SCHOOL OF
ECONOMICS AND
MANAGEMENT

LUND
UNIVERSITY

Retiring from Happiness?

Analysis of retirement and mental health using SHARE data

by

Bardh Manxhuka & Ea Dumancic

August 2022

Department of Economics

NEKN06

Master Essay I (15 credits)

Supervisor: Ulf Gerdtham

Abstract

We utilize data from the *Survey of Health, Ageing and Retirement in Europe* (SHARE) [1] to investigate the impact of retirement on mental health in a multi-country setting. To deal with endogeneity in retirement behaviour, we employ an individual-fixed effects IV strategy where pension eligibility thresholds at which financial incentives to retire are exploited to predict retirement behaviour. The combination of these quasi-experimental methods, with some borrowed intuition from the regression discontinuity literature, is the premise on which we are able to distinguish between short-, medium-, and long-term effects of retirement on mental health. Retirement is found to have no significant impact on mental health in the short- to medium-term. However, we find solid evidence of a large and negative impact of retirement on mental health in the long-term. The mental health effect of retirement is found to be homogeneous in terms of gender and marital status, but heterogeneous across educational attainment levels.

Keywords: retirement, mental health, SHARE, fixed effects, IV

---

[1] This paper uses data from SHARE Waves 4, 5, 6, and 7 *(DOIs:* 10.6103/SHARE.w4.800, 10.6103/SHARE.w5.800, 10.6103/SHARE.w6.800, 10.6103/SHARE.w7.800, 10.6103/SHARE.w8.800,). *The SHARE data collection has been funded by the European Commission, DG RTD through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782, SHARE-COVID19: GA N°101015924) and by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, and VS 2020/0313. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged* (see www.share-project.org).
[1] Börsch-Supan, A., M. Brandt, C. Hunkler, T. Kneip, J. Korbmacher, F. Malter, B. Schaan, S. Stuck, S. Zuber (2013). *Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)*. International Journal of Epidemiology. DOI: 10.1093/ije/dyt088.

Table of contents

List of tables

1. Introduction

The sustainability of pension systems has been at the forefront of social planning in European countries in recent years. The 2021 Ageing Report of EU member states projects that by 2070 there will be 1.7 workers per retiree, as opposed to 2.9 workers per every retiree recorded in 2019, signalling a shift in the EU countries' demographic composition in favour of the elderly. The global phenomenon of demographic transition is therefore expected to put substantial pressure on Europe's pay-as-you-go pension systems, health care, and economic growth (Earl & Archibald, 2014). Subsequently, the literature surrounding retirement and its effects has seen an expansion over a vast number of disciplines – from health economics to sociology, medicine, and psychology; all with an aim to study the implications of ageing and retirement for the sake of policy planning. Particular emphasis has been put on the effects of retirement on the mental- and physical health of newly retired. If the act of retiring worsens individual wellbeing, it is in society's best interest to prolong workers' participation in the labour force, postpone the act of retiring, and alleviate the pressure on the fiscal sustainability of pension systems. Contrary, if retirement has a protective impact on health, policy implications should shift in the reverse direction. This paper adds to the literature by using SHARE data in order to assess how entering retirement may affect one's mental health and wellbeing. Previous research on this topic has not provided a uniform answer, with retirement being found to have a negative, zero, as well as a positive effect on mental health. The direction of previous results is indicated to depend, among other factors, on the panel data utilized, period observed, and identification strategy employed. Additionally, we try to identify and compare short-term effects of retirement on mental health with medium- (minimum of two years after retirement) and long-term effects (minimum of four years after retirement) by employing an individual-fixed effects IV approach. The main intuition stems from using state pension eligibility ages as a source of exogenous variation in retirement behaviour. Discontinuities in retirement probability arise upon reaching the state pension-eligibility age thresholds, where financial incentives to retire increase the likelihood of retirement. Thus, our identification strategy allows for a fuzzy RDD intuition in a 2SLS estimation. We make certain to address all potential weaknesses of our estimation method, particularly those pertaining to the endogeneity of retirement and mental health. The novelty of this paper is in distinguishing between short-, medium-, and long-term effects of retirement, as well as by capturing effects on a unique sample of SHARE countries. Our main findings suggest no significant mental health effects of retirement in the short-to-medium term. Still, we present solid evidence of negative long-term effects of retirement on mental health. Results are heterogenous by educational attainment levels. Finally, surviving an appropriate line of robustness checks, the main results prove insensitive to model specification, choice of bandwidth, and panel attrition.

The paper is organized as follows; Section 2 introduces Grossman's human capital theory as the theoretical basis for this research, Section 3 provides a detailed overview of previous relevant literature, Section 4 organizes the working sample and main variables of interest, and Section 5 lays out the identification strategy. Results of this study can be found in Section 6 and the subsequent discussion and conclusion are given in Sections 7 and 8 respectively. Overall, the implications of this paper's findings are important, not just for policy planning, but for independent retirement planning as well.

2. Theoretical background

Let us ask the relevant question; why and how ought retirement affect mental health? Consistent with relevant literature, we opt to frame this question within the human capital model of the demand for health, in which health is portraited as a durable capital stock that depreciates with time and increases with investments (Grossman, 1972, pp.28-47)[2]. According to Grossman's model, health is a source of utility given that it produces satisfaction. Conversely, the amount of health stock determines an individual's wealth. Thus, the demand for health comes down to two primary reasons: health capital increases individuals' utility, directly entering their preference functions, and increases healthy time, consequently increasing time for work and hence lifetime earnings. Therefore, health can be viewed as both a consumption and investment commodity. However, the demand for health changes its properties after retirement, as wage-related earning incentives are no longer present, implying that the health investment motive may be neutralized. In this case, one is to expect a decline in health post retirement. Nonetheless, Grossman (1972) argues that improved health increases both market and nonmarket productivity. This implies that, since retirees experience a rise in leisure-time, the consumption value of health may rise as well. In this case, we may expect retirees to invest more in their health by e.g. sleeping and exercising more, and therefore experience an increase in health post retirement[3]. Ultimately, the net effects of retirement on health remain ambiguous and depend on whether the marginal utility of health increases or

[2] Taken from (Bhattacharya, Hyde & Tu, 2018, pp.28-47) textbook on Health Economics.

[3] Increased leisure time may intuitively decrease the marginal cost of health investments. Retirees have more time for health favourable activities, e.g. sleep, physical activity, social interactions, regular hospital visits etc.

decreases after retirement. On the other hand, there may be several explanations to how retirement might affect health, which do not necessarily rely on rational choice.

For example, going into retirement certainly changes individual's day-to-day practices, likely having less to no interactions with former colleagues and perhaps more interactions with family and friends. As prior empirical work has shown an individual's social capital to positively impact self-reported health and life satisfaction (Becchetti, Pelloni & Rossetti, 2008; Cohen, 2004; d'Hombres et al., 2010; Elgar et al., 2011; Folland, 2008; Glass et al., 1999; Melchior et al., 2003; Nieminen et al., 2013; Pinquart & Sörensen, 2000; Powdthavee, 2008; Ronconi, Brown & Scheffler, 2012; Saffer, 2005; Taylor et al., 2001), mental health may theoretically rely on how an individual's social network changes after retirement, i.e. whether it increases or decreases [4]. Then again, increased social capital must not necessarily have a positive impact on mental health. For instance, partnered retirees may spend more time than what is optimal with their spouse and children[5], potentially harming their mental health as well as experiencing negative cross-partner 'spillover' effects (e.g. Bertoni & Brunello, 2017; Müller & Shaikh, 2018)[6]. Within these lines, it may be difficult to predict the average effect of retirement on mental health. Increased leisure time reduces work-related responsibilities which in turn may reduce stress. Conversely, besides that retirement itself may be a stressful event, withdrawal from the labour force is also associated with reduction in income and consumption, which may very well increase stress, and consequently affect mental health negatively. Furthermore, retirement may affect individuals heterogeneously by education and income. In context of the Grossman model, we can expect more educated retirees to be more efficient producers of health since information concerning health-preserving inputs increases with schooling (Grossman, 1972). Additionally, assuming that income increases with education, we can expect less educated retirees to have lower socioeconomic status and retire from more physically demanding occupations. Elaborating further, however, the sign of this causal

[4] For instance, d'Hombres et al. (2010) find that social isolation is negatively and significantly associated with health. Ronconi, Brown & Scheffler (2012) present evidence of higher levels of social capital increasing health, irrespective of gender. Taylor et al. (2001) find that social relations and network factors (e.g. family support, number of friends, and frequency of contact with neighbours) are all significantly related with self-reported life satisfaction and happiness. Pinquart and Sorenson (2000) present a positive association between frequency of contact with acquaintances and self-reported life satisfaction among the elderly.

[5] For instance, Pinquart and Sorenson (2000) show that increased time spent with friends is more beneficial for self-reported life satisfaction than increased contact with adult children among the elderly.

[6] Both Bertoni and Brunello (2017) and Müller and Shaikth (2018) find that subjective health and mental health are negatively affected by spousal retirement but positively affected by own retirement. Results are heterogenous by gender.

pathway of retirement on health remains ambiguous at second glance: more educated individuals are more efficient producers of health and should therefore have better mental health outcomes as retirees, but less educated individuals retire from more physically demanding jobs and should therefore be happier to withdraw from the labour force.

It is essential to note that initial mental health effects of retirement may not necessarily reflect those of later stages. This because the process of retirement is likely to have different stages, all potentially affecting individual's mental health differently (Atchley, 1982). Initially, retirees are expected to go through the 'honeymoon stage' and perhaps perceive retirement as vacation rather than a permanent withdrawal from the labour force. During this stage, we may expect retirees to feel carefree and relaxed. After the initial stage, retirees are expected to undergo a period of disenchantment as reality of a life-changing event settles in, likely causing a decline in mental health. These stages are followed by the 'reorientation' stage and 'routine' stage where retirees settle into their new lifestyle and recognize new interests and opportunities. Assuming that there are several phases post retirement, it is very likely that retirement may affect mental health differently in the short-, medium-, and long-term perspective. Overall, it remains difficult to predict the direction of the hypothesized pathways.

3. Previous literature

In conjunction with pension system reforms during recent decades, there has been a substantial, and still flourishing, literature exploring the impact of retirement on objective and subjective health. Since early research, ranging across scientific fields, notions of retirement having negative effects on health have been fairly consistent (Minkler, 1981)[7]. Withdrawal from the labour force is traditionally considered by many as harmful to an individual's social capital, and the post-retirement era is argued to be associated with sentiments of loneliness, monotony, and lack of purpose. On the contrary, as we further discuss in this section, a fair share of recent findings presents protective effects of retirement on mental health.

Even with this long-lived debate, previous research investigating the *retirement* — *mental health* relationship exhibits no clear consensus. Correlational studies have found positive (Dave, Rashad & Spasojevic, 2008a; Jokela et al., 2010; Mein et al., 2003; Midanik et al., 1995; Oksanen et al., 2011; Ostberg & Samuelsson, 1994) as well as negative (Lindeboom, Portrait & van den Berg, 2002; Mosca & Barrett, 2016; Szinovacz & Davey, 2004; Vo et al., 2015) effects of retirement on mental health. Nonetheless, as the decision of retirement is not random,

[7] We refer to Minkler (1981) for an extensive review of this.

the methodological approaches of these studies are unable to account for reverse causality and consequently fail to determine causal inference, further discussed in Section 5**.**

Trying to solve the endogeneity problem of the retirement decision, various researchers have utilized state pension eligibility ages at which financial incentives to retire are exploited to predict retirement behaviour, thus attempting to obtain exogenous variation in the retirement probability. Intuitively, the idea is that reaching the state pension eligibility age should increase the probability of retirement. Reaching the threshold should, however, not have any impact on health through any other channel than retirement itself (Angrist & Pischke, 2009, pp.113-138). This is thoroughly discussed in Section 5. The nature of these thresholds has presented the opportunity for researchers to implement an instrumental variables (IV) framework and/or fuzzy regression discontinuity design (RDD), where reaching the state pension eligibility age serves as instrument for retirement (Belloni, Meschi & Pasini, 2016; Bonsang, Adam & Perelman, 2012; Celidoni, Dal Bianco & Weber, 2017; Eibich, 2015a; Fé & Hollingsworth, 2016; Fonseca et al., 2014; Gorry & Slavov, 2021; Heller-Sahlgren, 2017a; Hessel, 2016; Horner, 2014; Kesavayuth, Rosenman & Zikos, 2016; Kolodziej & García-Gómez, 2019a; Mazzonna & Peracchi, 2017; Müller & Shaikh, 2018; Picchio & Ours, 2020; Rohwedder & Willis, 2010). Alternatively, some utilize retirement reforms to implement a difference-in-differences approach, thus attempting to capture the average health effects of retirement on a group level (e.g. Messe & Wolff, 2019).

It must be highlighted, however, that previous literature is not absent from limitations. For instance, studies analysing the effects of retirement on mental health tend to include various *bad* controls in their regression models, i.e. control variables that are themselves outcome variables of retirement (Angrist and Pischke, 2009, pp.64-68). Examples of such bad controls are earnings and marital status, that are likely to be related to mental health and be endogenous to retirement itself, thus producing a version of selection bias. Including bad controls as such might disguise the total causal effect of retirement, which is likely to run through certain demographic factors, e.g. income and marital status (Angrist and Pischke, 2009, pp.64-68). Further, only a handful of studies explore differences between short- and long-term effects of retirement (Celidoni, Dal Bianco & Weber, 2017; Fé & Hollingsworth, 2016; Heller-Sahlgren, 2017a; Horner, 2014). Overlooking this potential pitfall can produce biased estimates, e.g. if retirement has a significant long-term impact on mental health, then only observing immediate changes in mental health may very well underestimate the total effect of retirement, further discussed in Section 5. Moreover, numerous studies do not consider heterogeneity in the impact of retirement on mental health, which proves to be highly relevant in many studies discussed further on. Related to our methodology, previous literature implementing instrumental variables approach with intuitions from the RDD literature tend to exclude vital sensitivity checks. Firstly, one should investigate if the results are robust to narrowing down the observation 'bandwidth' around the discontinuity. There is a common trade-off between bias and precision when determining the bandwidth since a broader bandwidth provides more observation and consequently higher precision, but observations further away from the thresholds are less comparable, hence increasing the bias. Secondly, researchers have on occasions ignored potential non-linear effects of age, by not providing results where the effect of age is allowed to differ on each side of the pension-eligibility ages used as instruments in their model specifications.

The strand of literature we consider most relevant to this paper, summarized in Table 1, implements similar methodology as us and focuses on subjective as well as objective measures of health[8]. Investigating the effects of retirement on cognitive abilities, Rohwedder and Willis (2010) use cross-country data from eleven European countries together with household data from the US and England. The authors implement an IV methodology while utilizing retirement eligibility ages to account for the endogeneity of the retirement decision and find that retirement has a negative effect on cognitive ability. Comparable results are found by Bonsang et al. (2012) who employ a similar IV approach using US data. Then again, other studies find mixed results. For instance, Bonsang and Klein (2012) investigate retirement effects on life satisfaction by using German Socio-Economic Panel data employing a fixed-effects approach. Allowing the effects to differ by terms of retirement, they find that voluntary retirement has no significant effect on life satisfaction, whereas involuntary retirement has a clear negative effect. These findings are confirmed by Abolhassani and Alessie (2013) using the same data in a similar methodological framework. Horner (2014) utilizes pension eligibility ages in an instrumental variable approach and analyses international data from sixteen European countries in Western Europe and the US. The author concludes that retirement has a positive, but temporary, effect on well-being. Investigating SHARE-data from eleven European countries, Fonseca et al. (2014) analyse the effects of retirement induced by state pension eligibility on depression. Once solving for endogeneity using instrumental variables, they find weak evidence that retirement may be protective against depression. Eibich (2015) exploits financial incentives in the German pension system employing a regression discontinuity design for identification. Retirement is concluded to have a positive effect on

[8] E.g., self-assessed health, well-being, life satisfaction, cognitive ability etc.

Table 1: Overview of relevant literature

Study	Country(ies)	Data	Time	Age	Outcome variable	Method	Sign of effect	Het.
(Rohwedder & Willis, 2010)	8	Numerous	2004	60-64	Cognitive abilities	IV – RAE	-	No
(Bonsang, Adam & Perelman, 2012)	US	HRS	1998-2008	51-75	Cognitive abilities	IV – RR	-	No
(Bonsang & Klein, 2012)	DE	GSOEP	1995-2010	50-70	Life satisfaction	FE	+ (Vol), – (Invol)	No
(Abolhassani & Alessie, 2013)	DE	GSOEP	1994-2009	50-70	Life satisfaction	FE	+ (Vol), - (Invol)	No
(Horner, 2014)	17	Numerous	2004-2006	50-70	Well-being	IV – RAE	+ (temporary effect)	No
(Fonseca et al. 2014)	11	SHARE	2004-2010	50+	Depression	IV – RAE	-	No
(Eibich, 2015)	DE	GSOEP	2002-2009	55-70	Mental health	RDD	+	E
(Belloni, Meschi & Pasini, 2016)	10	SHARE	2004-2013	55-70	Mental health	IV – RAE	+ (Male), O (Female)	O
(Kesavayuth, Rosenman & Zikos, 2016)	UK	BHPS	2005-2012	50-75	Well-being	IV – RR	O	G, P
(Hessel, 2016)	12	EU-SILC	2009-2012	50-74	Self-reported health	IV	+	No
(Fe & Hollingsworth, 2016)	UK	BHPS	1991-2005	50-80	Health indicators	RDD	O	No
(Mazzonna & Peracchi, 2017)	10	SHARE	2004-2006	50-70	Health & Cog. Abilities	IV – RAE	-	O
(Bertoni & Brunello, 2017)	JPN	PPS	2008-2013	42-69	Mental health	DID – RR	-	No
(Celidoni, Dal Bianco & Weber, 2017)	10	SHARE	2004-2011	50+	Cognitive abilities	IV – RAE	-	No
(Heller-Sahlgren, 2017)	10	SHARE	2004-2012	50+	Mental health	IV – RAE	-	No
(Gorry, Gorry & Slavov, 2018)	US	HRS	1992-2014	50-93	Health & Life satisfaction	IV – RAE	O (ST), - (LT)	E, G, O
(Muller & Shaikh 2018)	19	SHARE	2004-2013	45-91	Subjective health	IV – RAE	+	No
(Messe & Wolff, 2019)	FR	LFS	2013-2016	50-70	Self-reported health	RDD – RAE	O (Male), - (Female)	No
(Kolodziej and García-Gómez, 2019)	11	SHARE	2004-2013	55-69	Mental health	DID	+	O
(Atalay, Barrett & Staneva, 2019)	AU	HILDA	2012	55-74	Cognitive abilities	FD – IV	-	G
(Picchio & Ours, 2020)	NL	LISS	2007-2017	65+	Mental health	RDD – RAE	+	G, M
(Gorry & Slavov, 2021)	ENG	ELSA	2002-2015	50+	Health biomarkers	IV – RAE	0 (Obj), + (Subj)	No

Notes: DATA: BHPS = British Household Panel Survey, EU-SILC = European Union Statistics on Income and Living Conditions; LFS = Labour Force Survey; GSOEP = German Socio-Economic Panel; HRS = Health and Retirement Study; HILDA = Household, Income and Labour Dynamics in Australia; LISS = Longitudinal Internet Studies for the Social Sciences; SHARE = Survey of Health Aging and Retirement in Europe; CHARLS = China Health and Retirement Longitudinal Survey.
COUNTRIES: US = United States; DE = Denmark; UK = United Kingdom; JPN = Japan; FR = France, AU = Austria, CH = China, ENG = England, NL = Netherlands
METHOD: D – RR = Difference-in-differences using a retirement reform; FE = Fixed effects; IV – RAE = Instrumental variables using retirement age eligibility; RDD = Regression Discontinuity Design; FD-IV = First difference IV
Het = Heterogeneity: E = Educational attainment, G = Gender, M = Marital status, H = (mental) Health status, I = Income, O = Occupation, P = Personality traits.

subjective health status and mental health, where the key mechanisms through which retirement affects health are indicated to be increased sleep duration, relief from work-related stress, and increased physical activity during leisure-time. The health effects of retirement are heterogeneous across educational attainment levels, as lower educated workers benefit more in physical health, whereas higher educated workers benefit more in mental health post retirement. Using SHARE-data from ten European countries and applying a fixed-effects IV approach, Belloni et al. (2016) find that retirement improves mental health for men, whereas women are on average unaffected. Effects are heterogenous in the sense that the effect is greater for blue-collar workers working in regions most affected by economic crises. Kesavayuth et al. (2016) use data on older men and women from the British Household Panel Survey (BHPS) and apply an instrumental variables approach. By exploiting retirement eligibility ages to obtain exogenous variation in retirement behaviour, the authors find that retirement on average has no impact on well-being. They find evidence, however, of heterogeneity related to gender and personality traits. Specifically, the retirement effect on well-being for women who score high in 'openness' or low in 'conscientiousness' is stronger compared to other women, whereas personality traits have no significant importance for men. Hessel (2016) uses longitudinal data from twelve European countries to investigate the effect of retirement on self-reported and physical health. Implementing an instrumental variables approach, the author concludes that retirement improves self-reported health for men and women homogenously across all educational levels. Fé and Hollingsworth (2016) use BHPS data to explore the short- and long-term effects of retirement on health. In order to identify the short-term effects, the authors apply a RDD approach, whereas a parametric panel data model is used to estimate the long-term effects. Results indicate that retirement has little to no short- or long-term effects on health. Exploiting the panel dimension of the Survey of Health Ageing and Retirement in Europe, Mazzonna and Peracchi (2017) implement an instrumental variables approach using old age retirement rules across ten European countries to obtain an exogenous source of variation in retirement behaviour. The authors conclude a negative retirement effect on health and cognitive abilities, increasing by time after retirement. Moreover, the authors present evidence of considerable heterogeneity in the retirement effect across occupational groups, as retirement has a positive and immediate effect on both health and cognitive abilities for physically demanding occupations. Bertoni and Brunello (2017) analyse cross-partner retirement effects on mental health by using Japanese micro-data to investigate the so-called "Retired Husband Syndrome" which anecdotally discusses the mental health effects of wives of retired men. Results suggest that retirement of husbands has negative effects on both their own and their wife's mental health. Investigating how retirement affects cognition, Celidoni et al. (2017) find that retirement is beneficial at first but causes a decline in cognitive abilities in the long-term. Retirement speeds up cognitive decline for individuals who retire at statutory retirement age, whereas it has a protective role for those who take early retirement. Heller-Sahlgren (2017) uses SHARE data from ten European countries and implements a fixed-effects IV approach to distinguish the short- and long-term effects of retirement on mental health. Exploiting thresholds created by state pension ages, the author finds no short-term effects of retirement on mental health. However, long-term effects of retirement prove to be large and significantly negative on mental health. The effects are homogenous with respect to gender, educational level, and occupational background. In line with relevant previous literature, Gorry et al. (2018) use retirement eligibility ages in the US in instrumental variables approach to investigate the impact of retirement on health, life satisfaction, and healthcare utilization. Findings suggest that retirement has a positive and immediate effect on life satisfaction, while other health improvements appear later on. However, the authors find no evidence of retirement affecting healthcare utilization. Müller and Shaikh (2018) investigate the causal effect of spousal retirement on subjective health across nineteen European countries. Applying a fuzzy RDD and using retirement eligibility ages as an exogenous variation in retirement behaviour, they find that subjective health is negatively affected by spousal retirement, but positively affected by own retirement. The effects are heterogenous with respect to gender. While spousal retirement has no significant effect on husbands' health, retirement of the husband has a negative impact on wife's subjective health. These results are in line with Bertoni and Brunello (2017), who present negative spillover retirement effects on female mental health in Japan. On the other hand, applying a semi-parametric difference-in-differences methodology on French Labour Force Survey data, Messe and Wolff (2019) find no significant cross-partner spillover effects of retirement on health. The authors find, however, positive short-term effects of retirement on own self-reported health for those retiring from occupations with low physical burden. Unique to previous literature, Kolodziej and García-Gómez (2019) use SHARE data from eleven European countries to estimate not only the average effects of retirement on mental health, but also investigate whether retirement effects are unequally distributed across the mental health distribution. By applying an instrumental variables approach based on retirement eligibility age, they conclude a protective impact of retirement on mental health. The positive effects of retirement are unequally distributed and larger for those just around the threshold of clinical depression. Using a similar methodology on HILDA data, Atalay et al. (2019) find that retirement has a negative but modest effect on cognitive abilities. The long-term decline in cognition is greater for men than women. Consistent with previous literature, Picchio and Ours (2020) study the effects of retirement on mental health by applying a fuzzy RDD based on state-pension eligibility ages in the Netherlands. The authors find heterogenous health effects of retirement by marital status and gender. While retirement of partnered men positively impacts both their own and partner's mental health, retirement of partnered women has no significant impact on neither. These findings contradict some previous studies (Bertoni & Brunello, 2017; Müller & Shaikh, 2018) and somewhat confirm other (Messe & Wolff, 2019). Further, single individuals experience no mental health effect of retirement. In contrast to many previous studies, Gorry and Slavov (2021) focus on objective health biomarkers rather than subjective self-reported health and life satisfaction. Applying an instrumental variables approach on ELSA data, the authors find mixed and mostly statistically insignificant results of retirement on objective measures of health biomarkers. They confirm, however, that retirement appears to consistently improve self-reported health.

Overall, previous literature presents ambiguous results on objective general- and physical health, whereas retirement appears to mostly have favourable effects on self-reported mental health and life satisfaction. The presented strand of literature indicates that inconsistent findings can potentially be explained by the choice of countries under study and whether or not heterogeneous effects of retirement are considered. Yet, more than anything, differences in results across the literature appear to be highly attributed to the choice of estimation method (Nishimura, Oikawa & Motegi, 2017). This study aims to account for highlighted limitations and complement previous literature by allowing the mental health effects of retirement to operate in a fairly unique (to the best of our knowledge) time-dimension. Similar to our identification strategy, (Heller-Sahlgren, 2017) attempts to distinguish between short- and long-term effects of retirement on mental health by observing changes in mental health 2-4 years after one has entered the retirement era. Utilizing newer SHARE waves we are able to control for potential time-related effects for which (Heller-Sahlgren, 2017) could not: *Medium-term effects of retirement on mental health*[9]. We utilize four waves of the Survey of Health, Ageing, and Retirement in Europe (SHARE) and use state pension eligibility ages across 10 European countries as exogenous instruments for retirement behaviour. Our complete methodology is thoroughly discussed in section 5.

[9](Heller-Sahlgren, 2017) uses the first, second, and fourth waves of the SHARE data. Thus, he is not able to account for, or control for, neither changes in mental health nor time-specific effects between the second and fourth wave. If significant factors lie within that time-period, estimators are likely to be exposed to under- or overestimation of the long-term causal effect.

4. Data

This study's main aim is to employ panel data collected on Europe's aging population to investigate the effect of retirement on mental health of respondents. This section therefore gives an insight into data employed for that purpose, as well as a detailed overview of the working sample, variable construction, and corresponding theoretical background.

4.1 SHARE

For the purpose of this research we utilise four waves of the Survey of Health, Ageing, and Retirement in Europe (SHARE), a biennial survey of individuals aged 50 or older, jointly coordinated by Munich Center for the Economics of Aging and Max-Planck-Institute for Social Law and Social Policy. The survey was first conducted in 2004, with a lengthy questionnaire of 20 modules covering topics of health, health care, wealth, retirement, social networks, and living conditions of residents of eleven European countries. Nowadays, survey boasts eight survey waves, more than 30 modules, two special COVID surveys, and 530,000 in-depth interviews with 140,000 people from 28 European countries and Israel. (The Survey of Health, Ageing and Retirement in Europe (SHARE): Home, 2022).

As with any panel database, SHARE is not without its limitations. It suffers from irregularities in participation across countries and waves, as well as from missing values and minor differences in sampling procedures (Lusa & Huebner, 2021). Its weaknesses, however, are not detrimental to our analysis as it is still one of the largest cross-country panel studies, providing invaluable multidisciplinary micro data on Europe's aging population used by both researchers and policy makers alike (Maskileyson, Seddig & Davidov, 2021). Most importantly, the longitudinal aspect of SHARE provides great opportunities for identification strategies. As discussed in Section 5, we take significant advantage of this.

4.2 Sample

Data for the four waves employed in this paper were collected in 28 countries over the span of eight years, from 2011 to 2017. Number of unique observations in each wave ranges from 58,000 to 77,000, which totals 269,352 observations over four waves. Following previously outlined methodology we exploit the fact that each survey wave is, on average, two years apart, and broaden our research to an often overlooked question of capturing and contrasting short-, medium-, and long-term effects of retirement on mental health (Heller-Sahlgren, 2017; Messe & Wolff, 2018; Nielsen, 2018). As discussed in Section 2, focusing only on immediately visible effects can potentially expose our results to bias from the '*honeymoon effect*' (Atchley,1982), and thus underestimate the total effect of retirement on mental health. Therefore, we consider that it is unlikely to accurately predict the timing of the impact, as retirement might affect mental health differently in the short-, medium-, and long-term perspective. As further elaborated in Section 5, we utilise pension eligibility ages as instruments for retirement to obtain an exogenous variation in retirement probability. This in turn allows for a fuzzy RDD intuition. Thus, relevant to the methodological approach presented in Section 5, we restrict our sample to individuals aged 50-76 years old at the time of the fifth wave interview, establishing a bandwidth of 10 years above and below the maximum and minimum threshold, presented by Table A1.[10] Further, we restrict the sample strictly to individuals who are interviewed in all four of the waves utilised, namely SHARE waves four, five, six, and seven. This narrows down the number of countries observed to ten: Austria, Germany, Italy, Spain, France, Belgium, Denmark, Sweden, Switzerland, and Czech Republic. Ultimately, our working sample is restricted to a maximum of 15,748 individuals observed at four separate occasions. The number of observed individuals in our main estimation models is dependent on retirement definition utilized, further discussed in Subsection 4.3. When employing relevant robustness tests in terms of bandwidth choice, discussed in Subsection 5.2, we further restrict our working sample to a maximum of 9,365 individuals, all observed at four separate occasions. In line with the fuzzy RDD intuition, we treat those who cross their state pension eligibility age between waves 4 and 5 as our treatment group. By definition, the control group consists of those individuals who did not cross their state pension eligibility age between waves 4 and 5. Since we are able to follow individuals for two additional waves after they retire, corresponding to approximately 4 years, we believe our identification strategy is able to capture the immediate effect, as well as the lagged medium- and long-term effects of retirement in a feasible manner[11]. This is thoroughly discussed in Section 5. Table 2 displays the descriptive statistics by working sample and treatment status.

[10] The state pension eligibility ages presented in Table A1 are obtained from the U.S.A.'s Social Security Administration's survey and are dependent on country, birth cohort, and gender. Ages were valid at the time of the survey waves four and five, as well as in the period in-between the two waves, which is the relevant period at which respondents belonging to our treatment group crossed the pension threshold.

[11] As later discussed, our methodological set-up only provides estimated effects on compliers (LATE), i.e. those who change retirement status as a result of reaching their state pension eligibility age, but would not have done so otherwise.

Table 2: Descriptive statistics reported by working sample and treatment status

Variables	Main sample				Treatment group	Control group
	Mean	Standard deviation	Min	Max	Mean	Mean
WAVE 4						
Age	62.78	6.59	48.16	75.16	63.39	62.71
Gender	0.57	0.49	0	1	0.57	0.57
Retirement definition (1)	0.48	0.49	0	1	0.52	0.48
Retirement definition (2)	0.61	0.48	0	1	0.68	0.60
Retirement definition (3)	0.57	0.49	0	1	0.64	0.56
Euro-D	2.19	2.09	0	12	2.11	2.20
Clinical Depression	0.23	0.42	0	1	0.21	0.23
WAVE 5						
Age	64.69	6.60	50.08	76.91	65.32	64.62
Gender	0.57	0.49	0	1	0.57	0.57
Retirement definition (1)	0.55	0.49	0	1	0.71	0.54
Retirement definition (2)	0.67	0.46	0	1	0.82	0.65
Retirement definition (3)	0.64	0.47	0	1	0.81	0.62
Euro-D	2.16	2.10	0	12	2.07	2.16
Clinical Depression	0.22	0.41	0	1	0.21	0.22
WAVE 6						
Age	66.65	6.59	52.16	78.91	67.29	66.59
Gender	0.57	0.49	0	1	0.57	0.57
Retirement definition (1)	0.62	0.48	0	1	0.84	0.60
Retirement definition (2)	0.72	0.44	0	1	0.93	0.70
Retirement definition (3)	0.79	0.45	0	1	0.93	0.68
Euro-D	2.15	2.09	0	12	2.11	2.15
Clinical Depression	0.22	0.41	0	1	0.20	0.22
WAVE 7						
Age	68.74	6.60	54.25	80.83	69.37	68.68
Gender	0.57	0.49	0	1	0.57	0.57
Retirement definition (1)	0.58	0.49	0	1	0.87	0.66
Retirement definition (2)	0.68	0.46	0	1	0.92	0.71
Retirement definition (3)	0.67	0.46	0	1	0.95	0.76
Euro-D	2.25	2.16	0	11	2.26	2.25
Clinical Depression	0.23	0.42	0	1	0.25	0.23
n	*15,748*				*1,484*	*14,264*

Note: Working sample consists of max. 15,748 individuals in each wave, out of which 1,484 belong to the treatment group (those who reached their state pension eligibility age in-between waves four and five). Remaining 14,264 are by definition assigned to the control group (those who did not reach their state pension eligibility age in-between waves four and five). n = 15,748 applies for the first and second definition of retirement whereas n =13,818 applies for the third definition of retirement.

4.3 Retirement

Retiring is an act of leaving the workforce, done either because individuals are elderly and no longer able to work or because they no longer want to work. In 1889, Germany was the first country in the world to introduce retirement benefits, setting a precedent for publicly funded retirement programs (Social Security History, 2022). Countries all around the world have followed suit and installed an age threshold at which individuals become eligible for financial payment should they choose to exit the labour force and declare themselves retired. Today, retiring is synonymous both with exiting from the labour force and claiming public pension benefits[12]. Eligibility for retirement benefits is the main reason why individuals decide to take up the offer, as they can stop working but still receive something akin to a monthly salary. In countries such as England and France, with good pension systems, elderly individuals tend to retire immediately upon reaching state pension age (Motegi, Nishimura & Oikawa, 2020). If a person decides to retreat into retirement before the legal retirement age they are considered 'early' retired and receive only partial pension benefits, or no benefits at all (Larimore et al., 2009). We refrain from using early retirement ages in our model since not all European countries provide the opportunity for early retirement, and if they do, such opportunities are often limited to certain career paths. This gives rise to selection bias as individuals might purposefully choose jobs with an early retirement scheme (Coe & Zamarro, 2011; Heller-Sahlgren, 2017). It is also possible that taking up an early retirement offer is more prevalent among workers who report exhaustion with their working lives, again biasing the sample (Knoll, 2011). All these factors make early retirement ages particularly vulnerable to measurement errors (Heller-Sahlgren, 2017) and not relevant for our research question[13].

To correctly assess potential effects of retirement on mental health we need to clearly define what makes an individual retired. In their review of global retirement literature Nishimura et al. (2020) find that models are not particularly sensitive to the type of retirement definition used. Nonetheless, a single definition of retirement could potentially drive the results, so we employ three of them. Our first definition classifies individuals as retired if they claim to be retired. In the SHARE survey, individuals can define themselves as retired, employed, self-employed, permanently ill or disabled, homemaker, or engaged in other activities. Under this definition, everyone, no matter whether they are participating in the labour force or not, can

[12] Although it is often the case that some people continue to work whilst claiming pension payments.

[13] Other problems when using early retirement ages lie with different age thresholds for different occupations, within and across countries, as well as with different lengths of labour force participation. Therefore, it is difficult to ascertain one singular early retirement eligibility age for each country sample.

respond to retirement incentives and retire once they've reached the state pension age (Heller-Sahlgren, 2012, 2017). The dummy variable for definition 1 thus distinguishes between officially retired individuals and those not officially retired and belonging to other categories. Under the second definition, respondents who claim to be retired or out of the work force (homemakers, permanently ill or disabled, engaged in other activities) are counted as retired as long as they do not work part time (Eibich, 2015; Kolodziej & García-Gómez, 2019) which in the SHARE survey respondents indicate as having not done any paid work in the last four weeks. Thus, in retirement definition 2, individuals who are not officially retired, and work part time, are not defined as retirees. In the third definition, respondents are either retired, in the labour force (employed, self-employed), or out of the labour force (permanently ill or disabled, homemakers, engaged in other activities) (Apouey, Guven & Senik, 2019; Heller-Sahlgren, 2017). Third definition is the most restrictive one out of the three and significantly narrows down the number of observations. These three definitions of retirement are used as dummy variables indicating whether a person is retired or not, and we perform our main estimation as three separate equations with said definitions.

4.4 Mental health

Many studies have tried to assess retirement's impact on physical health. However, health variables are often self-reported and thus suffer from measurement errors (French & Jones, 2017; Motegi, Nishimura & Terada, 2016), or justification bias if respondents claim themselves to be in bad health to justify their retirement decision (Heller-Sahlgren, 2012). Additionally, retirement is an event with a profound impact on individuals' lifestyle. Thus, researchers have argued that mental wellbeing is more likely to be the vulnerable aspect of a person's health post retirement (Fé & Hollingsworth, 2012). For a measure of mental health to be comparable in a cross-country environment, the primary concern is to ensure that respondents with different cultural backgrounds interpret the question in the same manner (Maskileyson, Seddig & Davidov, 2021). Any differences in norms between countries should not be allowed to bias the measurement of mental health, more so when the symptoms are self-reported (Prince et al., 1999). To ensure validity, SHARE uses the Euro-D scale which is developed by European consortium in order to compare symptoms of depression across European countries.

The Euro-D scale consists of twelve questions that examine whether respondents have experienced a certain symptom of depression: pessimism, guilt, problems with sleep, irritability, appetite loss, fatigue, concentration problems, lack of enjoyment, tearfulness, depression, suicidal thoughts, and loss of interest (Prince et al., 1999). By indicating presence of a symptom, the respondent moves up on the Euro-D scale, with as little as four symptoms present being enough to mark the respondent as clinically depressed. By being a subjective measure, Euro-D is still weak to measurement error (French & Jones, 2017), but the scale has been cleared for use in cross-national comparative research as it is internally consistent, reliable, and accurate (Maskileyson, Seddig & Davidov, 2021). With the scale already incorporated in the SHARE questionnaire and used to evaluate depressive symptoms, we follow the same approach in our model and construct two measures of mental health: first is a cumulative score on the Euro-D scale which can range from 0, indicating no depression symptoms present, to 12, indicating severe depression. The second measure of mental health employed is a dummy variable indicating whether the respondent has reached the threshold of clinical depression, a condition prescribed to anyone scoring four or more on the Euro-D scale (Heller-Sahlgren, 2017; Lusa & Huebner, 2021).

4.5 Control variables and heterogeneity

Control variables are used in the second part of this study for the purpose of evaluating potential heterogeneous effects of retirement on mental health, and perhaps obtaining more information on the causal pathways through which retirement might affect mental health (Heller-Sahlgren, 2017; Heß et al., 2021). In the 2017 article, Nishimura et al. give guidance to future researchers to take educational differences into consideration. Referring back to Section 2 and the theory of human capital by Grossman (1972), we assume there is a likely relationship between educational attainment levels and health investment behaviour (Bhattacharya, Hyde & Tu, 2018, pp.28-47; Fe & Hollingsworth, 2012). Hence one of the covariates is education level, classified following ISCED-1997 classification, a self-reported measure of highest educational attainment (International Standard Classification of Education (ISCED), 2022). Whereas ISCED-1997 classifies respondents into categories of low, medium, and high education, we choose to focus only on the effect of low educational attainment. It has been implied by Leopold & Engelhardt's (2012) paper that health disparity between educational levels increases with age and thus we investigate whether a low educational attainment may affect mental wellbeing later in life. Furthermore, we evaluate whether the impact of retirement differs based on the marital status of the respondents. Currently, the literature focuses on the effects of retirement on marital satisfaction as well as on cross-gender spillover effects of spouses' retirement (Messe & Wolff, 2018; Zang, 2020). Article by Dave et al (2008) concludes that married people are physically healthier after retirement, but a definitive answer on the effects of marital status on mental health is yet to be presented. Lastly, it is without a doubt that impact of retirement on mental health varies across countries, with cultural, environmental, and geopolitical differences being implicit or explicit reasons as to why. Thus, we consider that respondent's country of residence influences their wellbeing given their age as well as at their retirement age, and account for that by including country dummies in the model.

4.6 Attrition

When using a longitudinal dataset, one of the main concerns is always selective attrition (Dave, Rashad & Spasojevic, 2008). Attrition happens when observations drop out of the sample, which is not uncommon in panel data sets as they span over a longer time period. As long as attrition is random it does not pose a problem to researchers, but if it is not and happens systematically, it may subject the data set to sample selection bias (Miller & Hollist, 2007). For example, observations lost due to attrition might differ from the remaining ones by a specific characteristic which invertedly causes biased estimates. As expected, the SHARE panel exhibits a considerable attrition rate. Approximately 37 percent of those interviewed in the fourth wave are still present in wave seven, corresponding to an attrition rate of approximately 63 percent.

To evaluate whether the main results are robust to panel attrition, we introduce an inverse probability weighting technique which allows for attrition to be non-random, conditional on observable characteristics. This technique is built upon the premise that disappearing from the sample is related to the outcome variable by a vector of endogenous observable variables c – also known as selection on observables (Fitzgerald, Gottschalk & Moffitt, 1998). Thus, we start by testing which covariates may affect respondent's probability of remaining in the sample after which we use the inverse of those probabilities[14] as weights in our main model. Bigger weights are given to observations who we have predicted will leave the sample early, with the rationale being that if we successfully predicted attrition, the weighted model's estimates will not differ greatly from estimates in the main regression as the only attrition left in the model will be random (Fitzgerald, Gottschalk & Moffitt, 1998). The variables chosen as observable characteristics indicative of the probable attrition are age, gender, marital status, employment status, educational attainment, EURO-D score, self-assessed physical health, BMI, number of limitations with activities of daily living, numeracy score, number of chronic diseases, grip strength, and number of mobility limitations[15]. We use values from wave four, the first wave

[14] Weight = 1/probability of survival.

[15] Observable characteristics consist of baseline characteristics and time-varying factors that predict survival status (Dave, Rashad & Spasojevic, 2008).

in which respondents participated, to predict their probability of remaining until wave seven (Comi, Cottini & Lucifora, 2022). The delimitation of this method is that we cannot be certain we have accounted for all covariates that may influence a person's ability to leave the survey, hence we cannot confirm we have completely gotten rid of attrition bias. However, if the results of our robustness check prove to be in line with the main model's results, we would have proven that attrition isn't a debilitating issue for our model.

5. Empirical Analysis

To successfully ensure a causal interpretation of the research question of interest, we must address issues of endogeneity, causal effects, and time dimension analysed. This section provides detailed insight into the construction of the research model used to accomplish this.

5.1 Endogeneity problem and model set-up

The straightforward way to measure the causal relationship between retirement and mental health would be to employ a standard OLS and evaluate in which way retirement might predict mental health. For this approach to be valid we must assume that $Cov\ (X_i, u_i) = 0$ holds[16]. However, as previous literature on the topic has noted, trying to estimate the causal relationship between retirement and mental health via OLS will not work on the account that the act of retirement is not random (Coe & Zamarro, 2011; Dave, Rashad & Spasojevic, 2008; Eibich, 2015; Picchio & van Ours, 2019), which will cause the model to be miss-specified (Angrist & Pischke, 2009, pp.113-138). The standard OLS model is exposed to three potential sources of endogeneity. Firstly, unobserved covariates may jointly affect retirement and mental health which gives rise to the omitted variables problem. Secondly, while it is presumed that retirement affects mental health, the causal relationship might be two-way[17], implying that endogeneity is caused by reverse causality. Lastly, any measurement errors in the independent variable will give rise to biased coefficients. Since the retirement decision is presumed to be a non-random act, the OLS approach is not fit to account for these potential threats of endogeneity, neither can it ensure a valid causal interpretation of our findings. Instead, we must employ a research strategy which provides an exogenous variation in the retirement probability.

[16] OLS is valid if the vector of explanatory covariates is not correlated with the error term, i.e. we have exogeneity. Here the variable for retirement is contained within the vector of explanatory variables X_i and is by assumption not correlated with the error term u_i.

[17] Entering retirement is often a choice one makes after assessing physical and/or mental health, so causality is difficult to measure (Coe & Zamarro, 2011).

As noted in Section 3, previous literature presents inconsistent findings on the mental health effects of retirement. These inconsistencies are strongly suggested to be caused by differences in estimation method (Nishimura, Oikawa & Motegi, 2017). By taking advantage of the longitudinal nature of the SHARE data, we identify favourable opportunities for employing a 'hybrid' identification strategy. Explicitly, we intend on utilizing an *individual-fixed effects* methodology within an *instrumental variables* approach, thus combining two quasi-experimental methods in order to ensure an improved causal interpretation of this long-debated research question. We believe the combination of different quasi-experimental methods in our 'hybrid' identification strategy to be highly feasible for causal inference, and thus remedy previous shortcomings among prevailing literature. Namely, utilization of individual-fixed effects is the premise on which we are able to distinguish between short-, medium-, and long-term effects of retirement. A fixed effects model on its own, however, is highly unlikely to allow for a credible causal interpretation of the mental health effects of retirement.

Consider this linear panel data model, in which we observe individuals across several time periods:

$$MH_{it} = \alpha + X_{it}\beta + \upsilon_{it}$$

$$i = 1, \dots, N \quad t = 1, \dots, T$$

where MH_{it} is our measure of mental health, X_{it} is a vector of exogenous regressors, and υ_{it} is the error term. According to Angrist & Pischke (2009), we may assume the error term υ_{it} to be the sum of two addends, such that:

$$\upsilon_{it} = \eta_i + \varepsilon_{it}$$

This implies that the error term υ_{it} can be separated into two parts; the part that varies over time, ε_{it}, and the part that does not vary over time, η_i. The latter reflects unobserved individual factors that are time-invariant, such as race, gender, genetics, ability, personality traits etc.
In a *fixed effects* model, presented below, assumptions of no endogeneity are relaxed. We assume that unobserved individual time-invariant factors η_i are indeed dependent on the values of X for all time periods T. This indicates that we do not dismiss correlation between η_i and X_{it} in any time period:

$$MH_{it} = \alpha + X_{it}\beta + \eta_i + \varepsilon_{it}$$

$$E[\eta_i | X_{i1}, \dots, X_{iT}] \neq 0$$

Nonetheless, although the zero conditional mean assumption is violated in this case, it may still be possible to obtain consistent estimates of the causal effect using *within estimation*. This is simply done by estimating the individual-specific mean values over time, and subtracting them from the main model, thus eliminating the unobserved individual time-invariant factors from our model and consequently getting rid of the fixed effects.

$$\overline{MH_i} = \bar{X}_i\beta + \bar{\eta}_i + \bar{\varepsilon}_i$$

$$MH_{it} - \overline{MH_i} => \widetilde{MH_{it}} = \widetilde{X_{it}}\beta + \widetilde{\varepsilon_{it}}$$

We can obtain the within-estimator, $\widehat{\beta_{within}}$, by employing an OLS strategy.
To summarize, by removing the fixed effects, η_i, we can control for all time-invariant individual factors, regardless of whether they can be measured or not. Our estimated effect of retirement, β, thus captures the within-subject change over time, rather than between-subject variability. This is the premise on which we are able to exploit different time-dimension in which the causal effect might operate.

What makes the fixed-effects estimator unfeasible, however, is that we are further required to assume that regressors are *strictly exogenous*, i.e. that the part of the error term that *does* vary over time, ε_{it}, is at all time periods unrelated to the value of the treatment indicator or other covariates:

$$E[\varepsilon_{it}|X_{i1}, \dots, X_{iT}, \eta_i] = 0$$

$$i = 1, \dots, N \quad t = 1, \dots, T$$

The strict exogeneity assumption is likely to fail as we cannot account for unobservable time-specific events that may affect both the outcome and the causal variable of interest, e.g. what we may perceive as an effect of retirement on mental health could likely be caused by some time-variant 'outside-model' effects that affect both mental health and are endogenous to retirement status. Thus, although the fixed effects estimator may potentially get rid of a large portion of the omitted variable bias in the model, it can only account for the bias caused by time-invariant unobservables. Hence, the fixed-effects estimator is still exposed to time-varying omitted variables and simultaneity bias, as well as susceptible to measurement errors which cause attenuation bias (Angrist & Pischke, 2009, pp.221-227).

Therefore, we combine the individual-level fixed effects strategy with an instrumental variables approach to rid the model of aforementioned endogeneity problems. In order to obtain a source of exogenous variation in retirement behaviour, we instrument retirement with reaching one's pension eligibility age as it has been noted that age thresholds, at which one becomes eligible for state pension benefits, create discontinuities in retirement probability. These discontinuities are then exploited to predict retirement behaviour, with this approach being proven to be feasible for causal inference on the relevant topic and well established in the prevailing literature investigating the mental health effects of retirement (Apouey, Guven & Senik, 2019; Coe & Zamarro, 2011; Eibich, 2015; Fletcher, 2014; Heller-Sahlgren, 2017; Heß et al., 2021; Messe & Wolff, 2018; Picchio & van Ours, 2019). We may therefore expect state pension eligibility ages to be good predictors of retirement behaviour. Worth of mention, however, is that state pension eligibility ages can only be a viable instrument if they meet the following assumptions (Angrist & Pischke, 2009, pp.113-138; Fe, 2020; Lousdal, 2018) presented below in Table 3.

Table 3: Validity of IV

Assumptions	Definition	Validity
Relevance assumption i.e. existence of first stage: $Cov(z, x) \neq 0$	Instrument has to have a causal effect on the dependent variable X.	State pension eligibility ages induce people to retire by providing a financial incentive in the form of pension benefits.
Exclusion restriction: $Cov(z, u) = 0$	Instrument is correlated with covariate of interest X but is uncorrelated with both observable and unobservable determinants of dependent variable Y.	State pension eligibility ages only affect mental wellbeing indirectly through retirement by serving as a proxy for retirement decision, not through any other channels.
Independence assumption	Instrument is independent of the outcome variable Y.	State pension eligibility ages should have no effect on mental health as they are a government-imposed threshold and cannot be manipulated.
Monotonicity assumption	Observations affected by the instrument must all be affected in the same way.	Everyone who reaches the state pension eligibility age has an increased probability of retiring.

It is important to note that the financial incentives to retire, which become available upon reaching legal retirement ages, are not powerful enough to force individuals into retirement. Instead, reaching the threshold merely creates discontinuities at which respondent's probability of entering retirement significantly increases (Picchio & van Ours, 2019). We can, therefore, apply a fuzzy regression discontinuity intuition to our model, where actual retirement is instrumented by discontinuities in retirement probability and age acts as the 'assignment' variable in a Two-Stage Least Squares model.

Discontinuities created by legal retirement ages are defined as following:

$$Pr(R_i = 1 \mid age_i) = \begin{cases} f_1(age_i) \; if \; age_i \geq pension\_age_i \\ f_0(age_{i)} \; if \; age_i \; < pension\_age_i \end{cases}$$

$$\text{where } f_1(pension\_age_i) \neq \; f_0(pension\_age_i)$$

The above expression posits that probability of the respondent being retired given their age is higher if the respondent's age is greater than or equal to the state pension age[18]. In the model we use the dummy variable of pension age to instrument retirement:

$$\overline{pa}_i = 1 \; if \; age_i \geq pension\_age_i \text{ otherwise } \overline{pa}_i = 0$$

where dummy variable $\overline{pa}_i$ indicates whether the respondent has reached state pension age or not. Essentially, in what follows, we are able to combine individual-fixed effects with the fuzzy regression discontinuity intuition, allowing the mental health effects of retirement to differ with time.[19] Estimating the lagged, long-term effects of retirement on mental health, we conduct the following 2SLS estimation:

$$\begin{aligned} R_{it-2} = \; & \beta_0 + \beta_1 \overline{pa}_{it-2} + \beta_2(\widetilde{age}_{it-2}) + \beta_3(\widetilde{age}_{it-2})^2 + \beta_4\gamma_c(\widetilde{age}_{it-2}) \\ & + \beta_5\gamma_c(\widetilde{age}_{it-2})^2 + \beta_6\overline{pa}_{it-2}(\widetilde{age}_{it-2}) + \beta_7\overline{pa}_{it-2}(\widetilde{age}_{it-2})^2 \\ & + \beta_8\gamma_c[\overline{pa}_{it-2}(\widetilde{age}_{it-2})] + \beta_9\gamma_c[\overline{pa}_{it-2}(\widetilde{age}_{it-2})^2] + \delta_i + \eta_t + \theta_t + \varepsilon_{it} \end{aligned} \quad (1)$$

$$\begin{aligned} euro_{d_{it}} = \; & \beta_0 + \beta_1\widehat{R_{it-2}} + \beta_2(\widetilde{age}_{it-2}) + \beta_3(\widetilde{age}_{it-2})^2 + \beta_4\gamma_c(\widetilde{age}_{it-2}) \\ & + \beta_5\gamma_c(\widetilde{age}_{it-2})^2 + \beta_6\overline{pa}_{it-2}(\widetilde{age}_{it-2}) + \beta_7\overline{pa}_{it-2}(\widetilde{age}_{it-2})^2 \\ & + \beta_8\gamma_c[\overline{pa}_{it-2}(\widetilde{age}_{it-2})] + \beta_9\gamma_c[\overline{pa}_{it-2}(\widetilde{age}_{it-2})^2] \\ & + \delta_i + \eta_t + \theta_t + \varepsilon_{it} \end{aligned} \quad (2)$$

We estimate the first stage regression (equation 1), where we predict retirement behaviour between wave 4 and 5 by reaching one's state pension eligibility age, $\overline{pa}_{it-2}$, over that same period. The presence of flexible continuous functions of age, $\widetilde{age}_{it-2}$ and $(\widetilde{age}_{it-2})^2$, is noticeable in the first- and second-stage as it is recommended to let the impact of the assignment variable, in this case the age of the respondent, differ on each side of the threshold (Angrist & Pischke, 2009, pp.251-269). We do so by centring age, so that the coefficient of the causal variable of interest is able to estimate the jump in mental health at the threshold. This is done by subtracting pension eligibility age from respondent's actual age[20] (Ashwin, Keenan & Kozina, 2021; Heller-Sahlgren, 2012). We include linear and quadratic age trends as they have been observed to have substantially smaller standard errors than models with high order

[18] In other words, reaching state pension age makes retiring more probable, for which we must also assume the following: $f_1(pension\_age_i) > f_0(pension\_age_i)$

[19] Whereas a limitation of using an RDD is that it can only estimate short-run effects of retirement (Nielsen, 2018).

[20] Time-demeaning is essential for estimating fixed effects (Angrist & Pischke, 2009, pp.221-227), thus $\widetilde{age}_{it-2}$ and $(\widetilde{age}_{it-2})^2$ are shorthand for $(age_{it-2} - pa_{it-2})$ and $(age_{it-2} - pa_{it-2})^2$ respectively.

polynomials (Gelman & Imbens, 2019). Further, we add interactions between centred age variables and the legal retirement age threshold: $\overline{pa}_{it-2}(\widetilde{age}_{it-2})$ and $\overline{pa}_{it-2}(\widetilde{age}_{it-2})^2$, with the purpose of isolating the effect of retirement by discerning between the nonlinearity of continuous variable of age and the discontinuous jump in retirement probability happening at $\overline{pa}_i$ (Lee & Lemieux, 2010; Picchio & van Ours, 2019). The variables should not exhibit any kink, otherwise our results will be confounded by variation stemming from covariates other than the retirement status (Fe & Hollingsworth,2012).

In the second-stage regression (equation 2) we regress the mental health measure $euro\_d_{it}$ on the predicted probability of retirement between wave 4 and 5 from the first-stage regression, denoted as $\widehat{R_{it-2}}$. We recognize that the effect of age may differ across countries[21] and we counter that by incorporating interactions between country dummy variables and centred ages, $\gamma_c(\widetilde{age}_{it-2})$ and $\gamma_c(\widetilde{age}_{it-2})^2$, as well as their interaction with the state pension eligibility threshold; $\gamma_c[\overline{pa}_{it-2}(\widetilde{age}_{it-2})]$ and $\gamma_c[\overline{pa}_{it-2}(\widetilde{age}_{it-2})^2]$. By doing so, we allow the general effect of age, as well as the effect of age around the discontinuity, to differ across observed countries. The model includes individual-fixed effects, δ_i, as well as year- and month-fixed effects, denoted by η_t and θ_t respectively. Year- and month-fixed effects are added to control for yearly and monthly factors of change that are common within countries (Allison, 2009, n.p.), with the common example being the need to account for the effects of seasons.

Thus, this model is essentially analysing the long-term longitudinal changes in mental health (wave 5-7) after a change in retirement status has been established between wave 4 and 5. The intuition is that by measuring all variables, except mental health, at $t-2$, we capture the lagged, long-term effect on mental health. Furthermore, in separate estimations, we control for lagged mental health in the models alongside the variables mentioned above for two reasons: it prevents mean reversion since we're controlling for individual's mental health in previous waves (Angrist & Pischke, 2009, pp.243-246; Heller-Sahlgren, 2017), and secondly, it serves as a further check on our model's ability to capture causal effects of retirement as coefficients between two models should not differ greatly. All models are estimated with standard errors clustered at the individual level as we assume that there is within-individual correlation. Intuitively, we estimate the medium-term effects by restricting the sample to observations in waves 4-6 and re-estimating equations (1) and (2), only now all regressors are measured at $t-1$, instead of at $t-2$. Same logic applied, we estimate all regressors at time t

[21] To rid the model of correlation between ages and country-specific pension system (Motegi, Nishimura & Oikawa, 2020).

for the short-term effects of retirement, where the sample is restricted to observations in waves 4 and 5.

By exploiting lags of relevant variables as well as multiple wave observations in our longitudinal data we are able to alternate the window of cause and effect across the four years after the change in retirement status took place, observing within-unit changes at multiple points in time (Fe & Hollingsworth, 2012). Our FE-IV models, based on a fuzzy RDD intuition, capture the local average treatment effect (LATE) of retirement on mental health (Imbens & Angrist, 1994; Zang, 2020). In the set-up of this identification strategy, given the four assumptions provided in Table 3, the LATE can be interpreted as the effect of retirement on mental health for those whose retirement status was changed by the instrument, i.e. compliers[22].

Limitations of this model setup are few but important to address. Based on the approximately two-year window between the wave four and five interviews, there is a potential risk of misspecification in the model. Namely, individuals assigned to the treatment group cross their state pension eligibility age at different times between wave four and five, depending on country and gender. Thus, for those who retire immediately after the wave four interview, mental health will not be reported until the wave five interview, happening almost two years after the change in retirement status took place. Intuitively, this may cause our estimators to ignore potential short-term effects arising immediately after the change in retirement status. We firmly believe, however, that this potential risk of misspecification is significantly mitigated by the inclusion of lagged mental health, which should account for any differentials in mental health trends between compliers in the treatment and control groups over the fourth and fifth wave interviews.

5.2 Robustness tests

First robustness test conducted pertains to narrowing down the age window of our working sample at the fifth wave interview. In the benchmark specifications we include individuals aged 50-76 years at wave 5, which corresponds to a bandwidth of 10 years below and above the minimum and maximum state pension eligibility ages presented in Table A1. Thus, to determine the sensitivity of our results to the choice of bandwidth, we restrict the sample to individuals aged 57-69 years old at wave 5. This corresponds to a bandwidth of 3 years below

[22] Local average treatment effect for respondents affected by the instrument means both our treatment group (retired because they reached state pension age in-between waves four and five) and our control group (didn't reach state pensione eligibility age in that time so they didn't retire) are compliers in this scenario, and we compare changes in mental health between two groups.

and above the minimum and maximum state pension eligibility ages. Narrower age window should decrease the likelihood of bias, however, it is also possible that by having fewer observations we increase the variance (Heller-Sahlgren, 2017; Lee & Lemieux, 2010). The smaller age window is estimated with and without quadratic age trend, an intuition borrowed from the RD design. It has been noted that RD estimates get less precise as the sample around the discontinuity gets smaller, and to counter this it is recommended to lower the number of polynomials included in the model (Angrist & Pischke, 2009, pp.251-269).

To conclude, we investigate potential heterogeneous effects of retirement depending on marital status, gender, and educational attainment. We include interaction variables between our instrument and chosen covariates in separate estimations to investigate potential differences in the impact of retirement on mental health depending on marital status, gender, and education.

6. Results

This section presents the estimated *impact of retirement on mental health*, based on three definitions of retirement and two separate outcome measures. We provide the main results of this paper in Subsection 6.1 by analysing the short-, medium-, and long-term effect of retiring between interview waves four and five, where the endogenous retirement decision is instrumented by state pension eligibility age. In Subsection 6.2, we evaluate the robustness of our results by examining their sensitivity to (1) choice of bandwidth and (2) panel attrition. This section concludes with Subsection 6.3, where we explore heterogenous effects of retirement.

6.1 The impact of retirement on mental health

As noted in Section 5, the retirement decision is likely to be endogenous. Hence, inconsideration of this poses great threats to the validity of our results. Table 4 displays estimates of the impact of retirement on mental health utilizing OLS models. While these models incorporate month-, year-, and individual-fixed effects, they yet cannot account for endogenous retirement behaviour across individuals. Consistent with a greater part of previous literature, discussed in Section 3, the coefficients are noticeably small and statistically insignificant independent of outcome measure and definition of retirement. Thus, our OLS estimates suggest that mental health is unaffected by retirement in the long-, medium-, and long-term respectively.

Table 4: OLS results

Definition of Retirement	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Outcome Variable	*Euro-D*	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Short-term relationship						
R_{it}	0.04 (0.05)	0.03 (0.05)	0.09 (0.06)	-0.01 (0.01)	-0.01 (0.01)	-0.01 (0.01)
Medium-term relationship						
R_{it-1}	-0.07 (0.01)	-0.01 (0.03)	-0.04 (0.04)	-0.01 (0.01)	0.00 (0.01)	-0.01 (0.01)
Long-term relationship						
R_{it-2}	0.04 (0.06)	0.07 (0.06)	0.14 (0.07)	0.01 (0.02)	0.02 (0.02)	0.03 (0.02)
n	15 748	15 748	13 818	15 748	15 748	13 818

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors clustered at the individual level are in parentheses. Year-, month-, and individual-fixed effects are included in all models. Country dummy interactions are included in all models. Linear and quadratic age trends are included in all models.

As discussed in Subsection 5, we deal with the endogeneity problem by utilizing our main FE-IV research strategy. Table 5 displays the estimates obtained from our 2SLS model in equations (1) and (2). Short-, medium-, and long-term effects are displayed in panels 1-5. The first row of each panel presents second-stage results, whereas the second row of each panel presents first-stage results. Based on the results of the second row of each panel, assessed by the coefficient of the instrument $\overline{pa}_i$, we can establish that the instrument has a sufficiently strong first-stage effect across all definitions of retirement, with the F-statistics continuously exhibiting values significantly higher than the critical values established by Stock and Yogo (2005) for non-weak instruments.[23] Depending on the definition of retirement, the first-stage coefficients suggest that reaching the state pension eligibility age increases the probability of retirement by 10-13 percentage points, respectively statistically significant at a 1 percent significance level. Thus, reaching one's state pension eligibility age meets expectations of being a strong predictor of retirement, which in turn implies an ample variation in the retirement probability within the working sample.

The second-stage results of the first panel indicate a fairly small, but positive, association between retirement and mental health in the short-term across all definitions of retirement. These indications hold true for both outcome measures. In absence of statistical significance, however, these coefficients display no evidence of short-term effects on mental health following a change in retirement status between the fourth and fifth wave interviews, assessed by the coefficient of $\widehat{R_{it}}$. As mentioned in Section 2, however, an important question while

[23] Our F-statistics are consistently higher than the Stock & Yogo threshold bias of 5% so we conclude our state pension eligibility age is a sufficiently strong instrument.

assessing the mental health effects of retirement is when those effects might appear, and the extent to which they may persist. Therefore, in what follows we analyse the potential mental health effects approximately 2-4 years after retirement has ensued. It could very well be that causal inference is not achieved in previous literature attributable to inconsideration of time-dimensions in which the impact of retirement may operate.

Contrary to short-term implications, the second-stage results of the second panel suggest a negative medium-term association between retirement and mental health across all definitions of retirement, somewhat mitigated by the models in the third panel including lagged mental health. Correspondingly to the short-term estimates, however, the second-stage coefficients exhibit no statistical significance and allow for no causal interpretation of the mental health effects of retirement, regardless of outcome measure. Thus, these results display no evidence of any effects on mental health over the fifth and sixth wave interviews following a change in retirement status over the fourth and fifth wave interviews, assessed by the coefficient of $\widehat{R_{it-1}}$. The second-stage results of the fourth panel, however, indicate a substantial negative long-term effect on mental health following a change in retirement status over the fourth and fifth wave interviews, assessed by the coefficient of $\widehat{R_{it-2}}$. Depending on which definition of retirement is employed, the second-stage coefficients display a long-term average increase of 1.14-1.57 Euro-D points among those who retired between the fourth and fifth wave interviews, relative to the control group. This corresponds to 0.52-0.72 standard deviations, based on summary statistics for wave 7, as seen in Table 2. Similarly, substantial negative long-term effects of retirement are found when analysing the probability of clinical depression, which is denoted as scoring 4 or higher on the Euro-D scale. Depending on which definition of retirement is employed, the probability of becoming or remaining clinically depressed increases on average by 19-26 percentage points for those who retire between the fourth and fifth wave interviews, relative to the control group. This corresponds to 0.45-0.61 standard deviations. These results hold true in the models including lagged mental health in the fifth panel, although the negative effects are slightly mitigated, and somewhat more precise. Nonetheless, the long-term effects of retirement on mental health are statistically significant and uniformly negative, irrespective of outcome measure, model specification, and which definition of retirement is employed. Worth of mention, the absence of statistically significant short- to medium-term effects on mental health has an important implication. With no evidence of any positive effects on mental health prior to the long-term estimation, one can arguably dismiss any '*honeymoon-stage*' effects[24].

Table 5: FE-IV models – Mental health effects of retirement

Definition of Retirement	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Outcome Variable	*Euro-D*	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Short-term effects						
$\widehat{R_{it}}$ *Second-stage*	-0.11 (0.44)	-0.16 (0.60)	-0.14 (0.54)	-0.05 (0.10)	-0.07 (0.14)	-0.06 (0.12)
$\overline{pa}_{it}$ *First-stage*	0.13*** (0.01)	0.09*** (0.02)	0.10*** (0.01)	0.13*** (0.02)	0.09*** (0.02)	0.10*** (0.01)
F-statistic	78.36	46.88	57.17	78.36	46.88	57.17
Hausman test	0.58	0.51	0.33	0.52	0.50	0.30
n	15 748	15 748	13 818	15 748	15 748	13 818
Mid-term effects						
Excluding lagged mental health						
$\widehat{R_{it-1}}$ *Second-stage*	0.42 (0.41)	0.71 (0.57)	0.64 (0.51)	0.04 (0.10)	0.05 (0.13)	0.05 (0.12)
$\overline{pa}_{it-1}$ *First-stage*	0.13*** (0.01)	0.09*** (0.02)	0.10*** (0.01)	0.13*** (0.01)	0.09*** (0.01)	0.10*** (0.01)
F-statistic	78.30	46.45	56.97	78.30	46.45	56.97
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Including lagged mental health						
$\widehat{R_{it-1}}$ *Second-stage*	0.26 (0.38)	0.56 (0.52)	0.51 (0.47)	0.03 (0.09)	0.04 (0.11)	0.03 (0.11)
$\overline{pa}_{it-1}$ *First-stage*	0.13*** (0.01)	0.09*** (0.02)	0.10*** (0.01)	0.13*** (0.01)	0.09*** (0.02)	0.10*** (0.01)
F-statistic	78.36	46.43	56.95	78.26	46.43	56.92
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	15 748	15 748	13 818	15 748	15 748	13 818
Long-term effects						
Excluding lagged mental health						
$\widehat{R_{it-2}}$ *Second-stage*	1.14*** (0.42)	1.57*** (0.58)	1.42*** (0.52)	0.19** (0.11)	0.26** (0.13)	0.23** (0.09)
$\overline{pa}_{it-2}$ *First-stage*	0.13*** (0.01)	0.09*** (0.02)	0.10*** (0.01)	0.13*** (0.01)	0.09*** (0.02)	0.10*** (0.01)
F-statistic	78.19	46.38	56.89	78.19	46.38	56.89
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Including lagged mental health						
$\widehat{R_{it-2}}$ *Second-stage*	1.10*** (0.37)	1.54*** (0.50)	1.38*** (0.45)	0.16** (0.09)	0.22* (0.12)	0.20* (0.11)
$\overline{pa}_{it-2}$ *First-stage*	0.13*** (0.01)	0.09*** (0.02)	0.10*** (0.01)	0.13*** (0.01)	0.09*** (0.02)	0.10*** (0.01)
F-statistic	78.13	46.36	56.85	78.15	46.36	56.94
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	15 748	15 748	13 818	15 748	15 748	13 818

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 (1) and (2).

[24] As discussed in Section 2, retirees are initially expected to go through the "honeymoon stage" and perhaps perceive retirement as vacation rather than a permanent withdrawal from the labour force. During this stage, we may expect retirees to feel carefree and relaxed, and hence experience positive changes in their mental health.

This in turn indicates that the coefficients of the fourth panel are capturing true negative long-term effects of retirement, with no threats of the effect being triggered by diminishing positive short- to medium-term effects. This implication is also supported by the long-term estimates of the fifth panel, where models including lagged mental health exhibit highly comparable estimates with somewhat higher precision. We identify, therefore, no threats of mean reversion. Importantly, the Hausman specification test uniformly rejects the null hypothesis of no endogeneity in the models assessing mid- to long-term effects. This implies that OLS estimates in Table 4 are underestimating the long-term effects of retirement and are to be perceived as downwards biased. Ultimately, our main results strongly imply that retirement has substantial negative long-term effects on mental health, although no significant short- to mid-term effects.

6.2 Robustness checks

6.2.1 Sensitivity to the choice of bandwidth

As discussed in Section 2, it is vital to investigate if the results are robust to changes in the size of the bandwidth around the threshold.[25] Thus, exploring the implications of the bandwidth choice, we re-estimated our FE-IV models by restricting the working sample to individuals within an age interval of approximately three years above and below the maximum and minimum age threshold at the fifth wave interview. Doing so, we only include individuals who were 57-69 years old at that point in time. Given the considerably smaller neighbourhood of observations around the thresholds, we employed a linear age trend. Consequently, the working sample is reduced by approximately 40 percent. As displayed by the estimates of Table 6, even when significantly reducing observation window, the negative long-term results remain considerable and are highly comparable to the benchmark specifications. These results hold true even when combining a quadratic age trend with the narrowed bandwidth, displayed by the estimates of Table A2. Evidently, the results are robust to the choice of bandwidth. The sample size significantly changes, but the basic causal inference does not. Establishing the internal validity of these results, our research design proves to capture the causal effects of retirement on mental health for the population under study.

[25] As discussed in Section 5, there is a common trade-off between bias and precision when determining the bandwidth since a broader bandwidth provides more observation and consequently larger precision/information, but observations further away from the thresholds are less comparable, hence increasing the bias.

Table 6: FE-IV models – Mental health effects of retirement (3-year age window)

Definition of Retirement	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Outcome Variable	*Euro-D*	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Short-term effects						
$\widehat{R_{it}}$ *Second-stage*	-0.02 (0.31)	-0.02 (0.40)	-0.07 (0.38)	-0.01 (0.07)	-0.03 (0.12)	-0.02 (0.09)
$\overline{pa}_{it}$ *First-stage*	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)
F-statistic	132.42	87.11	96.64	132.42	87.11	96.64
Hausman test	0.87	0.88	0.77	0.86	0.83	0.79
n	9 365	9 365	8 282	9 365	9 365	8 282
Mid-term effects						
Excluding lagged mental health						
$\widehat{R_{it-1}}$ *Second-stage*	0.38 (0.30)	0.48 (0.38)	0.44 (0.36)	0.01 (0.06)	0.02 (0.09)	0.01 (0.08)
$\overline{pa}_{it-1}$ *First-stage*	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)
F-statistic	132.53	87.49	97.02	132.53	87.49	97.02
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Including lagged mental health						
$\widehat{R_{it-1}}$ *Second-stage*	0.36 (0.26)	0.46 (0.34)	0.42 (0.31)	0.01 (0.05)	-0.02 (0.08)	0.01 (0.07)
$\overline{pa}_{it-1}$ *First-stage*	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)
F-statistic	132.51	87.43	96.90	132.55	87.44	97.01
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	9 365	9 365	8 282	9 365	9 365	8 282
Long-term effects						
Excluding lagged mental health						
$\widehat{R_{it-2}}$ *Second-stage*	1.10*** (0.36)	1.39*** (0.37)	1.29*** (0.30)	0.20*** (0.05)	0.26*** (0.10)	0.23*** (0.08)
$\overline{pa}_{it-2}$ *First-stage*	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)
F-statistic	132.78	87.95	96.93	132.78	87.85	96.93
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Including lagged mental health						
$\widehat{R_{it-2}}$ *Second-stage*	0.78*** (0.27)	1.01*** (0.33)	0.92*** (0.21)	0.13*** (0.04)	0.16*** (0.08)	0.15*** (0.05)
$\overline{pa}_{it-2}$ *First-stage*	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)
F-statistic	132.76	87.93	71.27	123.79	87.94	72.56
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	9 365	9 365	8 282	9 365	9 365	8 282

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 (1) and (2).

6.2.2 Inverse probability weighting

As noted in Subsection 4.6, although attrition is characteristic and to be expected in longitudinal research, we ought to treat it as detrimental to our results a priori. The idea is quite intuitive: if reasons of attrition are unknown, the same applies for its nature. Frankly, without accounting for it, there is no way of knowing whether attrition is random or systematic, and therefore, whether it poses a threat to the internal and-/or external validity of our findings. In order to account for possible systematic attrition, and thus any attrition bias, we assign different weights to each individual in our working sample depending on their inverse probability of remaining in all four waves of the panel, conditional on observables. That is, we evaluate the sensitivity of our benchmark specifications in Table 5 and Table 6 by re-estimating them, although now employing an inverse probability weighting on our working sample. Worth of mention, this approach slightly limits our sample size as a consequence of missing values on observables used to estimate each individual's inverse probability of remaining in all four waves of the panel. As Table A3 displays, the estimated long-term effects of retirement on mental health are highly comparable to those of Table 5 and Table 6, although larger and somewhat more precise. This holds true irrespective of outcome measure and definition of retirement. Indeed, this strongly indicates that we are dealing with random attrition, which further suggests that participants who left the panel do not differ in specific ways from those who remained. Ultimately, the validity of our results proves to be solid, and is unlikely to be affected by selective attrition.

6.3 Heterogeneity analysis

Evaluating the findings so far, we identify solid evidence of substantial negative long-term effects of retirement on mental health. However, based on inconsistent findings in previous literature discussed in Section 3, we have reasons to believe that the impact of retirement is not uniform across individuals hinged on demographic factors. Hence, we seek to obtain a better understanding of what factors potentially drive the results obtained in Section 6.1, and perhaps remedy some of the shortcomings of previous literature discussed in Section 3. Inconsideration of heterogenous retirement effects may very well explain the lack of consensus in previous findings. Table 7 displays long-term estimates of our FE-IV models using a 10-year age window, now allowing the impact of retirement to vary by educational attainment levels, gender, and marital status. As presented by the second and third panels, we find no evidence of heterogenous retirement effects in terms of gender and marital status.

Table 7: Heterogenous treatment effects – the impact of retirement on mental health

Definition of Retirement	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Outcome Variable	*Euro-D*	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Long-term effects						
Educational attainment level						
$\widehat{R_{it-2}}$ *Second-stage*	1.19*** (0.43)	1.61*** (0.76)	1.42** (0.60)	0.19** (0.12)	0.25** (0.19)	0.22** (0.16)
$\boldsymbol{\widehat{R_{it-2}}}$ × ***LEA***	0.22*** (0.06)	0.16*** (0.05)	0.36*** (0.11)	0.05*** (0.02)	0.12*** (0.02)	0.10*** (0.06)
n	15 748	15 748	13 818	15 748	15 748	13 818
Gender						
$\widehat{R_{it-2}}$ *Second-stage*	1.39*** (0.52)	1.97*** (0.81)	1.74** (0.63)	0.18*** (0.12)	0.29*** (0.11)	0.24** (0.16)
$\boldsymbol{\widehat{R_{it-2}}}$ × ***Female***	-0.09 (0.42)	0.16 (0.74)	0.11 (0.56)	-0.08 (0.10)	0.13 (0.18)	0.07 (0.12)
n	15 748	15 748	13 818	15 748	15 748	13 818
Marital status						
$\widehat{R_{it-2}}$ *Second-stage*	1.02*** (0.51)	1.36*** (0.78)	1.14** (0.58)	0.09** (0.13)	0.17** (0.14)	0.16** (0.18)
$\boldsymbol{\widehat{R_{it-2}}}$ × ***Married***	-0.22 (0.46)	-0.27 (0.74)	-0.38 (0.59)	-0.09 (0.10)	-0.04 (0.15)	-0.07 (0.14)
n	15 748	15 748	13 818	15 748	15 748	13 818

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Standard errors in parentheses (clustered at individual level). All models include the variables in equations (1) and (2). All models include lagged mental health. The minimum Cragg-Donald F statistic is 28.51. The Hausman test displays p-values lower than 0.01.

The estimated coefficients are fairly small and lack statistical significance across all definitions of retirement. Contrary, we find solid evidence of heterogeneity in the long-term impact of retirement hinged on educational attainment levels, independent of retirement definition and outcome measure. Explicitly, the first panel displays that retirement over the fourth and fifth waves causes an increase of 0.16-0.36 Euro-D points score for individuals with low educational attainment in the long-term, relative to the counterpart. Likewise, the probability of becoming or remaining clinically depressed increases by 5-12 percentage points for individuals with low educational attainment in the long-run, relative to highly educated individuals. Generally, this implies that the negative long-term effects of retirement on mental health are homogeneous across genders and marital status, but do not apply uniformly across educational levels. As discussed in Section 2, one can expect more educated retirees to be more efficient producers of health in context of the Grossman model, since information concerning health-preserving inputs increases with schooling. Thus, based on these results, lower education arguably reveals itself as a potential pathway of the negative mental health effects of retirement.

7. Discussion

Having established that the 2SLS estimates obtained in Subsection 6.1 are robust to model specification, choice of bandwidth, and panel attrition, we direct our focus towards the validity of our identification strategy. Prior to any policy implications, we need to evaluate whether or not our 2SLS estimates are fit for a causal interpretation of the impact of retirement on mental health. Therefore, we consider this discussion as a critical part of our paper. In what follows, we assess the internal and external validity of the main results.

It has been noted that the 2SLS estimator yields consistent but biased estimates in samples with restricted size, given that the exclusion restriction holds. This suggests that the 2SLS estimator is able to yield estimates close to the causal effect of interest only when samples are large in size. Consequently, small samples cause the 2SLS estimator to differ significantly from the estimand of the full population. Worth of mention, the bias of the 2SLS estimator is largest in presence of *weak* instruments, implying low correlation between the instruments and endogenous regressors. This especially holds true if the instruments are many compared to the endogenous regressors. If instruments are many and weak, the 2SLS estimator is biased towards OLS (Angrist & Pischke, 2009, pp.173-188). Further, researchers have shown that weak instruments cause the 2SLS estimator to be biased even in a large sample, emphasizing the importance of a strong first-stage (Bound et al., 1995). An appropriate approximation of the 2SLS estimator bias, relative to the OLS estimator, is assumed to be the inverse of the F-statistic of the excluded instruments (Bound et al., 1995; Stock & Yogo, 2005; Olea & Pflueber, 2013). Assessing our main results, as presented in Subsection 6.1, all benchmark specifications of our 2SLS estimation display large F-statistics for the excluded instrument, slightly varying between definitions of the endogenous regressor. Corresponding F-statistics are obtained in our sensitivity analysis, where the instrument becomes even more powerful, as presented in Subsection 6.2. Establishing our instrument as strong, the F-statistics uniformly imply a trivial bias of the 2SLS estimator relative to the OLS estimator, given that the assumption of exclusion restriction is satisfied. On that note, we ought to address the assumption of exclusion restriction. As discussed in Subsection 5.1, the exclusion restriction is of major importance in terms of obtaining valid causal estimates in an instrumental variables approach. Namely, for a valid causal interpretation of the results, the instrument must be independent of potential outcomes, mental health, but not of the causal variable of interest, retirement. This would indicate that the instrument is as good as randomly assigned, conditional on covariates (Angrist & Pischke, 2009, pp.113-147). Unfortunately, there are no valid tests that may prove if the exclusion restriction is satisfied or not. Our 'hybrid' identification strategy allows for slightly different identification assumptions. While a conventional fuzzy RDD would assume that individuals just around the pension eligibility threshold differ only in terms of retirement probability, we assume that crossing one's pension eligibility threshold (which serves as instrument for retirement) has no impact on mental health for observations around the given threshold other than via retirement. This is even more likely to hold when a smaller bandwidth of observations around the threshold is employed. As discussed in Subsection 5.1, we may assume this after controlling for direct effects of age in flexible specifications of our FE-IV models, where we focus on within-unit changes over time rather than across across-unit variation. It is strongly implied that one should be cautious giving causal interpretations of estimates when the exclusion restriction is likely to be violated. Given the nature of our identification strategy, however, we have reasons to believe that the exclusion restriction is more likely to be satisfied than not. Nonetheless, the analysis identifies a causal effect for the population under study which survives an appropriate line for sensitivity checks. As causal inference is achieved for the working sample, we conclude a high internal validity of our results, as expected when estimating the local average treatment effect (Angrist & Pischke, 2009, pp.173-216). Commonly known among researchers, 2SLS estimates conventionally have high internal validity but low external validity. In contrast to the average treatment effect, a LATE is not informative regarding the effect on individuals whose treatment status is unaffected by the instrument, thereby minimizing the degree to which the results can be generalized. Then again, this is of concern only if we are interested in the average treatment effect on the entire population. Some researchers argue that the LATE could be misleading for policy implications even if it's consistently estimated by 2SLS, since it may substantially differ from the average treatment effect (Heckman & Urzúa, 2010). Others imply that this poses no problems since policies are unlikely to affect an entire population, thereby making LATE the parameter of interest (Imbens, 2010). In context of our research question, differences between the average treatment effect on a stable population and the LATE should be of no concern. Pension eligibility does not force individuals into retirement, but merely incentivizes them towards it. Intuitively, the effect on those whose retirement status changes because of the instrument, i.e., compliers, should lay the foundations for policy implications, as they are the ones actually affected by the policy. In any case, an accumulation of different LATEs from different subpopulations can help build evidence on the topic of interest.

8. Conclusion

In this study, we investigated effects of entering retirement on one's mental health by applying an individual-fixed effects model with instrumental variables approach on four waves of Survey of Health, Ageing, and Retirement in Europe. Our choice of model allowed us to apply RDD intuition and treat state pension eligibility ages as sources of discontinuities in order to obtain exogenous variation in retirement decision, whereas inclusion of individual-fixed effects allowed us to stratify the observed effects on mental health over a short-, medium-, and long-term period. Thus, this thesis addresses shortcomings of previous studies, whose research designs did not allow the effect of retirement to operate in different time-dimensions. The obtained results show that retirement has no significant effect on mental health of respondents in short- and medium-term, however, in long-term the effect is both significant and negative. We thus conclude that retirement has a lagged negative effect on mental health, measured both by the Euro-D scale of depression symptoms as well as by the probability of remaining or becoming clinically depressed. Additionally, our main results survive an appropriate line of robustness checks, and prove insensitive to model specification, choice of bandwidth, and panel attrition. By using panel data on the sample of ten European countries in period from 2011 to 2017, we were able to estimate a local average treatment effect and reach a conclusion that negative long-term effects of retirement are homogenous in terms of gender and marital status, but heterogenous across educational attainment levels. While it is still mostly unclear through which mechanisms the effect of retirement operates, this paper suggests that educational attainment is a factor of interest in the ongoing discussion, hence this particular channel of effect should be thoroughly investigated by future researchers. Conversely, we recognize that there may be other causal pathways through which retirement may affect mental health and therefore further research on this topic is advised.

Research conducted in this paper presents results with high internal validity, however, we argue that LATE's obtained by this research design can have broader implications on policy making. Specifically, policy makers should consider different time dimensions in which causal effects could operate and work to find a solution that alleviates the pressure on fiscal sustainability of pay-as-you-go pension systems while simultaneously postponing negative effects of retirement on mental health.

References

Allison, P. D. (2009). Fixed Effects Regression Models, [e-book] SAGE Publications, Available Online: https://books.google.hr/books?id=3UxaBQAAQBAJ.

Angrist, J. & Pischke, J.-S. (2009). Mostly Harmless Econometrics: An Empiricist's Companion, in *Mostly Harmless Econometrics: An Empiricist's Companion*.

Apouey, B. H., Guven, C. & Senik, C. (2019). Retirement and Unexpected Health Shocks, *Economics & Human Biology*, [e-journal] vol. 33, pp.116–123, Available Online: https://www.sciencedirect.com/science/article/pii/S1570677X18302776 [Accessed 17 July 2022].

Ashwin, S., Keenan, K. & Kozina, I. M. (2021). Pensioner Employment, Well-Being, and Gender: Lessons from Russia, *American Journal of Sociology*, [e-journal] vol. 127, no. 1, pp.152–193, Available Online: https://www.journals.uchicago.edu/doi/abs/10.1086/715150 [Accessed 23 July 2022].

Atchley, R. C. (1982). Retirement: Leaving the World of Work, *The ANNALS of the American Academy of Political and Social Science*, [e-journal] vol. 464, no. 1, pp.120–131, Available Online: https://doi.org/10.1177/0002716282464001011 [Accessed 23 May 2022].

Becchetti, L., Pelloni, A. & Rossetti, F. (2008). Relational Goods, Sociability, and Happiness, *Kyklos*, vol. 61, no. 3, pp.343–363.

Belloni, M., Meschi, E. & Pasini, G. (2016). The Effect on Mental Health of Retiring During the Economic Crisis, *Health Economics*, vol. 25 Suppl 2, pp.126–140.

Bertoni, M. & Brunello, G. (2017). Pappa Ante Portas: The Effect of the Husband's Retirement on the Wife's Mental Health in Japan, *Social Science & Medicine (1982)*, vol. 175, pp.135–142.

Bhattacharya, J., Hyde, T. & Tu, P. (2018). Health Economics, Bloomsbury Publishing.

Bonsang, E., Adam, S. & Perelman, S. (2012). Does Retirement Affect Cognitive Functioning?, *Journal of Health Economics*, vol. 31, no. 3, pp.490–501.

Celidoni, M., Dal Bianco, C. & Weber, G. (2017). Retirement and Cognitive Decline. A Longitudinal Analysis Using SHARE Data, *Journal of Health Economics*, vol. 56, pp.113–125.

Coe, N. B. & Zamarro, G. (2011). Retirement Effects on Health in Europe, *Journal of Health Economics*, vol. 30, no. 1, pp.77–86.

Cohen, S. (2004). Social Relationships and Health, *The American Psychologist*, vol. 59, no. 8, pp.676–684.

Comi, S. L., Cottini, E. & Lucifora, C. (2022). The Effect of Retirement on Social Relationships, *German Economic Review*, [e-journal] vol. 23, no. 2, pp.275–299, Available Online: https://www.degruyter.com/document/doi/10.1515/ger-2020-0109/html [Accessed 23 May 2022].

d'Hombres, B., Rocco, L., Suhrcke, M. & McKee, M. (2010). Does Social Capital Determine Health? Evidence from Eight Transition Countries, *Health Economics*, vol. 19, no. 1, pp.56–74.

Dave, D., Rashad, I. & Spasojevic, J. (2008a). The Effects of Retirement on Physical and Mental Health Outcomes, *Southern Economic Journal*, vol. 75, no. 2, pp.497–523.

Dave, D., Rashad, I. & Spasojevic, J. (2008b). The Effects of Retirement on Physical and Mental Health Outcomes, *Southern Economic Journal*, [e-journal] vol. 75, no. 2, pp.497–523, Available Online: https://www.jstor.org/stable/27751397 [Accessed 17 July 2022].

Earl, J. & Archibald, H. (2014). RETIREMENT PLANNING AND RESOURCES, *European Journal of Management*, vol. 14, pp.21–36.

Eibich, P. (2015a). Understanding the Effect of Retirement on Health: Mechanisms and Heterogeneity, *Journal of Health Economics*, vol. 43, pp.1–12.

Elgar, F. J., Davis, C. G., Wohl, M. J., Trites, S. J., Zelenski, J. M. & Martin, M. S. (2011). Social Capital, Health and Life Satisfaction in 50 Countries, *Health & Place*, vol. 17, no. 5, pp.1044–1053.

Fe, E. (2020). Pension Eligibility Rules and the Local Causal Effect of Retirement on Cognitive Functioning, SSRN Scholarly Paper, 2993152, Available Online: https://papers.ssrn.com/abstract=2993152 [Accessed 5 August 2022].

Fé, E. & Hollingsworth, B. (2012). Estimating the Eect of Retirement on Mental Health via Panel Discontinuity Designs, MPRA Paper, University Library of Munich, Germany, Available Online: https://econpapers.repec.org/paper/pramprapa/38162.htm [Accessed 18 May 2022].

Fé, E. & Hollingsworth, B. (2016). Short- and Long-Run Estimates of the Local Effects of Retirement on Health, *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, vol. 179, no. 4, pp.1051–1067.

Fitzgerald, J., Gottschalk, P. & Moffitt, R. (1998a). An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics, *The Journal of Human Resources*, [e-journal] vol. 33, no. 2, pp.251–299, Available Online: https://www.jstor.org/stable/146433 [Accessed 27 May 2022].

Fletcher, J. M. (2014). Late Life Transitions and Social Networks: The Case of Retirement, *Economics Letters*, [e-journal] vol. 125, no. 3, pp.459–462, Available Online: https://www.sciencedirect.com/science/article/pii/S0165176514003838 [Accessed 23 July 2022].

Folland, S. (2008). An Economic Model of Social Capital and Health, *Health Economics, Policy, and Law*, vol. 3, no. Pt 4, pp.333–348.

Fonseca, R., Kapteyn, A., Lee, J., Zamarro, G. & Feeney, K. (2014). A Longitudinal Study of Well-Being of Older Europeans: Does Retirement Matter?, *Journal of Population Ageing*, vol. 7, no. 1, pp.21–41.

French, E. & Jones, J. B. (2017). Health, Health Insurance, and Retirement: A Survey, SSRN Scholarly Paper, 3017727, Available Online: https://papers.ssrn.com/abstract=3017727 [Accessed 17 July 2022].

Gelman, A. & Imbens, G. (2019). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs, *Journal of Business & Economic Statistics*, [e-journal] vol. 37, no. 3, pp.447–456, Available Online: https://doi.org/10.1080/07350015.2017.1366909 [Accessed 24 July 2022].

Glass, T. A., de Leon, C. M., Marottoli, R. A. & Berkman, L. F. (1999). Population Based Study of Social and Productive Activities as Predictors of Survival among Elderly Americans, *BMJ : British Medical Journal*, vol. 319, no. 7208, pp.478–483.

Gorry, D. & Slavov, S. N. (2021). The Effect of Retirement on Health Biomarkers, *Economics & Human Biology*, vol. 40, p.100949.

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.

Heckman, J. J. & Urzúa, S. (2010). Comparing IV With Structural Models: What Simple IV Can and Cannot Identify, *Journal of econometrics*, [e-journal] vol. 156, no. 1, pp.27–37, Available Online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2861784/ [Accessed 18 May 2022].

Heller-Sahlgren, G. (2012). Work 'Til You Drop: Short- and Longer-Term Health Effects of Retirement in Europe, SSRN Scholarly Paper, 2153191, Available Online: https://papers.ssrn.com/abstract=2153191 [Accessed 17 July 2022].

Heller-Sahlgren, G. (2017). Retirement Blues, *Journal of Health Economics*, vol. 54, pp.66–78.

Heß, M., Naegele, L., Becker, L., Mäcken, J. & De Tavernier, W. (2021). Planned Retirement Timing in Europe: Are Europeans Adapting to the Policy of Extending Working Lives, *Frontiers in Sociology*, vol. 6, p.150.

Hessel, P. (2016). Does Retirement (Really) Lead to Worse Health among European Men and Women across All Educational Levels?, *Social Science & Medicine (1982)*, vol. 151, pp.19–26.

Horner, E. (2014). Subjective Well-Being and Retirement: Analysis and Policy Recommendations, *Journal of Happiness Studies*, vol. 15, no. 1, pp.125–144.

Imbens, G. W. (2010). Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009), *Journal of Economic Literature*, [e-journal] vol. 48, no. 2, pp.399–423, Available Online: https://www.aeaweb.org/articles?id=10.1257/jel.48.2.399 [Accessed 11 June 2022].

Imbens, G. W. & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects, *Econometrica*, [e-journal] vol. 62, no. 2, pp.467–475, Available Online: https://www.jstor.org/stable/2951620 [Accessed 12 May 2022].

International Standard Classification of Education (ISCED). (2022). , Available Online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=International_Standard_Classification_of_Education_(ISCED) [Accessed 3 June 2022].

Jokela, M., Ferrie, J. E., Gimeno, D., Chandola, T., Shipley, M. J., Head, J., Vahtera, J., Westerlund, H., Marmot, M. G. & Kivimäki, M. (2010). From Midlife to Early Old Age: Health Trajectories Associated with Retirement, *Epidemiology (Cambridge, Mass.)*, vol. 21, no. 3, pp.284–290.

Kesavayuth, D., Rosenman, R. & Zikos, V. (2016). RETIREMENT, PERSONALITY, AND WELL-BEING, *Economic Inquiry*, vol. 54, no. 2, pp.733–750.

Knoll, M. A. Z. (2011). Behavioral and Psychological Aspects of the Retirement Decision, *Social Security Bulletin*, vol. 71, no. 4, pp.15–32.

Kolodziej, I. W. K. & García-Gómez, P. (2019). Saved by Retirement: Beyond the Mean Effect on Mental Health, *Social Science & Medicine*, vol. 225, pp.85–97.

Larimore, T., Lindauer, M., Ferri, R. A. & Dogu, L. F. (2009). The Bogleheads' Guide to Retirement Planning, John Wiley & Sons.

Lee, D. S. & Lemieux, T. (2010). Regression Discontinuity Designs in Economics, *Journal of Economic Literature*, [e-journal] vol. 48, no. 2, pp.281–355, Available Online: https://www.jstor.org/stable/20778728 [Accessed 24 July 2022].

Lindeboom, M., Portrait, F. & van den Berg, G. J. (2002). An Econometric Analysis of the Mental-Health Effects of Major Events in the Life of Older Individuals, *Health Economics*, vol. 11, no. 6, pp.505–520.

Lousdal, M. L. (2018). An Introduction to Instrumental Variable Assumptions, Validation and Estimation, *Emerging Themes in Epidemiology*, [e-journal] vol. 15, no. 1, p.1, Available Online: https://doi.org/10.1186/s12982-018-0069-7 [Accessed 1 August 2022].

Lusa, L. & Huebner, M. (2021). Organizing and Analyzing Data from the SHARE Study with an Application to Age and Sex Differences in Depressive Symptoms, *International Journal of Environmental Research and Public Health*, vol. 18, p.9684.

Maskileyson, D., Seddig, D. & Davidov, E. (2021). The EURO-D Measure of Depressive Symptoms in the Aging Population: Comparability Across European Countries and Israel, *Frontiers in Political Science*, [e-journal] vol. 3, Available Online: https://www.frontiersin.org/article/10.3389/fpos.2021.665004 [Accessed 25 April 2022].

Mazzonna, F. & Peracchi, F. (2017). Unhealthy Retirement?, *Journal of Human Resources*, vol. 52, no. 1, pp.128–151.

Mein, G., Martikainen, P., Hemingway, H., Stansfeld, S. & Marmot, M. (2003). Is Retirement Good or Bad for Mental and Physical Health Functioning? Whitehall II Longitudinal Study of Civil Servants, *Journal of Epidemiology and Community Health*, vol. 57, no. 1, pp.46–49.

Melchior, M., Berkman, L. F., Niedhammer, I., Chea, M. & Goldberg, M. (2003). Social Relations and Self-Reported Health: A Prospective Analysis of the French Gazel Cohort, *Social Science & Medicine (1982)*, vol. 56, no. 8, pp.1817–1830.

Messe, P.-J. & Wolff, F.-C. (2019). The Short-Term Effects of Retirement on Health within Couples: Evidence from France, *Social Science & Medicine*, vol. 221, pp.27–39.

Midanik, L. T., Soghikian, K., Ransom, L. J. & Tekawa, I. S. (1995). The Effect of Retirement on Mental Health and Health Behaviors: The Kaiser Permanente Retirement Study, *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, vol. 50, no. 1, pp.S59–S61.

Miller, R. & Hollist, C. (2007). Attrition Bias, *Encyclopaedia of measurement and statistics*, vol. 1.

Minkler, M. (1981). Research on the Health Effects of Retirement: An Uncertain Legacy, *Journal of Health and Social Behavior*, vol. 22, no. 2, pp.117–130.

Mosca, I. & Barrett, A. (2016). The Impact of Voluntary and Involuntary Retirement on Mental Health: Evidence from Older Irish Adults, *The Journal of Mental Health Policy and Economics*, vol. 19, no. 1, pp.33–44.

Motegi, H., Nishimura, Y. & Oikawa, M. (2020). Retirement and Health Investment Behaviors: An International Comparison, *The Journal of the Economics of Ageing*, [e-journal] vol. 16, p.100267, Available Online: https://www.sciencedirect.com/science/article/pii/S2212828X20300323 [Accessed 17 July 2022].

Motegi, H., Nishimura, Y. & Terada, K. (2016). Does Retirement Change Lifestyle Habits?: Does Retirement Change Lifestyle Habits?, *The Japanese Economic Review*, vol. 67, pp.169–191.

Müller, T. & Shaikh, M. (2018). Your Retirement and My Health Behavior: Evidence on Retirement Externalities from a Fuzzy Regression Discontinuity Design, *Journal of Health Economics*, vol. 57, pp.45–59.

Nielsen, N. F. (2018). Sick of Retirement?, SSRN Scholarly Paper, 3096512, Available Online: https://papers.ssrn.com/abstract=3096512 [Accessed 16 June 2022].

Nieminen, T., Prättälä, R., Martelin, T., Härkänen, T., Hyyppä, M. T., Alanen, E. & Koskinen, S. (2013). Social Capital, Health Behaviours and Health: A Population-Based Associational Study, *BMC Public Health*, vol. 13, no. 1, p.613.

Nishimura, Y., Oikawa, M. & Motegi, H. (2017). WHAT EXPLAINS THE DIFFERENCE IN THE EFFECT OF RETIREMENT ON HEALTH? EVIDENCE FROM GLOBAL AGING DATA: EVIDENCE FROM GLOBAL AGING DATA, *Journal of Economic Surveys*, vol. 32.

Oksanen, T., Vahtera, J., Westerlund, H., Pentti, J., Sjösten, N., Virtanen, M., Kawachi, I. & Kivimäki, M. (2011). Is Retirement Beneficial for Mental Health? Antidepressant Use before and after Retirement, *Epidemiology (Cambridge, Mass.)*, vol. 22, no. 4, pp.553–559.

Ostberg, H. & Samuelsson, S. M. (1994). Occupational Retirement in Women Due to Age. Health Aspects, *Scandinavian Journal of Social Medicine*, vol. 22, no. 2, pp.90–96.

Picchio, M. & Ours, J. C. van. (2020). Mental Health Effects of Retirement, *De Economist*, vol. 168, no. 3, pp.419–452.

Pinquart, M. & Sörensen, S. (2000). Influences of Socioeconomic Status, Social Network, and Competence on Subjective Well-Being in Later Life: A Meta-Analysis, *Psychology and Aging*, vol. 15, no. 2, pp.187–224.

Powdthavee, N. (2008). Putting a Price Tag on Friends, Relatives, and Neighbours: Using Surveys of Life Satisfaction to Value Social Relationships, *The Journal of Socio-Economics*, vol. 37, no. 4, pp.1459–1480.

Prince, M. J., Reischies, F., Beekman, A. T., Fuhrer, R., Jonker, C., Kivela, S. L., Lawlor, B. A., Lobo, A., Magnusson, H., Fichter, M., van Oyen, H., Roelands, M., Skoog, I., Turrina, C. & Copeland, J. R. (1999). Development of the EURO-D Scale--a European, Union Initiative to Compare Symptoms of Depression in 14 European Centres, *The British journal of psychiatry*, [e-journal] vol. 174, pp.330–338, Available Online: https://doi.org/10.1192/bjp.174.4.330 [Accessed 25 April 2022].

Rohwedder, S. & Willis, R. J. (2010). Mental Retirement, *Journal of Economic Perspectives*, vol. 24, no. 1, pp.119–138.

Ronconi, L., Brown, T. T. & Scheffler, R. M. (2012). Social Capital and Self-Rated Health in Argentina, *Health Economics*, vol. 21, no. 2, pp.201–208.

Saffer, H. (2005). The Demand for Social Interaction, SSRN Scholarly Paper, 875732, Available Online: https://papers.ssrn.com/abstract=875732 [Accessed 15 May 2022].

Social Security History. (2022). , Available Online: https://www.ssa.gov/history/age65.html [Accessed 16 June 2022].

Stock, J. & Yogo, M. (2005). Identification and Inference for Econometric Models, New York: Cambridge University Press, pp.80–108.

Szinovacz, M. E. & Davey, A. (2004). Retirement Transitions and Spouse Disability: Effects on Depressive Symptoms, *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, vol. 59, no. 6, pp.S333-342.

Taylor, R. J., Chatters, L. M., Hardison, C. B. & Riley, A. (2001). Informal Social Support Networks and Subjective Well-Being among African Americans, *Journal of Black Psychology*, vol. 27, no. 4, pp.439–463.

The 2021 Ageing Report: Economic and Budgetary Projections for the EU Member States (2019-2070). (2022). , Available Online: https://economy-finance.ec.europa.eu/publications/2021-ageing-report-economic-and-budgetary-projections-eu-member-states-2019-2070_en [Accessed 6 June 2022].

The Survey of Health, Ageing and Retirement in Europe (SHARE): Home. (2022). , Available Online: http://www.share-project.org/home0.html [Accessed 9 July 2022].

Vo, K., Forder, P. M., Tavener, M., Rodgers, B., Banks, E., Bauman, A. & Byles, J. E. (2015). Retirement, Age, Gender and Mental Health: Findings from the 45 and Up Study, *Aging & Mental Health*, vol. 19, no. 7, pp.647–657.

Zang, E. (2020). Spillover Effects of a Husband's Retirement on a Woman's Health: Evidence from Urban China, *Social Science & Medicine*, [e-journal] vol. 245, p.112684, Available Online: https://www.sciencedirect.com/science/article/pii/S0277953619306793 [Accessed 16 April 2022].

A. Tables

A.1 Legal Retirement Ages across Countries and Genders

Table A1: Legal retirement ages across countries and genders

Country	Men	Women
Austria	65	60
Belgium	65	65
Czechia	62 and 6 months	61 and 4 months
Denmark	65	65
France	60	60
Germany	65 and 1 month	65 and 1 month
Italy	66	62
Spain	65	65
Sweden	65	65
Switzerland	65	64

Note: State pension ages are provided by U.S.A.'s Social Security Administration's Social Security Program Survey for year 2012. These pension ages were valid for people retiring between SHARE waves four and five which is the interval of interest in our research. Ages with specific months of retiring (such as state pension ages in Czechia and Germany) were transformed into decimals.

A.2 The Mental Health Effects of Retirement: Bandwidth choice

Table A2: FE-IV models – Mental health effects of retirement (3-year age window with quadratic age trend)

Definition of Retirement	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Outcome Variable	*Euro-D*	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Long-term effects						
Excluding lagged mental health						
$\widehat{R_{it-2}}$ *Second-stage*	1.91*** (0.82)	2.31*** (1.09)	2.11*** (0.96)	0.32*** (0.26)	0.48*** (0.41)	0.39*** (0.33)
$\overline{pa}_{it-2}$ *First-stage*	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)
F-statistic	64.79	32.45	42.30	64.85	29.56	42.41
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Including lagged mental health						
$\widehat{R_{it-2}}$ *Second-stage*	1.55*** (0.66)	1.98*** (0.97)	1.67*** (0.81)	0.21*** (0.18)	0.35*** (0.30)	0.24*** (0.24)
$\overline{pa}_{it-2}$ *First-stage*	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)	0.19*** (0.01)	0.15*** (0.01)	0.16*** (0.01)
F-statistic	64.56	27.89	41.92	64.89	27.95	41.98
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	9 365	9 365	8 282	9 365	9 365	8 282

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 (1) and (2).

A.3 The Mental Health Effects of Retirement: Inverse probability weighting

Table A3: FE-IV models – using inverse probability weighting

Definition of Retirement	**(1)**	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Outcome Variable	*Euro-D*	*Euro-D*	*Euro-D*	*Clinical depression*	*Clinical depression*	*Clinical depression*
Long-term effects						
10-year age window with lagged mental health						
$\widehat{R_{it-2}}$ *Second-stage*	1.25*** (0.47)	1.93*** (0.69)	1.46*** (0.55)	0.19*** (0.13)	0.27*** (0.19)	0.23*** (0.15)
F-statistic	72.27	43.87	52.80	73.54	44.62	52.93
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	13 530	13 530	11 669	13 530	13 530	11 669
3-year age window with lagged mental health						
$\widehat{R_{it-2}}$ *Second-stage*	1.13*** (0.32)	1.59*** (0.39)	1.37*** (0.40)	0.17*** (0.13)	0.23*** (0.11)	0.20*** (0.10)
F-statistic	128.92	87.20	69.73	110.02	85.97	66.33
Hausman test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
n	6 916	6 916	5 919	6 916	6 916	5 919

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 (1) and (2).