in *Journal of Health Economics*

Abstract

This paper analyses the short- and longer-term effects of retirement on mental health in ten European countries. It exploits thresholds created by state pension ages in an individual-fixed effects instrumental-variable set-up, borrowing intuitions from the regression-discontinuity design literature, 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 people of different educational and occupational backgrounds similarly. Overall, the 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, Instrumental Variables

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).

† Department of Social Policy and Centre for Economic Performance, London School of Economics, Houghton Street, WC2A 2AE, London, UK

‡ Research Institute of Industrial Economics, Stockholm, Sweden.

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

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 has 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. 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 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 covering ten European countries, it exploits thresholds created by state pension ages in an individual-fixed effects instrumental-variable (FE-IV) set-up, borrowing intuitions from the regression-discontinuity design (RDD) literature, 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 age effects are held constant in a flexible way. Previous research has utilised similar strategies – most often finding positive or zero short-term effects of retirement on mental health – but tends to ignore common pitfalls that threaten their validity and the possibility that the effects of retirement are not instant. To the best of our knowledge, this paper is the first to analyse both short- and longer-term, lagged effects of retirement on mental health in a set-up exploiting age-dependent discontinuities to predict retirement.

The results display no short-term impact of retirement on mental health, but strong negative effects that become apparent a couple of years following the event. The longer-term effect does not differ between men and women, or between people with different educational and occupational backgrounds. It also survives a battery of robustness tests, which include analyses of narrower age windows and using inverse probability weighting to deal with panel attrition. The differences compared with 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 prior to retirement as some research suggests – at least if these reforms postpone eligibility rules late in people's lives – this paper indicates that they may pay a mental health dividend after some time by delaying the negative longer-term impact of retirement per se.

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. 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' 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 create 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). While retirement could potentially induce couples to spend more time together, it may also increase the probability of divorce (see Stancanelli 2014). A similarly equivocal story applies to stress. While retirement may decrease work-related stress, it is in itself a life event that can be very stressful. And retirement is often accompanied by a decrease in income and consumption (e.g. Finnie and Spencer 2013), which may affect mental health negatively directly and via increased stress. In fact, such ambiguous stories apply to most theoretical mechanisms potentially linking retirement to mental health.[1]

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 lower/higher investments do 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 (e.g. Atchley 1976). If so, one may also expect people's mental health to improve – or at least not deteriorate – 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 subside 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, it is clearly important to take into account that short- and longer-term effects of retirement may differ, although it is difficult to predict how and in what ways.

3 Previous literature

The impact of retirement on health 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, since the act of retirement is not random, these studies cannot tease out its causal effects on mental health.

Improving the methodology, some researchers have employed IV models using eligibility ages at which state pension benefits can be drawn to predict retirement. The idea is that reaching these eligibility ages gives rise to significant

[1] Of course, there is a plethora of other theoretical mechanisms through which retirement could potentially affect mental health in different ways, such as changes in health insurance status, although these are also often related to the rational choice perspective.

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 as instruments for retirement in IV or fuzzy RDD frameworks. Another approach has been to utilise pension reforms in difference-in-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 (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, this research also suffers from some limitations. 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 immediate and lagged 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. Bertoni, Brunello, and Mazzarella 2016; de Grip, Lindeboom, and Montizaan 2012; Montizaan, Cörvers, and de Grip 2010). This violates the assumption that the reforms used for identification affect mental health solely through retirement, thereby casting doubt on the studies' internal validity.

Another potential problem in most previous research is that it uses instruments constructed from both regular and early retirement ages. This

[2] Using a regular IV set-up, Mazzonna and Peracchi (2014) also find weak evidence in robustness tests that longer time spent in retirement increases the likelihood of depression on average. However, the instrument used is the distance of respondents' actual age from the relevant state pension thresholds. As discussed in Section 5.1, this variable should normally be included as a control to allow the age trend to differ on both sides of the threshold and thus decrease the risk that the binary retirement indicator, which in their study is only used to estimate the immediate impact of retirement, picks up non-linear effects of age. This is not a trivial concern, especially since the authors use a relatively broad age window and control for a linear age trend only.

neglects potential 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 of the population. While discontinuities arising at the state pension age should help ameliorate measurement error in retirement status as observed in survey data – since the discontinuities are shaped by institutional rules and are therefore uncorrelated with potential measurement error – it is thus 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 events have the same effects; as countries are in the process of increasing regular state pension ages specifically, it is important to disentangle the separate impact of retirement at these ages.

Another general problem with previous IV studies based on RDD intuitions is that they often ignore common pitfalls associated with such designs. For example, they do not present results in which the impact of age is allowed to differ on each side of the discontinuities used as instruments (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 discontinuities (Angrist and Pischke 2015).[3] Instead, they choose a rather wide range of data, without sufficiently exploring non-linear effects of age or the results' sensitivity to the specific data range.[4]

Overall, therefore, while previous research most often finds no or positive effects of retirement on mental health, it suffers from some limitations. Perhaps most important is that previous studies do not generally separate short- from longer-term effects of retirement. This study aims to remedy the shortcomings highlighted and provide a more rigorous evaluation of the impact of retirement on mental health in Europe. Section 5 discusses this strategy in detail.

[3] Much previous multi-country research has also 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.
[4] The exception is Eibich's (2015) paper, which finds a short-term positive impact of retirement on mental health in Germany.

4 Data

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.[5] Analysing panel data spanning over several SHARE waves allows us to investigate both the short- and longer-term effects of retirement on mental health. Table 1 displays the timeline for interviews over the period analysed.

Table 1: Timeline for interviews

	Mean	Min	Max
Wave 1	October 2004	March 2004	December 2005
Wave 2	January 2007	September 2006	October 2007
Wave 4	May 2011	February 2011	March 2012
Time passed (Waves 1-2)	2 years, 3 months	11 months	3 years, 4 months
Time passed (Waves 2-4)	4 years, 4 months	3 years, 5 months	5 years, 6 months
Time passed (Waves 1-4)	6 years, 7 months	5 years, 6 months	7 years, 10 months

Note: the timeline refers to the main sample within the ten-year age window over and below the relevant state pension age. It is essentially identical for the sample within the three-year age window.

As Table 1 shows, the second interview is held on average two years and three months after the first interview, whereas the fourth interview is held on average four years and four months after the second interview, and six years and seven months after the first interview.[6] In our main analysis, we study the impact of a change in retirement status over the first and second waves on the change in mental health over the second and the fourth waves. Since the fourth interview takes place several years after the change in retirement status took place

[5] 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).

[6] In unreported robustness tests, we restricted the sample to (1) only individuals with a time span between the first and second waves of minimum two years and maximum three years, and (2) only individuals with a time span between the second and fourth waves of minimum four years and maximum five years. Overall, the main results were very similar.

between the first and the second waves, we believe our set-up captures a lagged, longer-term effect well. We also seek to compare and contrast this impact with the effect of a change in retirement status between the first and second waves on changes in mental health over the same period, which captures the immediate, short-term impact. The methodology utilised to tease out the different effects is discussed in detail in Section 5.

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.[7] As discussed in Section 5, the study exploits discontinuities in retirement that occur at the state pension age to obtain exogenous variation in retirement behaviour. The state pension ages for the ten European countries utilised in the study – which take into account pension reforms in recent decades and thus vary across cohorts within countries – 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 set-up outlined in Section 5 – were between 60 and 65.[8] This gives an age window of about ten years over and under the lowest and highest thresholds utilised, calculated at the second wave, which is in line with Moreau and Stancanelli (2013) and Stancanelli and Van Soest (2012).

The working sample thus consists of maximum 8,566 individuals across the ten countries, all observed three times, with the exact sample size depending on the restrictions discussed in Section 4.2. Table 2 provides summary statistics for the working sample as well as differences in means between individuals assigned to the treatment group (those who crossed the state pension age between the first and second interviews), and individuals assigned to the control group (those

[7] Since the third wave was a special survey (SHARELIFE) that did not enquire respondents about their mental health, it cannot be used in this paper.

[8] The only 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 age window is calculated from age 65 also for the Danish sample. However, all results are essentially identical if the Danish age window is instead extended by two years on the right-hand side, which only increases the sample by about 30 respondents. Results are also robust to extending the right-hand side of the age window to 67 for all countries.

who did not cross the state pension age between the first and second interviews).[9] In robustness tests, we also utilise a narrower age window of three years around the lowest and highest state pension ages, which means that people who were 57-68 years old at the second wave are included. This decreases the sample size to maximum 4,704 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 are engaged in paid work, may respond to retirement incentives at the state pension age. This is because they may still be able to access age-dependent benefits at that point, which could trigger permanent withdrawal from the labour market and make them regard themselves to be retired.[10] Furthermore, for policy purposes, the mental health impact of retirement is relevant regardless from which category respondents officially retire.

However, to ensure that the above definition of retirement does not drive the results, we also employ an alternative definition based 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 do not also report having done any paid work in the past four weeks. Respondents in these categories who report having

[9] The FE-IV design described in Section 5 analyses compliers only: respondents assigned to the treatment group who retired between waves 1 and 2 and respondents assigned to the control group who did not retire in that period. In robustness tests, we simultaneously analyse the lagged impact of retiring between waves 1 and 2 and the direct effect of retiring between waves 2 and 4, both in the full sample and when excluding individuals who were retired and/or above the state pension age at the first interview. As Table A5 shows, the results are very similar in these analyses.

[10] This argument is supported by the fact that the F statistic for the instrument is the largest when utilising this definition, indicating that respondents who are not engaged in paid work prior to retirement do respond to the incentives in the regular state pension system.

done paid work are still included as non-retirees. Finally, in the third definition, the sample is simply restricted to individuals who claim to be retired or to be working. The latter category includes employed/self-employed respondents and other non-retired individuals who report having done paid work in the past four weeks.

4.3 Mental health

Following previous research in the field, our 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, we also analyse the likelihood that respondents have reached 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 behaviour, the only necessary covariates to ensure a causal interpretation of the estimates are flexible controls for respondents' age. 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. It is also possible to interact certain background variables with the retirement indicator in order to investigate potential heterogeneous effects. These issues are discussed more formally in Section 5.1.

4.5 Attrition

It is important to consider potential selection problems due to non-random panel attrition, which is often ignored by studies exploiting panel surveys to evaluate various outcomes of retirement (e.g. Bonsang, Adam, and Perelman 2012; Eibich 2015). Like in most panel surveys, attrition in SHARE is substantial: about 50 per cent of people interviewed in the first 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. 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 as well as their retirement behaviour and mental health. This should at least mitigate the attrition problem.

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. 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 achievement.[11] 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 weight in robustness regressions analysing the impact of retirement on mental health.[12] If attrition poses a serious problem for the study, 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 impossible to demonstrate conclusively an absence of attrition bias, the combination of individual-level fixed effects and inverse probability weighting leave little room for any remaining bias.

---

[11] 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.

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

Table 2: Descriptive statistics

Variable	*Working sample*				*Treatment group*	*Control group*
Wave 1	Mean	SD	Min	Max	Mean	Mean
Age	60.63	6.65	47.23	74.46	62.53	60.42
Retired (1)	0.44	0.50	0	1	0.53	0.43
Retired (2)	0.59	0.49	0	1	0.70	0.57
Retired (3)	0.52	0.50	0	1	0.64	0.51
Euro-D	2.16	2.10	0	12	1.98	2.19
Depression	0.23	0.42	0	1	0.20	0.23
Wave 2	Mean	SD	Min	Max	Mean	Mean
Age	62.95	6.65	50	75.92	64.95	62.73
Retired (1)	0.51	0.50	0	1	0.81	0.48
Retired (2)	0.65	0.48	0	1	0.91	0.62
Retired (3)	0.60	0.49	0	1	0.90	0.57
Euro-D	2.09	2.06	0	11	1.91	2.11
Depression	0.21	0.41	0	1	0.19	0.21
Wave 4	Mean	SD	Min	Max	Mean	Mean
Age	67.24	6.66	54.08	80.77	69.20	67.02
Retired (1)	0.65	0.48	0	1	0.87	0.62
Retired (2)	0.78	0.42	0	1	0.96	0.75
Retired (3)	0.74	0.44	0	1	0.96	0.72
Euro-D	2.27	2.11	0	12	2.25	2.27
Depression	0.24	0.43	0	1	0.25	0.24
n	8,566				849	7,717

Note: Respondents assigned to the treatment group crossed the state pension age between waves 1 and 2, while respondents assigned to the control group did not cross the state pension age in that period. Retired (1): respondents who report to be retired or give a retirement date preceding the interview date. Retired (2): retired (1) plus homemakers, the permanently ill or disabled, and those engaged in other activities (as long as they do not do paid work). Retired (3): retired (1) but excluding all non-workers in the non-retired category (for the last definition: n = 7,113 in the total sample, n = 692 assigned to the treatment group and n = 6,421 assigned to the control group).

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:

$$mh_i = \alpha + \beta_1 r_i + \beta_2 x_i + \varepsilon_i \tag{1}$$

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 mean that $Cov(r_i, \varepsilon_i|x_i) \neq 0$ (Angrist and Pischke 2009). As noted in Section 3, the critical assumption is not likely to hold. To ensure a causal interpretation of the paper's findings, it is thus key to obtain exogenous variation in retirement behaviour.

To do so, the paper proposes an individual-fixed effects IV design, with intuitions borrowed from the RDD literature. The design is 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 continuous variable determining the discontinuities (Angrist and Piscke 2009; Imbens and Wooldridge 2009). As noted in Section 3, this idea has been exploited in previous research, but not in a way that allows researchers to distinguish between short- and longer-term effects. 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.

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 $age_i < sp_i$. The strategy is dependent on the ability to separate smooth effects of age_i from the impact at $\overline{sp}_i$, which serves as instrument for r_i. The paper follows previous studies and assumes a quadratic age trend in the baseline estimates. Recent research shows that estimates including higher-order polynomials in similar designs may be misleading and that it is therefore preferable to reduce the range of data around the threshold (Gelman and Imbens 2014).

The principal difference between our set-up and a regular fuzzy RDD is the inclusion of individual-level fixed effects, which means that we focus on the variation within individuals across time rather than the variation between individuals. This also means that the identification assumptions are different. While a traditional fuzzy RDD would hinge on the assumption that people who are on different sides of, but close to, the state pension age only differ in terms of the probability of being retired, once controlling flexibly for the direct impact of age, our individual fixed-effects IV estimator hinges on the assumption that merely crossing the threshold serving as instrument does not impact an individual's mental health around the threshold apart from via retirement. We note that similar designs, combining individual-fixed effects with intuitions from the RDD literature, have been utilised in previous research in this and other fields (e.g. Eibich 2015; Lemieux and Milligan 2008; Petterson-Lidbom 2012).

We thus model the lagged, longer-term impact of retirement on mental health using a 2SLS model. Unlike most previous research, we allow the impact of age to differ on both sides of the eligibility threshold used as instrument. To do so, we centre age and its polynomial – by subtracting the state pension age from the respondent's age – which ensures that the coefficient of the retirement indicator still measures the jump in the dependent variable at the threshold serving as instrument (see Angrist and Pischke 2009, 2015). We also take into account cross-country differences in age effects, below and above the threshold.[13] The estimation then reads:

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

[13] Thus, a key difference between this paper's set-up and the one used by Coe and Zamarro (2011) is the fact that we exploit the longitudinal aspect of SHARE, allowing us to include individual-fixed effects and to separate short- from longer-term effects of retirement. Other differences include the fact that they analyse only men, do not take into account that the effect of age may differ across countries, ignore the fact that the impact of age may differ below and above the threshold serving as instrument, use both early and regular state pension ages as instruments, and control for potentially endogenous variables, such as marital status and income.

$$mh_{it} = \alpha + \beta_1 \widehat{r_{it-1}} + \beta_2 \widetilde{age}_{it-1} + \beta_3 \widetilde{age}_{it-1}{}^2 + \beta_4 \overline{sp}_{it-1}(\widetilde{age}_{it-1})$$
$$+\beta_5 \overline{sp}_{it-1}(\widetilde{age}_{it-1}{}^2) + \gamma_c(\widetilde{age}_{it-1}) + \gamma_c(\widetilde{age}_{it-1}{}^2) + \gamma_c[\overline{sp}_{it-1}(\widetilde{age}_{it-1})]$$
$$+\gamma_c[\overline{sp}_{it-1}(\widetilde{age}_{it-1}{}^2)] + \delta_i + \mu_t + \varrho_t + \varepsilon_{it} \quad (3)$$

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; $\widetilde{age}_{it-1}$ and $\widetilde{age}_{it-1}{}^2$ denote $(age_{it-1} - sp_{it-1})$ and $(age_{it-1} - sp_{it-1})^2$ respectively; δ_i denotes individual-fixed effects; and μ_t and ϱ_t represent separate year- and month-fixed effects respectively. Including interactions between the age variables and γ_c, which denote country dummies, means that the effect of age is allowed to differ across countries.[14]

The model thus effectively analyses the impact of a change in retirement status between the first and second waves (on average spanning two years and three months) on the change in mental health between the second and fourth waves (on average spanning four years and four months). 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, the model focuses only on the change in mental health that occurs after the shift in retirement status has taken place. Of course, the set-up thus also risks ignoring potential immediate effects of a shift in retirement status between the first and the second waves. Yet if we also control for lagged mental health, any difference between compliers in the treatment and control groups due to differential mental health trends between the first and the second waves is ignored. Also, in robustness tests, we simply analyse the change in mental health between the first and fourth waves, thus taking into account any short-term impact directly.

An important rationale behind the study is also to investigate whether the short-term effect of retirement differs from the longer-term impact. Thus, in models analysing the short-term effect of retirement, we estimate equations (2) and (3) but with all variables measured at t instead of $t-1$.[15] Effectively, we then analyse the impact of retiring between the first and the second waves on the

[14] The results are very similar if we also include separate indicators for $\widetilde{age}_{it}$ and $\widetilde{age}_{it}{}^2$ and their interactions with γ_c when analysing longer-term effects. This is expected if the strategy induces random variation in retirement behaviour.

[15] In estimations analysing 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 can be compared among the same individuals. However, results for the direct association and impact over the second and fourth waves are presented in Tables A2, A3, and A5 in the Appendix.

change in mental health over the same period. In this way, by altering the observation window across two and four waves respectively, it is possible to compare and contrast the short- and longer-term effects of retirement on mental health using the same spike in retirement that occurs at the state pension age as instrument, while at the same time holding constant differential linear and non-linear age trends under and above the thresholds.[16]

We also investigate potential heterogeneous effects depending on gender, education level, as well as physical and psychological occupational burden. We create one dummy that takes the value 1 for women and 0 for men, and one dummy that takes the value 1 for education levels equal to or below lower-secondary school and 0 otherwise. We then utilise Kroll's (2011) indexes measuring the physical burden (OPB) and psychosocial burden (OSB) of different occupations, calculated from their ISCO-88 codes, which we link to the SHARE dataset.[17] The indexes range from 1 to 10, with higher values indicating higher occupational strain. We create two dummies indicating physical and psychosocial occupational strain above the value of 5. By interacting these variables with the retirement indicator, and in turn instrumenting the product with the interaction between the variables and the state pension threshold, it is possible to analyse potential heterogeneous effects within the above framework.

5.2 Assumptions

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 choose not to retire because they reach the state pension age (Imbens and Angrist 1994; Hahn, Todd, and Van der Klaauw 2001). Previous research shows that age-dependent

[16] Our modelling approach thus differs from Mazzonna and Perrachi's (2014) attempt to separate shorter- and longer-term effects of (early and regular) retirement on general health. Whereas these authors effectively use the interaction between $\overline{sp}_i$ and $(age_{it} - sp_{it})$ as instrument for time spent in retirement, we include the lagged version of this variable and its polynomial as well as their interactions with country dummies as controls to ensure that differential age trends above the thresholds are held constant.

[17] We use the ISCO-88 code for respondents' last job in the first instance. For some respondents, the code is only available for their second or first job. To maximise the number of observations, we thus use the code for the second job in the second instance and the code for the first job in the third instance.

financial incentives in public pension systems are important for retirement behaviour (e.g. Börsch-Supan, Brugiavini, and Croda 2009; Gruber and Wise 1999, 2004; Hurd, Michaud, and Rohwedder 2012). It is also unlikely that reaching the state pension age would induce some people not to retire, which would violate the monotonicity requirement. And since we hold constant individual-fixed effects and control very flexible for the continuous variable from which the binary instrument is constructed – in ways similar to RDD strategies – we believe the instrument satisfies the validity requirement in our set-up.[18]

Furthermore, having access to panel data also means that it is possible to investigate more thoroughly whether our set-up produces random variation in retirement. Indeed, if this is the case, the retirement coefficient should not differ much when including lagged mental health as independent variable, although precision may increase (Lee and Lemieux 2010). Including lagged mental health also makes it possible to test and control for potential mean reversion, which may affect the findings (e.g. Angrist and Pischke 2009).[19]

Finally, we note that the age window utilised may impact the findings. Choosing the window involves a trade-off between consistency and efficiency: a smaller window decreases the likelihood of bias, but fewer observations simultaneously increase the variance (Lee and Lemieux 2010). We thus restrict the age window in robustness tests to ensure that the results do not hinge on the one utilised in the main set-up.

6 Results

Table 3 displays estimates from OLS models that include individual-fixed effects, but do not correct for endogeneity in retirement behaviour, using the ten-year age window calculated at the second wave. The coefficients are insignificant for both measures analysed. This holds true when analysing the short-term association between the first and the second waves in the first panel, and when analysing the longer-term association in the second panel. Meanwhile, the results

[18] As always, the estimates capture 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 retire because they reach the relevant state pension eligibility age.

[19] Because individual-fixed effects are included, lagged mental health is mechanically correlated with ε_{it} (Nickell 1981), but this is not a problem for this study's purposes, as long as lagged mental health is orthogonal to the instrument – which the exercise is supposed to test.

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. The OLS estimates thus support previous findings of zero or positive effects of retirement on mental health.

Table 3: 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	-0.06	0.00	0.00	0.00
	(0.07)	(0.07)	(0.08)	(0.01)	(0.02)	(0.02)
Long-term associations						
r_{it-1}	-0.03	0.02	-0.11	0.00	0.01	-0.01
	(0.07)	(0.07)	(0.08)	(0.02)	(0.02)	(0.02)
n	8,566	8,566	7,113	8,566	8,566	7,113

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.

Turning to our main research strategy to deal with endogeneity, Table 4 displays the estimates from the FE-IV model in equations (2) and (3). 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 coefficients indicate that reaching the state pension age threshold increases the likelihood of retirement by 10-17 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.

The second-stage results in the first panel, in turn, display no evidence that a change in retirement status over the first and second waves, measured by the coefficient of $\widehat{r_{it}}$, has any short-term effects on changes in mental health over the same period. As displayed in Tables A3 and A5, this also holds true in the FE-IV analyses of the direct impact over the second and fourth waves, in sharp contrast to the OLS estimates in Table A2.

Table 4: Estimates from FE-IV models

	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Short-term effects						
$\widehat{r_{it}}$	-0.16	-0.26	-0.44	-0.04	-0.06	-0.09
Second stage	(0.49)	(0.78)	(0.65)	(0.11)	(0.18)	(0.15)
$\overline{sp}_{it}$	0.17***	0.10***	0.13***	0.17***	0.10***	0.13***
First stage	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
F statistic	88.56	38.32	48.62	88.56	38.32	48.62
Hausman test	0.81	0.77	0.56	0.72	0.77	0.50
Longer-term effects						
Excluding lagged mental health						
$\widehat{r_{it-1}}$	1.70***	2.74***	1.90**	0.31**	0.50**	0.42**
Second stage	(0.55)	(0.95)	(0.74)	(0.12)	(0.21)	(0.17)
$\overline{sp}_{it-1}$	0.17***	0.10***	0.13***	0.17***	0.10***	0.13***
First stage	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
F statistic	88.46	38.25	48.73	88.46	38.25	48.73
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Including lagged mental health						
$\widehat{r_{it-1}}$	1.61***	2.60***	1.69**	0.29***	0.47**	0.37**
Second stage	(0.50)	(0.87)	(0.67)	(0.11)	(0.18)	(0.15)
$\overline{sp}_{it-1}$	0.17***	0.10***	0.13***	0.17***	0.10***	0.13***
First stage	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
F statistic	88.42	38.23	48.67	88.46	38.24	48.75
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	8,566	8,566	7,113	8,566	8,566	7,113

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 (2) and (3).

However, the results in the second panel uniformly indicate that a change in retirement status over the first and second waves, measured by the coefficient of $\widehat{r_{it-1}}$, has a large negative longer-term impact on changes in mental health between the second and fourth waves. The coefficients display that retirement increases the overall Euro-D score by 1.55–2.44 points, relative to the control group, depending on which retirement definition that is utilised. Based on the descriptive statistics for the fourth wave in Table 2, this corresponds to 0.81–1.30 standard deviations. Meanwhile, retirement increases the likelihood of remaining or becoming clinically depressed, defined as scoring 4 or higher on the Euro-D scale, in the longer term by 31-50 percentage points, relative to the control group. This corresponds to an impact of 0.72–1.16 standard deviations, which is slightly lower compared with the effect size obtained when analysing the Euro-D index.[20] Thus, the initial results indicate that retirement has considerable negative longer-term effects on mental health.[21]

Since we do not detect any short-term effects on mental health, the results imply 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. This interpretation receives support from models including lagged mental health in the third panel, which display almost identical, but slightly more precise, longer-term estimates. Also, as Table A4 shows, the results are very similar when analysing 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, thus incorporating any short-term effects directly. Overall, therefore, the FE-IV estimates point in one direction: retirement has no impact on mental health in the short run, but a large negative effect in the longer term.[22]

Meanwhile, in models evaluating longer-term effects, the Hausman tests always reject the null hypothesis of no endogeneity, indicating that OLS estimates are biased downward. This is unsurprising given the results in Table 3, but may be unexpected since potential reverse causality is often thought to bias OLS estimates in the opposite direction, with poor mental health raising the probability of retirement. Yet the differences are plausible since omitted variables and measurement error may bias estimates downward more than potential reverse causality biases estimates upward.[23]

---

[20] The results are plausible since, as displayed in Table 2, the average Euro-D score increased from 1.91 to 2.25 (sd = 2.05), while the share of depressed individuals increased from 0.19 to 0.25 (sd = 0.46), between waves 2 and 4 among respondents who crossed the state pension age between waves 1 and 2. This should be compared with an increase in the average Euro-D score from 2.11 to 2.27 (sd = 2.17), and the share of depressed individuals from 0.21 to 0.24 (sd = 0.48), between waves 2 and 4 in the group who did not cross the state pension age between waves 1 and 2. The impact of retirement on mental health is effectively identified from these differences.

[21] The reduced-form effect of crossing the state pension age threshold is to generate an increase of 0.287 Euro-D points (standard error = 0.088) and raise the likelihood of depression by 0.053 (standard error = 0.019), using the specifications equivalent to columns 1, 2, 4, and 5 in panel 2. In the sample analysed in columns 3 and 6 in panel 2, the reduced-form impact is 0.253 Euro-D points (standard error = 0.092) and 0.056 (standard error = 0.021) respectively.

[22] As Table A5 shows, the findings are very similar when analysing the lagged impact of retirement and controlling for the direct effect between waves 2 and 4, both in the main sample and when excluding respondents who were retired and/or above the state pension age in wave 1. Overall, these results indicate that our main strategy is appropriate for capturing the longer-term effect of retirement on mental health.

[23] 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).

6.1 Robustness tests

6.1.1 Three-year age window

How sensitive are the results to the specific age window around the threshold? Table 5 displays estimates from models with the sample restricted to individuals aged within approximately three years over and under the lowest and highest thresholds at the second interview. We then include a linear age trend. Again, there is little evidence of any short-term effects. However, the longer-term effects are considerable, despite the fact that 45 per cent of the main sample is dropped. Estimates in Table A6 also show that the long-term effect remains when including a quadratic age trend, despite the narrower age window. Overall, therefore, the results are robust to using a narrower age window, which further supports the idea that our research design captures causal effects.

Table 5: Estimates from FE-IV models (3-year age window)

	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
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	-0.06	-0.22	0.01	0.01	-0.01
Second stage	(0.40)	(0.60)	(0.49)	(0.09)	(0.14)	(0.11)
$\overline{sp}_{it}$	0.20***	0.13***	0.17***	0.20***	0.13***	0.17***
First stage	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
F statistic	123.34	60.33	78.85	123.34	60.33	78.85
Hausman test	0.98	0.99	0.82	0.98	0.91	0.89
Longer-term effects						
Excluding lagged mental health						
$\widehat{r_{it-1}}$	1.30***	1.98***	1.18**	0.25**	0.39**	0.27**
Second stage	(0.44)	(0.69)	(0.53)	(0.09)	(0.15)	(0.12)
$\overline{sp}_{it-1}$	0.20***	0.13***	0.17***	0.20***	0.13***	0.17***
First stage	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
F statistic	123.66	60.64	79.27	123.66	60.64	79.27
Hausman test	<0.01	<0.01	0.02	<0.01	<0.01	0.01
Including lagged mental health						
$\widehat{r_{it-1}}$	1.28***	1.95***	1.08**	0.26***	0.40***	0.27**
Second stage	(0.40)	(0.63)	(0.49)	(0.09)	(0.14)	(0.11)
$\overline{sp}_{it-1}$	0.20***	0.13***	0.17***	0.20***	0.13***	0.17***
First stage	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
F statistic	123.63	60.62	48.67	123.67	60.64	79.28
Hausman test	<0.01	<0.01	0.01	<0.01	<0.01	<0.01
n	4,704	4,704	3,818	4,704	4,704	3,818

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

6.1.2 *Inverse probability weighting*

As noted in Section 4.5, it is important to investigate whether panel attrition threatens the validity of the findings. Table A7 displays results from the same specifications as those estimated in the third panels in Tables 4 and 5, 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. All estimates are very similar to the ones obtained without weights, irrespective of retirement definition utilised and outcome analysed. In fact, the coefficients become larger and 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 unlikely that selective attrition poses a threat to the study's findings.

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 it affects everybody in the same way. Table A8 shows results from models that allow the long-term impact of retirement to differ depending on gender, educational background, and physical as well as psychosocial occupational strain, as discussed in Section 5.1. There is no evidence of statistically significant heterogeneous effects.[24] The coefficients are generally small and/or inconsistent across retirement definitions. In unreported regressions, we found no evidence of heterogeneous short-term effects between the first and second waves either. Overall, this indicates that the zero immediate and negative longer-term effects of retirement on mental health generally apply similarly to men and women as well as individuals with different socio-economic and occupational backgrounds.[25]

[24] This also applies when analysing the sample within the three-year age window. We found some evidence that the negative impact of retirement was only detectable among respondents who were living with a partner at the time of the first wave. However, the number of single individuals being affected by the instrument threshold is small in our sample, which makes it difficult to draw strong conclusions from this exercise.

[25] In unreported regressions, we also estimated separate models for the different groups, but again found little consistent evidence of heterogeneity; no differences were statistically significant and no subgroups appeared to benefit from retirement.

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, they suffer from limitations. Perhaps most conspicuous is that nobody thus far has separated short- from longer-term effects of retirement in a rigorous framework that exploits discontinuities in retirement arising from state pension ages. This is an important shortcoming since there are good theoretical reasons to believe that the short- and longer-term effects of retirement differ.

This study has sought to remedy these issues by investigating the short- and longer-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 utilised an individual-fixed effects IV approach and age-based discontinuities within state pension systems as instruments. Although the results show no impact of retirement in the short run, there is strong evidence of a considerable negative lagged effect that appears within a couple of years' time. This effect, which survives a range of robustness tests, is apparent both when analysing the Euro-D scale as well as the cut-off point for clinical depression. It applies to women and men similarly and also appears to operate independently of individuals' educational background and level of occupational strain. This indicates retirement affects people of different socio-economic backgrounds and professions similarly in terms of mental health.

Yet while the study has found a negative longer-term effect 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 a longer-term perspective 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 immediate negative mental health effects operating independently of retirement, at least if these reforms affect people late in their lives. With time, however, our 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 by delaying the negative longer-term effect of retirement per se.

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.

Atchley, Robert C. 1976. *The Sociology of Retirement*. New York: John Wiley & Sons Inc.

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

Bertoni, Marco, Giorgio Brunello, and Gianluca Mazzarella. 2016. "Does Postponing Minimum Retirement Age Improve Healthy Behaviours Before Retirement? Evidence from Middle-Aged Italian Workers." IZA Discussion Paper No. 9834, Institute for the Study of Labor (IZA), Bonn.

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.

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.

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.

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.

Kroll, Lars E. 2011. "Job Exposure Matrices (JEM) for ISCO and KldB (Version 1.0)." Dataset, GESIS Data Archive. http://doi.org/10.7802/1097.

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.

Lemieux, Thomas and Kevin Milligan. 2008. "Incentive effects of social assistance: A regression discontinuity approach." *Journal of Econometrics* 142(2):807-828.

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.

Petterson-Lidbom, Per. 2012. "Does the size of the legislature affect the size of government? Evidence from two natural experiments." *Journal of Public Economics* 96(3-4):269-278.

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 across countries, gender, and cohorts

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). The state pension age varies by country, gender, and cohort (as indicated by the age ranges in the columns).

Table A2: The short-term association between retirement and mental health (waves 2-4)

	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*
r_{it}	-0.13**	-0.15**	-0.17**	-0.03**	-0.03**	-0.04**
	(0.06)	(0.06)	(0.07)	(0.01)	(0.01)	(0.01)
n	8,551	8,551	7,228	8,551	8,551	7,228

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 panel in Table 3, but instead analysing changes between waves 2 and 4. 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: The short-term impact of retirement on mental health (waves 2-4)

	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
$\widehat{r_{it}}$	-0.04	-0.06	-0.04	0.07	0.10	0.10
	(0.35)	(0.50)	(0.44)	(0.08)	(0.11)	(0.10)
F statistics	165.35	80.49	97.02	165.36	80.51	97.00
Hausman	0.97	0.99	0.97	0.23	0.27	0.24
n	8,551	8,551	7,228	8,551	8,551	7,228

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 panel in Table 4, but instead analysing changes between waves 2 and 4, with all controls measured at t, while also including lagged mental health.

Table A4: Longer-term effects between waves 1 and 4

	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
$\widehat{r_{it-1}}$	1.52***	2.46***	1.47**	0.27**	0.44**	0.33*
	(0.56)	(0.94)	(0.75)	(0.12)	(0.20)	(0.17)
F statistic	88.46	38.25	48.73	88.46	38.25	48.73
Hausman	<0.01	<0.01	0.02	0.03	0.02	0.04
n	8,566	8,566	7,113	8,566	8,566	7,113

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level in parentheses. The specification corresponds to the one used in the second panel in Table 4, but instead analysing changes in mental health between waves 1 and 4.

Table A5: Controlling for short-term effects between waves 2 and 4 (excluding and including further restrictions on the control group)

	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Full working sample						
$\widehat{r_{it-1}}$	1.91***	2.67***	1.62*	0.45***	0.67***	0.51**
	(0.62)	(1.03)	(0.89)	(0.14)	(0.24)	(0.21)
$\widehat{r_{it}}$	0.33	0.06	-0.02	0.18	0.17	0.12
	(0.56)	(0.81)	(0.73)	(0.13)	(0.19)	(0.17)
n	8,550	8,550	7,011	8,550	8,550	7,011
Excluding individuals who were retired at wave 1						
$\widehat{r_{it-1}}$	1.66***	1.61**	1.20**	0.32***	0.34**	0.30**
	(0.52)	(0.63)	(0.59)	(0.12)	(0.14)	(0.13)
$\widehat{r_{it}}$	0.26	0.38	0.15	0.15	0.10	0.10
	(0.44)	(0.53)	(0.48)	(0.10)	(0.12)	(0.11)
n	4,757	3,535	3,334	4,757	3,535	3,334
Excluding individuals who were retired and/or above the state pension age at wave 1						
$\widehat{r_{it-1}}$	2.50***	2.04**	1.64**	0.53***	0.40**	0.38**
	(0.94)	(0.82)	(0.77)	(0.20)	(0.18)	(0.17)
$\widehat{r_{it}}$	-0.03	-0.05	-0.18	0.10	0.03	0.04
	(0.54)	(0.70)	(0.61)	(0.12)	(0.15)	(0.14)
n	4,310	3,416	3,224	4,310	3,416	3,224

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level are in parentheses. The specifications correspond to the one used in the third panel in Table 4, while also including $\widehat{r_{it}}$ (with $\overline{sp}_{it}$ as instrument). The minimum F statistic is 18.45. The minimum Cragg-Donald F statistic is 14.29. The Hausman test always displays values lower/higher than 0.1 for the lagged/non-lagged coefficient

Table A6: Combining a quadratic age trend with the 3-year age window

	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
			Excluding lagged mental health			
$\widehat{r_{it-1}}$	2.61***	3.76***	2.75**	0.50***	0.72**	0.64***
	(0.84)	(1.33)	(1.07)	(0.18)	(0.28)	(0.24)
F statistics	45.92	27.43	28.71	45.92	23.43	28.71
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
			Including lagged mental health			
$\widehat{r_{it-1}}$	2.27***	3.27***	2.25**	0.39**	0.57**	0.48**
	(0.75)	(1.18)	(0.94)	(0.16)	(0.24)	(0.21)
F statistics	45.82	23.33	28.50	46.01	23.37	28.76
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	4,704	4,704	3,818	4,704	4,704	3,818

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

Table A7: Estimates from models using inverse probability weighting

	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
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 4			
$\widehat{r_{it-1}}$	2.21***	3.48***	2.67***	0.39***	0.61***	0.57***
	(0.61)	(1.05)	(0.86)	(0.11)	(0.22)	(0.09)
F statistic	71.95	34.00	39.62	72.01	34.02	39.70
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	8,493	8,493	7,050	8,493	8,493	7,050
			Weighted estimates from the 3rd panel in Table 5			
$\widehat{r_{it-1}}$	1.68***	2.49***	1.67***	0.33***	0.48***	0.38***
	(0.47)	(0.72)	(0.58)	(0.10)	(0.16)	(0.12)
F statistic	106.12	58.20	69.37	106.10	58.23	69.41
Hausman	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	4,661	4,661	3,782	4,661	4,661	3,782

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

Table A8: Heterogeneous effects

	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV	FE-IV
Retirement definition	1	2	3	1	2	3
	Euro-D	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
			Gender			
$\widehat{r_{\iota t-1}} * female$	0.07	1.53	0.04	-0.06	0.16	-0.05
	(0.48)	(0.98)	(0.60)	(0.10)	(0.20)	(0.14)
$\widehat{r_{\iota t-1}}$	1.57***	2.04***	1.67**	0.33***	0.41***	0.39***
	(0.54)	(0.71)	(0.67)	(0.12)	(0.15)	(0.15)
n	8,566	8,566	7,113	8,566	8,566	7,113
			Education			
$\widehat{r_{\iota t-1}} * low\ education$	-0.53	-0.28	-0.66	-0.10	-0.06	-0.14
	(0.57)	(0.81)	(0.69)	(0.12)	(0.17)	(0.16)
$\widehat{r_{\iota t-1}}$	2.02***	2.86***	2.22**	0.36**	0.51**	0.48**
	(0.66)	(1.02)	(0.92)	(0.15)	(0.22)	(0.21)
n	8,502	8,502	7,056	8,502	8,502	7,056
			Occupational Physical Burden			
$\widehat{r_{\iota t-1}}*high\ OPB$	0.04	0.14	0.24	-0.10	-0.11	-0.04
	(0.49)	(0.69)	(0.60)	(0.11)	(0.15)	(0.13)
$\widehat{r_{\iota t-1}}$	1.55***	2.33***	1.76**	0.34***	0.49***	0.38**
	(0.53)	(0.84)	(0.69)	(0.12)	(0.19)	(0.16)
			Occupational Psychosocial Burden			
$\widehat{r_{\iota t-1}}*high\ OSB$	0.21	0.17	0.57	0.07	0.08	0.13
	(0.48)	(0.68)	(0.57)	(0.11)	(0.15)	(0.13)
$\widehat{r_{\iota t-1}}$	1.43**	2.28**	1.47*	0.24*	0.39*	0.27
	(0.60)	(1.00)	(0.80)	(0.14)	(0.22)	(0.18)
n	7,607	7,607	6,665	7,607	7,607	6,665

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level are in parentheses. The minimum F statistic for the instrumented interaction is 61.70 and the minimum F statistics for $\widehat{r_{\iota t-1}}$ is 18.70. The minimum Cragg-Donald F statistic is 23.77. The Hausman test always displays p-values lower than or equal to 0.1. All models include lagged mental health. OPB = Occupational Physical Burden. OSB = Occupational Psychosocial Burden.