Work 'til You Drop: Short- and Longer-Term Health Effects of Retirement in Europe

Gabriel H. Sahlgren

Work 'til You Drop: Short- and Longer-Term Health Effects of Retirement in Europe[*]

Gabriel H. Sahlgren[†]

September 2012

Abstract

Declining fertility rates and increasing life expectancy necessitate a higher labor participation rate among older people in order to sustain pension systems and boost economic growth. At the same time, researchers have only recently begun to pay attention to the health effects of a longer working life, with rather mixed results thus far. Utilizing panel data from eleven European countries, and two distinct identification strategies to deal with endogeneity, we provide new evidence of the health effects of retirement. In contrast to prior research, we analyze both the impact of being retired and the effect of spending longer time in retirement. Using spouses' characteristics as instruments, while taking precautions to ensure validity, we find a robust, negative impact of being retired and spending longer time in retirement on self-assessed, general, mental and physical health. In addition, we show that the impact on self-assessed health remains similar in models using instruments from previous research while also including individual- and time-fixed effects to remove time-invariant unobserved heterogeneity between individuals as well as common health shocks. Overall, the results suggest that this innovation and the fact that we take lagged effects into account explain the differences in comparison to prior multi-country research using these instruments. While the short-term health impact of retirement in Europe remains uncertain, the medium- to long-term effects appear to be negative and economically large.

Keywords: Health, Retirement, SHARE, SHARELIFE

JEL Classifications: I10, J26, J14

1   Introduction

Since World War II, declining fertility rates in combination with increasing life expectancy, and the emergence of retirement as a widespread phenomenon, have lead to a significant reduction of the working population in industrialized countries. To a large extent, this is due to various pathways into early retirement and financial incentives to utilize those pathways, which have been shown to be important for labor market exit (e.g. Gruber and Wise 2004; Asch, Haider, and Zissimopolous 2005; Börsch-Supan, Brugiavini, and Croda 2009; Euwals and Trevisan 2011; Hanel 2011). In addition, the overall generosity of public pensions is also important in explaining differences in the average retirement age across countries (Hurd, Michaud, and Rohwedder 2012). In general, the reduction of the working population has increased pressure on countries' pension systems, but the issue concerns more than solvency of such systems. Countries with a disproportionate number of people outside the labor force are bound to suffer from lower savings rates, which, in turn, stifle investments and economic growth. In other words, maintaining a high labor force participation rate is also important for countries' long-run economic well being. As a response to the challenges, politicians in OECD and European Union countries have begun to reform their countries' pension systems to induce people to work longer, particularly by reducing financial incentives to retire.

While the direct economic implications of high labor force participation are relatively well studied, less is known about the health effects of such participation. Since people's decision to stop working gives rise to important life changes that may alter their health status, it is crucial that policymakers take this into consideration. This is because the population's health status affects health spending, which in most OECD countries to a large extent is publicly funded. If employment in old age worsens health, policymakers might face increasing health spending as a result of their policies. If, on the other hand, employment in old age preserves or improves health, public spending in this area could be reduced. Whether retirement is good or bad for health has thus important implications for the state of countries' public finances.

In this paper, we provide new micro-level evidence regarding the health effects of retirement using data from the Survey of Health, Ageing, and Retirement in Europe (SHARE), with two distinct strategies to deal with endogeneity in individuals' retirement decisions. The first strategy exploits, in a novel way, spouses' characteristics as instruments while taking precautions to ensure validity. In contrast to prior multi-country research, the results suggest rather strong negative impacts on self-assessed, general, physical and mental health.

Conspicuously, in our second strategy, we utilize the special third SHARELIFE wave and previously used instruments constructed from official retirement ages, while controlling for individual- and time-fixed effects as well as taking into account that the short- and long-term impacts differ. We find that the negative effects of retiring and spending more time in retirement remain. While data only permit us to use this strategy when analyzing self-assessed health, the effect size is similar or larger compared to our first strategy in this respect. We also show that including individual-fixed effects and assuming a lagged impact are crucial for these conclusions. Indeed, we find evidence that the short-term effect of retirement may differ significantly. Future country-specific research should be careful in taking both issues into account.

The differences between our findings and previous results from multi-country research are attributed to research design. In contrast to prior studies, we (1) analyze both the impact of labor market status and time spent in retirement; (2) take into account that short-term and longer-term effects of both variables may differ considerably; and (3) use alternative strategies to deal with endogeneity as noted above.

Overall, our findings lend credence to a win-win scenario in which countries' economic growth prospects, public finances and health are better off—at least in a medium- to long-term perspective—when older people engage in employment rather than retiring entirely. Inducing people to work longer is not only a necessary feature of sustainable pension systems and an important ingredient of any growth strategy, but also appears beneficial for individuals' well being. Our strategies can only analyze specific subpopulations, however, and future research should investigate effects among other subgroups. In order to stem the negative longer-term impact of retirement, future research should also investigate more thoroughly the mechanisms through which employment preserves health.

The paper proceeds as follows. Section 2 discusses the ambiguous theoretical impact of retirement on health; Section 3 reviews the literature; Section 4 discusses the data; Section 5 outlines our research strategies; Section 6 presents the results; and Section 7 concludes.

2   The Ambiguous Theoretical Impact of Retirement on Health

The theoretical health impact of retirement is far from straightforward. In Grossman's (2000) human capital model, health is crucial for utility maximization. Health, in this model, is both an investment and consumption good—it impacts utility directly through its effect on life satisfaction and happiness while also reducing work-related illness, which allows people to

raise their total earnings. The former mechanism may lead people to invest more in their health after retirement, whereas the latter mechanism may lead them to invest less. Whether incentives to invest in health increase or decrease after an individual retires thus depends on whether the marginal benefit of health is higher or lower compared to before—and there is no straightforward correct answer (Dave, Rashad, and Spasojevic 2006). Additionally, it is important to note that health investments may also change prior to retirement since individuals engage in retirement planning, which may kick in once they have retired.

It is also noteworthy that other mechanisms are equally ambiguous. The social capital literature, for example, indicates beneficial effects of trust and social interactions on health (e.g. Petrou and Kupek 2008; d'Hombres et al. 2010; Ronconi, Brown, and Scheffler 2012). Retiring can reduce social interactions since one loses former colleagues and work-related contacts in general. However, one also has more leisure with which to establish new social contacts outside work. Additionally, one has more time to devote to voluntary work, which is a base at which new contacts can be established.

Equally, while stress is clearly important for health, the impact of retirement on stress is less clear-cut. Permanent labor force exit is an important life-event that can be very stressful, but it can certainly decrease work-related pressure. The same applies for physical exercise. Some people get most of their exercise from work, whereas retirement may also alter their behavior so that voluntary exercise increases. The impact of retirement on exercise appears to be heterogeneous depending on the type of people and the type of job from which they exit (Chung et al. 2009; Kuvaja-Köllner et al. 2012).

Furthermore, it is important to note that the health impact of retirement is not necessarily linear as a function of time; immediate and short-term effects may differ significantly from medium- to long-term impacts. For example, a spike in leisure time immediately following retirement can have a short-term beneficial impact on health even if the marginal utility of health is lower in retirement. This is because effects of investments in health are not instantaneous. The impacts following lower investments in health—as well as from other channels, such as lost social networks—do not necessarily appear until at some point in the future. Of course, it could also be the other way around if the marginal utility of health is higher in retirement, in which case the positive impact of health investments appear later. Similarly, retirement may create stress that later subsides as the individual adapts to the new situation. Regardless, it is clearly theoretically plausible that longer-term health effects of retirement can differ significantly from the short-term impact.

3 Previous Literature

Having briefly discussed the theoretical ambiguity regarding the impact of working on health in old age, this section reviews the empirical literature. A key problem for any valid estimation strategy is that employment status is not random but rather a function of a wide variety of predictors that are not necessarily observed. For example, retirement is clearly endogenous to health (e.g. Coile 2004; Cai 2010; García-Gómez, Jones, and Rice 2010; Zucchinelli et al. 2010). Consequently, a plethora of studies, which display conflicting results, cannot separate causality from correlation (e.g. Litwin 2007; Alavinia and Burdorf 2008; Brockman, Müller, and Helmert 2009; Jang et al. 2009; Westerlund et al. 2009; Roberts et al. 2011). Similarly, research merely controlling for time-invariant unobserved hetereogeniety, such as Kerkhofs and Lindeboom (1997) and Lindeboom, Portrait, and van den Berg (2002), cannot partial out time-variant sources of endogeneity.

In general, the most rigorous research is mixed, with results depending on country analyzed and methodology utilized. A regular approach in the literature has been to use instrumental-variable (IV) models with eligibility ages to early and regular pension benefits as instruments for retirement. The rationale is that reaching the official retirement age induces significant financial incentives to exit the labor market, enabling researchers to use fuzzy regression-discontinuity designs (RDD) to study the impact of retirement on health. Another approach has been to use changes to the retirement age, and other features of the pension system, as instruments. The multi- and within-country research using these strategies generally finds positive or no health effects of retirement (Charles 2004; Bound and Waidmann 2007; Neuman 2008; Johnston and Lee 2009; Lindeboom and Lindegaard 2010; Bingley and Pedersen 2011; Bonsang and Klein 2011; Coe and Zamarro 2011; Fé and Hollingsworth 2011, 2012; Kantarci and van Soest 2011; Blake and Garrouste 2012; Hernaes et al. 2012; Insler 2012; Latif 2012).[1] The exceptions are Lei, Tan, and Zhao's (2011) study, which displays a large, negative impact of retirement on self-assessed health in China, and Kolodziej's (2011) paper, which displays some negative effects of retirement on mental health across eleven European countries. Some research finds that people's BMI increases as a result of retirement (Kantarci and van Soest 2011). Naturally, whether or not this is good or

---

[1] Kantarci and van Soest (2011) find some negative effects of retiring on mental health, but only when using regular OLS fixed-effects models. Insler (2012) also includes interactions between the eligibility ages and self-reported probabilities of working past these ages as additional instruments, which is likely to compound the retirement planning and selection issues discussed in Section 3.1. Indeed, Behncke (2012) actually controls for subjective expectations regarding retirement date to make sure that these do not bias estimates (see Section 3.2).

bad for health depends on the BMI level of the retiree, but other research indicates that already overweight or obese people gain the most (Chung, Domino, and Stearns 2009).

Researchers have also utilized unexpected retirement offered by U.S. employers to employees as instruments. With this approach, Coe and Lindeboom (2008) find no or positive effects of retirement on health. Calvo, Sarkisian, and Tamborini (2011) use both unexpected retirement windows and eligibility ages as instruments, and find that when retirement occurs on time, meaning at the established U.S. early retirement threshold, it is positive for self-reported health. On the other hand, if it occurs earlier or later, it is negative. This may be due to anticipation effects among individuals produced by the official retirement age; individuals are likely to adapt to later retirement in the long run once they get accustomed to new rules.

3.1 Concerns

The research discussed above is thus mixed, but a majority display either positive or no effects of retirement on health. However, some questions arise regarding the viability of the instruments utilized. Indeed, it is possible that official retirement ages, unexpected retirement windows and changes in pension systems are all endogenous to health. First, if people anticipate that there are strong financial incentives to retire at certain ages they may adjust their behavior well before their labor force exit. While retirement planning naturally applies to most regular pathways to labor market exit in old age, research suggests that the state pension system has traditionally been the key route because of the financial incentives involved (Gruber and Wise 2004), which is likely to make it susceptible to more significant and long-term retirement planning than other routes. Indeed, Calvo, Sarkisian, and Tamborini's (2011) finding that retirement earlier or later than at the early retirement age is negative for health suggests that people generally plan for retirement according to state pension rules. The problem also applies to changes in the state pension system if these are predictable, which is often the case due to public debate. Furthermore, since individuals have time to adapt to the reforms before they retire, unobserved changes in behavior induced by the change could bias estimates. For example, there is some evidence that reforms reducing pension rights lead workers to participate in training courses (Montizaan, Cörvers, and de Griep 2010).[2]

---

[2] We do not discuss studies analyzing the impact of retirement on cognitive ability since our explicit focus is on health and not cognitive achievement. Interestingly, these studies utilize similar strategies—and results vary also in this literature (e.g. Rohwedder and Willis 2010; Bingley and Martinello 2011; Coe and Zamarro 2011; Coe et al. 2012; de Griep et al. 2012; Mazzonna and Peracchi 2012; Perelman, Adam, and Bonsang 2012).

In addition, as Behncke (2012) argues, reaching the official eligibility age might also have an independent impact on health, especially mental, if it is considered a milestone in an individual's life. It might also be the case that individuals with certain health select into jobs with mandatory retirement—or jobs in which early retirement is more likely granted—which compounds the direction of causality further. This is especially true of the early retirement age, and changes to that age, which is because of increased likelihood of self-selection. All individuals do not necessarily even know whether they are eligible for early retirement, and those who do are more likely to have been searching for ways to retire prematurely. Similarly, a concern with the approach utilizing unexpected retirement windows as instrument is that employees of companies giving such offers may differ systematically from other individuals.

3.2   *Different Strategies—Different Findings*

Due to the above issues, it is conspicuous that other estimation strategies display different findings. Analyzing U.S. data, Dave, Rashad, and Spasojevic (2008) use spouses' retirement status as instruments, and find that retirement predicts worse physical and mental health. However, the authors take no steps to control for the potential direct impact of spouses' health, which could bias estimates. Research also finds negative effects of retirement in Austria. Kuhn, Wuellrich, and Zweimüller (2010)—using a region-specific unexpected change in the Austrian old age unemployment insurance system as instrument—find that early retirement among male blue-collar workers increases the probability of dying before age 67. However, there is no effect among women. It is noteworthy that very little planning for the new rules could have been taking place since they were instituted because of an economic slump. Take-up was dependent on region rather than age, which may also explain the different findings compared to the above noted papers. In Japan, Kajitani (2011) uses self-employment and marital statuses as instruments, and finds a positive impact of a more significant workload on self-assessed health among elderly men. The strategy, however, ignores that individuals' with specific health profiles may self-select into self-employment. Marital status, too, can potentially affect health significantly (Koball et al. 2010).

Analyzing English data, Behncke (2012) focuses on individuals who were employed in the first wave of interviews, and analyze whether retirement at the second wave affects their health in the third wave. She finds that retirement produces worse health in models assuming that all variables jointly affecting retirement and health are included, but also in IV models

using the normal retirement age as instrument. Yet, in the IV model, the effect is insignificant for subjective health, chronic conditions and mental health. Conspicuously, the author controls for individuals' subjective expectations regarding retirement date as well as potential to experience health shocks limiting their ability to work. This is supposed to take into account retirement planning for the official retirement age. It is unlikely, however, that controlling for subjective expectations are sufficient because of potential justification bias. In addition, it is implausible that individuals are able to predict their own health shocks.

Overall, therefore, the most methodologically convincing research on the health effects of retirement is rather mixed, which is likely to due to researchers' employing different research strategies and data.[3] It is also conspicuous that few studies, and no multi-country ones, take lagged effects into account properly. This is problematic since the short- and long-term effects of retirement may differ significantly. In addition, prior multi-country research has not controlled for prior health levels. Moreover, it has neither taken into account that age effects on health may differ between countries, nor that instruments based on retirement ages may be endogenous without taking sufficient precautions. The latter apply to many within-country studies too. Furthermore, prior research, multi-country and country-specific, has focused on the impact of being fully or partially retired, but no study evaluates the effect of time spent in retirement. Our paper attempts to remedy these shortcomings.

4   Data

Motivated by the mixed literature, and the above-noted flaws in multi-country research, we investigate the relationship between retirement and health in Europe in a similar multi-country framework to other research on the effect of retirement on health and/or cognitive achievement (e.g. Rohwedder and Willis 2010; Coe and Zamarro 2011; Kolodziej 2011; Mazzonna and Peracchi 2012). More specifically, we utilize the first two waves of SHARE, which allow us to control for previous health levels while at the same time taking into account that the health effects of being retired and spending longer time in retirement are not necessarily immediate. The third wave of SHARE, a special wave called SHARELIFE,

---

[3] De Griep, Lindeboom, and Montizaan (2012) use a regression-discontinuity design and find that a Dutch reform, which increased the retirement age and decreased the treated individuals' pension benefits, produced worse mental health among people who had not yet retired. While this is relevant for how changes to the retirement system affect health prior to retirement, it is less relevant for how individuals' actual retirement impacts their health. Indeed, the effect might simply be due to the fact that people expect their health to decrease if they have to work longer (c.f. Scheubel, Schunk, and Winter 2009).

generally focused on individuals' life histories, which means that respondents were not inquired about their health. The exception is that the survey registered respondents' self-assessed health, which we exploit in our second identification strategy. Similar to previous research using SHARE, we obtain data from representative samples of individuals in eleven European countries: Austria, Belgium, Denmark, France, Germany, Greece, Italy, the Netherlands, Spain, Sweden, and Switzerland. The key difference of our use of the data compared to prior studies is that we make full use of the survey's panel structure.

4.1 Sample Selection

Since our first identification strategy, which is discussed in depth in Section 5.2, utilizes spouses' characteristics as instruments, we observe in the first instance 6,751 individuals who live with a spouse or partner, all of whom were interviewed in the first two waves of SHARE, and who were 50-69 years old at the time of the first interview. We therefore include both women and men in our analysis, while taking into account that health effects of retirement may differ depending on gender in our robustness tests. When using our second strategy, discussed in Section 5.3, we obtain a larger sample of 9,123 individuals, of the same initial ages as in the first strategy, who were interviewed in the first two waves of SHARE as well as in the third special SHARELIFE wave. The total sample size varies slightly depending on which health indicator we analyze. Since our strategies hinge upon observing respondents several times, while also taking lagged effects into account, the panel is always balanced so that each individual is observed twice in the first strategy and thrice in the second strategy.

4.2 Retirement

Our definition of retirement is straightforward and mostly follows Coe and Zamarro (2011). We define workers as those who report to be either employed or self-employed. Consequently, individuals reporting they are retired, homemakers, the permanently sick and disabled as well as those who are separated from the labor force in other ways are defined as retired.[4] Since it is not possible to separate unemployed people who are currently looking for work from those who do not, and since the effect of short-term unemployment on health is

---

[4] The latter includes students, rentiers, people who live off their own property as well as those who do voluntary work.

likely to be different compared to the impact of retirement, we exclude unemployed individuals from the sample.[5] Following the discussion in Section 2, we lag the variable to take into account that short- and longer-term effects may differ. Due to data availability, and in common with much previous research, we cannot analyze differential effects of retirement from different types of jobs. Nevertheless, unveiling an average effect is certainly important for policymakers to consider.

4.3 Time Spent in Retirement

While previous research focuses on retirement, full or partial, no study has previously evaluated the health effects of time spent in retirement. We thus also use a question in the second wave of SHARE, which inquires for how long respondents have been retired. Additionally, for individuals who do not report to be retired but have nevertheless exited the labor force, we count the years and months since they last held a job.[6] Using the date of each interview in the first and second waves, we then calculate this variable at the point of the first interview. This allows us to investigate more thoroughly whether the potential effect of retirement declines or increases with time. It also allows us to take into consideration that the impact operates with a lag, which is important. Similar to the effect of retiring, the impact of spending additional time in retirement may very well take some time to manifest itself. Since the marginal effect may be expected to decrease with time, we log this variable.

4.4 Health

Our main dependent variables are commonly analyzed health indicators. Previous research consistently utilizes self-assessed health, with respondents rating their own health on a scale between one and five: "poor," "fair," "good," "very good," and "excellent."[7] We use this

---

[5] However, results are very similar when also including the unemployed among the people who have retired. The only difference is that the negative impact of being retired turns marginally insignificant in Model 1 in Table 5 (p=0.10) and Model 1 in Table 8 (p=0.12), while the non-lagged positive impact in Model 2 in Table 15 also turned marginally insignificant (p=0.14). However, the instruments become weaker when using our first strategy, which is expected since the joint retirement hypothesis applies to retirement, not unemployment.

[6] However, the overall results are qualitatively similar when using the date since the respondent last held a job as the date of retirement also for individuals who do report the date of their retirement. The only differences are that the negative impact of time spent in retirement did turn insignificant in Model 5 in Table 8 and marginally insignificant (p=0.10) in Model 17 in Table 13.

[7] In the first wave, a modified European version of this scale was also used. Yet, that version was abandoned in the second wave and since we exploit the panel dimension we do not analyze this measure.

index, but transform it so that a higher score means better health. Self-reported health is consistently found to be a good predictor of mortality (e.g. Ambrasat, Schupp, and Wagner 2011), but there are also problems with the index. First, it could suffer from justification bias if respondents report worse health to justify their labor market status. Additionally, different individuals may view the categories differently, which could produce significant measurement error.

We thus also analyze more objective health indicators. As a measure for mental health, we use the Euro-D depression scale. Again, we rescale this index so that higher values equal better health. Following other research, such as Bound et al. (1999) as well as Coe and Zamarro (2011), we also create a health index by regressing more objective health indicators on the self-assessed health scale:

$$sh_{it} = \alpha + \beta_1 ohi_{it} + \delta_i + \varepsilon_{it} \tag{1}$$

where sh_{it} is the self-assessed health scale, ohi_{it} includes objective health indicators, and δ_i represents individual-level fixed effects.[8] More precisely, we include an indicator for how limited individuals' are due to health reasons, the number of limitations in activities in daily living (ADLs), the number of instrumental activities in daily living (IADLs), whether or not the respondent suffers from any long-term illnesses, the number of diagnosed conditions, the number of symptoms, the number of drugs the respondent is taking, an index of mobility, the Euro-D depression scale, grip strength, indicators for obesity and overweight as measured by conventional thresholds on the BMI, as well as scores on numeracy, memory and verbal tests.[9] Note that our strategy differs compared to the only previous multi-country study on both physical and mental health. Coe and Zamarro (2011) estimate Equation (1) separately in each country to take into account that respondents in different countries may hold different norms when reporting self-assessed health. In contrast, we take into account that individuals may vary in this respect also within countries.[10]

---

[8] Including time-fixed effects too makes no difference for the results, but these are correlated with changes in age. We also control for age and time-fixed effects in our health equations, and, for these reasons, we exclude the latter from the construction of the index.

[9] While our focus is not cognitive achievement, we choose to include it since it is a variable that can be affected by retirement while also being a predictor of self-reported health.

[10] We note that estimating the same equation separately for each country without individual-fixed effects, using Ordered-Probit models, rendered very similar results regarding the impact of retirement on health. The only difference is that the main impact in Model 1 in Table 17 turned marginally insignificant (p=0.12). This is of little importance since the gender-retirement interaction is insignificant in the same model.

In robustness tests we also analyze the impact on physical health indicators only. As measures of physical health, we use the number of diagnosed physical conditions and the number of drugs taken for physical conditions. However, since grip strength is a good predictor of mortality (e.g. Sasaki et al. 2007), we also analyze this as a proxy for physical illness. The proxy may also be affected by mental health, and we thus control directly for it in the equation explaining grip strength. The other two measures may of course also be directly caused by mental health. Nevertheless, we are interested in the total impact of retirement on physical health, including the indirect effect operating through mental health, and since the variables are direct measures of physical health we do not control for the Euro-D index in these equations.[11]

Finally, we analyze cut-off points. Since we are using the U.S. version of the self-assessed index, we use "very good" and "excellent" as indicators of "good health." Second, we analyze the threshold for clinical depression, which is defined as a score of 4 or above on the Euro-D scale according to convention. Third, we estimate the impact on having been diagnosed with at least one chronic physical disease as well as on taking at least one drug for physical illness. Again, all variables are scaled so that a higher value means better health. Table 1 displays the descriptive statistics.

[INSERT TABLE 1]

5   Research Design

5.1   *The problem of identification*

The key issue involved in analyzing the impact of retirement on health is the estimation approach. As noted by prior research, the simplest strategy is to analyze the relationship using:

$$h_{it} = \alpha + \beta_1 e_{it-1} + \beta_2 a_{it-1} + \beta_3 x_{it} + \varepsilon_{it} \tag{2}$$

where h_{it} denotes one of the health measures described above; e_{it-1} represents the person's employment status—which can either take the value 1 (retired) or 0 (working for pay) —*or* the logged time the respondent has spent in retirement; a_{it-1} is a vector of control variables

---

[11] In unreported regressions, we did include the Euro-D index and results were almost identical.

measured at the same time as e_{it-1} and x_{it} is a vector of control variables measured at the same time as h_{it}.

The key assumption in the above model is that $Cov(e_{it-1}, \varepsilon_{it}|a_{it-1}, x_{it}) = 0$. But if x_{it} and a_{it-1} do not include all relevant variables jointly affecting e_{it-1} and h_{it}, or if h_{it} impacts e_{it-1} directly, the latter is endogenous to the former. Furthermore, measurement error in the variables can also produce bias in the estimates. Any of these problems, in turn, would mean that $Cov(e_{it-1}, \varepsilon_{it}|a_{it-1}, x_{it}) \neq 0$. In order to estimate a causal impact of retirement on health, therefore, it is crucial to find a variable that isolates the variation in e_{it-1} that is exogenous to h_{it}.

What variable should this be? The only two studies that analyze health using multi-country data, as well as many studies on the relationship between retirement and health in general, utilize dummy variables indicating whether or not individuals' are older than the official early and normal retirement ages, which we denote $\bar{r}$, as instruments for retirement (Coe and Zamarro 2011; Kolodziej 2011). The rationale behind this approach is that reaching the official retirement ages radically increases the likelihood of an individual retiring because of the strong financial incentives to do so. Yet, for the reasons discussed in Section 3.1, it is possible that $Cov(\bar{r}_{t-1}, \varepsilon_{it}|e_{t-1}, a_{it-1}, x_{it}) \neq 0$, which could bias the impact of retirement on health upwards.

5.2   Strategy 1: Spouses' Characteristics

In our first strategy, therefore, we use two entirely different instruments. First, we exploit the fact that spouses' labor market status is related to individuals' labor market status. If the spouse is working, it is also likely that the respondent will do so. Similarly, spouses' labor market status is likely to be significantly related to the time an individual has spent in retirement. With a working spouse, individuals are likely to have spent less time in retirement. These relationships are plausible due to various incentives for joint retirement, a hypothesis supported by recent research (e.g. Banks, Blundell, and Casanova 2010; Casanova 2010; Stancanelli 2012). Spouses' retirement status is thus likely to be a significant predictor of individuals' retirement status and the amount of time spent in retirement.[12] Our first

---

[12] If spouses have insufficient savings to counter a drop in income when one of them retires, there might be incentives to continue working. However, leisure complementarities provide stronger incentives to retire even if this is the case (Banks, Blundell, and Casanova 2010). Regardless, we control for both respondents' and spouses' education levels, which partials out socioeconomic differences in the first stage equation.

instrument, spe_{it-1}, then, is a dummy indicating whether or not the spouse is retired at the time of the first wave interview.

In addition, we complement this with the logarithm of spouses' ages, denoted spa_{it-1}. This is, to our knowledge, the first time this variable, and therefore the combination with the first variable, is used as an instrument for employment status and time spent in retirement. The instrument is also used because of the joint retirement hypothesis. The likelihood of individuals' retirement increases as a function of ageing, which could induce their partners to retire while expecting to be followed within a certain time frame.

A problem occurs since this variable can to a certain extent also be expected to pick up the effect of retirement at the official retirement ages, which could be problematic given the potential for self-selection into early retirement as described above. The closer to the official retirement ages, furthermore, the more likely people are beginning to adjust their behavior in anticipation of being eligible for retirement benefits. As explained below, however, we take further precautions to make sure that this does not bias our findings. Since our instruments do not exploit financial incentives in the state pension system, which has traditionally been the key route to retirement in most countries, and because we explicitly control for such incentives as explained below, it is less likely that estimates will be biased because of significant retirement planning prior to labor force exit. Furthermore, the other selection problems discussed in Section 3.1 do not apply to our first strategy.

5.2.1   Ensuring Instrument Validity

Nevertheless, our instruments are only valid if spouses' employment status and age are exogenous to individuals' health when conditioning on the other relevant variables. If, however, spouses' health has an independent causal impact on h_{it}, the exclusion restriction is violated since our instruments and spouses' health then both correlate with our outcome variable h_{it}. This is unlikely to be the case with physical health but individuals' mental health might be directly related to their spouses' health. The only available research on the topic that deals with endogeneity indicates that there is in fact no causal impact of spouses' health in general, with the exception that there is some evidence of causality running from husbands' general health to wives' mental health (Michaud and van Soest 2008). If so, $Cov(spe_{it-1}, spa_{it-1}, \varepsilon_{it} | e_{t-1}, a_{t-1}, x_{it}) > 0$ and this makes it possible that we find a false effect of retirement on mental health. However, it also indicates that our estimates that deal

with physical and self-assessed health are most likely not biased even if we do not control for an independent impact of spouses' health.

Nevertheless, it is clearly important to take this potential source of bias seriously. But how can we control sufficiently for the impact of spouses' health? It is clearly problematic to control directly for spouses' health, lagged or non-lagged, since if we acknowledge that spouses' health may impact individuals' health directly we also acknowledge that individuals' health may impact spouses' health directly.[13] By controlling for spouses' health, therefore, we are likely to condition for a portion of the individuals' health too. Furthermore, and more important, spouses' health is endogenous to our instruments spe_{it-1} and spa_{it-1}——since spouses' health is plausibly affected by spouses' labor market status and age——which means that if we condition for sh_{it-1} we simply control away for a significant portion of the impact of our instruments on h_{it} through e_{it-1}. Thus, conditioning on spouses' health constitutes a "bad control" problem (see Angrist and Pischke 2009).

Nevertheless, to display the robustness of our findings, we do present results of regressions were we include a new health index, $sh1_{it}$, which is created by regressing spouses' objective health indicators, $sohi_{it}$ as well as their subjective health, ssh_{it}, on respondents' self-assessed health while also controlling for individuals' objective health indicators, ohi_{it}:

$$sh1_{it} = \alpha + \beta_1 sohi_{it} + \beta_2 ohi_{it} + \beta_3 ssh_{it} + \delta_i + \varepsilon_{it} \qquad (3)$$

We lag this variable so it is measured at the same time as our main predictors. This means that we control for individuals' health as far as their spouses' objective and subjective health indicators can explain it, over and above respondents' own objective health conditions. In a final robustness test, we also simply include spouses' self-assessed health directly. It should be noted that spouses' self-assessed health is likely to be more endogenous to individuals' self-assessed health, since it is not clear that respondents separate their own overall health from spouses' entirely.

Yet, we can also simply control for respondents' prior health conditions, h_{it-1}. Doing so solves the potential problem regarding the independent impact of spouses' health since we then control for it indirectly with the same lag as we measure respondents' employment status. We thus argue that the models explaining changes in health, which include h_{it-1} as discussed below, sufficiently control for the separate impact of spouses' health. Nevertheless,

---

[13] Since health is a persistent and cumulative variable, prior realizations of spouses' health are not exogenous to future realizations of individuals' health.

to display robustness, we also report findings from models that include spouses' health, as discussed above, as well as h_{it-1}.

Finally, since individuals' and spouses' employment statuses, as well as spouses' ages, are correlated with the official retirement ages, we also take a further precaution against bias. We control directly for dummy variables indicating whether or not individuals and their spouses have reached the early retirement age when respondents' retirement status is measured. This removes the discontinuous impact of reaching the early retirement age in the household and thus the effect of retirement planning for early retirement in the state pension system, which is important given the discussion in Section 3. We are especially concerned that failing to control for this impact might bias the impact of retirement on mental health (see Behnke 2012). As explained in Section 5.3, in contrast to prior multi-country research, we use the retirement ages that were in place at the time when individuals faced their retirement decisions rather than when they were interviewed. This takes into account pension reforms in European countries, which often changed the retirement age significantly for different parts of the populations. The retirement ages are obtained from Mazzonna and Peracchi (2012).[14]

The two-wave panel allows us to (1) investigate how retirement status and time spent in retirement affect subsequent changes in health, (2) hold initial health levels constant, and (3) take into account that shorter- and longer-term health effects of retirement may differ. These innovations make our strategy significantly different compared to prior multi-country research.

Thus, we estimate two-stage least squares (2SLS) models in which the first and second stages read respectively:

$$e_{it-1} = \alpha + \beta_1 h_{it-1} + \beta_2 spe_{it-1} + \beta_3 spa_{it-1} + \beta_4 a_{it-1} + \beta_5 x_{it} + \varepsilon_{it} \tag{4}$$

$$h_{it} = \alpha + \beta_1 h_{it-1} + \beta_2 \hat{e}_{it-1} + \beta_3 a_{it-1} + \beta_4 x_{it} + \varepsilon_{it} \tag{5}$$

where $\hat{e}_{it-1}$ indicates the predicted values from the first stage. Our control variables in vector x_{it} include age and its square, gender, and educational level.[15] In addition, we control for spouses' educational level. By doing so, we are essentially controlling for household income,

---

[14] However, we make one small change by taking into account that the official retirement age in Denmark was 67 for people who turned 60 before July 1, 1999 (SSA and ISSA 2010).

[15] We also investigated whether including age cube affected our results. The overall findings are very similar, but it often produced a highly singular variance matrix in the first stage. Due to this, we leave out age cube in our first strategy. Results are also very similar when excluding age square.

since this is strongly correlated with education levels. Finally, the vector also includes country-fixed effects to partial out cultural differences that may be important for health. Since there might be separate age effects on health in different countries, which could bias our findings, we also estimate models including interactions between the age terms and country dummies in robustness tests. We note that this contrasts to previous multi-country research. As mentioned above, a_{it-1} includes a variable indicating whether or not individuals and their spouses' have reached the early retirement age as a control for retirement planning. Depending on model, x_{it} or a_{it-1} also includes spouses' health to ensure instrument validity.[16]

However, by including h_{it-1} as a predictor, the above strategy analyzes the causal effect of being retired and time spent in retirement on changes in health. But since we are also interested in the cumulative impact, we also estimate:

$$e_{it-1} = \alpha + \beta_1 spe_{it-1} + \beta_2 spa_{it-1} + \beta_3 a_{it-1} + \beta_4 x_{it} + \varepsilon_{it} \tag{6}$$

$$h_{it} = \alpha + \beta_1 \hat{e}_{it-1} + \beta_2 a_{it-1} + \beta_3 x_{it} + \varepsilon_{it} \tag{7}$$

where we again allow for a lag in the effect, but exclude h_{it-1}. Like Coe and Zamarro (2011), we utilize standard errors clustered at the individual level, but results are similar when clustering at the household level.

It is important to note that our strategy cannot analyze the average treatment effect (ATE) of being retired and time spent in retirement. Rather, we estimate a local average treatment effect (LATE) on individuals who respond to the instruments. In the case of retirement, this means people who retire because their spouses do so and because of their spouses' age. In the case of time spent in retirement up until the interview, it means people whose time in retirement increases as a result of having a spouse in retirement as well as because of their spouses' ages. Our first strategy, therefore, can only estimate the health impact of being retired and spending longer time in retirement among couples.[17]

---

[16] Since the self-assessed health index is ordinal, an Ordered Probit model could be a viable alternative. However, 2SLS is generally preferred for IV estimations even if the dependent variable is not continuous (Angrist and Krueger 2001; Angrist and Pischke 2009).

[17] This means people who report living with a spouse or partner. Since there is a small discrepancy between this and the indicator for marriage/registered partnership, we tested to control for the latter also when using our first strategy. Unsurprisingly, given the small variation in this sample, the variable was almost never significant, and it never affected the retirement coefficients more than marginally.

5.3 Strategy 2: IV Individual- and Time-Fixed Effects

In our second strategy, we switch our focus to instruments that have been utilized in previous multi-country studies, as well as a significant share of the country-specific research, while employing a different methodology. A special third wave of SHARE, called SHARELIFE, focused on respondents' life histories, and it did not inquire about their health indicators in general. It did, however, still register their self-assessed health. This means that we can observe individuals over three periods, allowing us to include individual-fixed effects—thereby removing all time-invariant unobserved heterogeneity between individuals— at the same time as taking into account that the impact of being retired and spending longer time in retirement could operate with a lag. As instrument, we first use a dummy indicating whether the respondent is above the normal retirement age that applied when they faced their retirement decisions.[18] We then also control for the impact of reaching the early retirement age, which, as argued above, is more likely to be endogenous to health compared to the normal retirement age. Nevertheless, we also include this variable as an additional instrument in other specifications. The retirement ages for the first wave are again obtained from Mazzonna and Peracchi (2012). These take into account large pension reforms up until that wave, which changed eligibility ages for specific groups. For the second wave, however, we must adjust these ages to take into account the changes that occurred between the first and second waves for people who were not already retired by the first wave. To do so, we obtain the relevant retirement ages from the U.S. Social Security Administration's survey of pension systems around the world (SSA and ISSA 2010).[19] Thus, the estimation reads:

$$e_{it-1} = \alpha + \beta_1 \bar{r}_{it-1} + \beta_2 h_{it-1} + \beta_3 a_{it-1} + \beta_4 x_{it} + \delta_i + \mu_t + \varepsilon_{it} \qquad (8)$$

$$h_{it} = \alpha + \beta_1 \hat{e}_{it-1} + \beta_2 h_{it-1} + \beta_3 a_{it-1} + \beta_4 x_{it} + \delta_i + \mu_t + \varepsilon_{it} \qquad (9)$$

where δ_i and μ_t represent individual- and time-fixed effects respectively. We note that no previous study, neither country-specific nor multi-country, has previously controlled for individual- and time-fixed effects simultaneously. However, results are very similar when

---

[18] The instruments constructed from spouses are not suitable for this exercise since they rely strongly upon the cross-unit variation. Indeed, when including individual-level fixed effects, these instruments turned out to be too weak to be of value.

[19] The overall results are similar when we, like all previous multi-country research, use instruments based on retirement ages that were in place at the time of each interview. But this strategy ignores changes to the pension system that affected people differently depending on date of birth.

excluding time-fixed effects. $\bar{r}_{it-1}$ is a vector including the dummies indicating whether or not individuals have reached the regular or early retirement age.[20] As discussed above, depending on estimation, we include both or only the one denoting whether or not individuals have reached the regular retirement age as instrument. Including δ_i makes h_{it-1} mechanically correlated with ε_{it} because of "Nickell bias" (Nickell 1981), and we therefore report findings when including and excluding h_{it-1}. In addition, we also estimate the models using the cut-off point as dependent variable. We cluster the standard errors at the individual level to allow for serial correlation, but results are very similar when clustering at the household level. Again, we also include age-country interactions in robustness tests to take into account that age effects may differ between countries.

It is important to note that including δ_i makes it more likely that the instruments, especially the one constructed from the regular retirement age, are valid. This is because a large part of the selection issues arise because of differences between individuals that do not necessarily differ over time or at the very least not over the time periods we analyze here. By investigating the impact of changes in retirement status on subsequent changes in health, therefore, we are more likely to provide valid estimates when using previously utilized instruments. Finally, by estimating models with Δh_{it} as dependent variable without including h_{it-1}, we also analyze the impact of changes in retirement status and time spent in retirement on *changes-in-changes* in health. This could further minimize the above-discussed biases in both instruments simply because we then focus on the changes in health trajectory rather than health *per se*—changes-in-changes in health are less likely to be dependent on longer-term health levels that could be affected by the instruments.

Again, we can only estimate LATEs, which in this case mean the effects among those who would retire if they reach the official retirement ages and continue to work otherwise. Thus, as in previous multi-country research, we are essentially employing a fuzzy RDD, but with the important differences that we both (1) take into account unobserved heterogeneity among individuals and common health shocks that apply equally to all individuals, as well as (2) focus on the lagged health impact of being retired or spending longer time in retirement.

---

[20] In these regressions, we do include three terms of age since the first stage results did not indicate singularity problems, and all three terms were highly significant in the second stage. Including only one or two terms of age made the age terms insignificant or only marginally significant. In the majority of presented models, results were robust to including only one or two terms of age, but the impact did turn insignificant in a couple of them. The Hansen-J test then often rejected the instruments as exogenous. Thus, three terms of age appear necessary when using IV fixed-effects models in order to pick up the non-linear impact of age on health, which is correlated with retirement.

6   Results

6.1   Strategy 1: Baseline Estimates

Table 2 displays our baseline findings. Model 1 indicates that being retired decreases an individual's self-assessed health by 1.00 standard deviation (SD) between the first interview and second interview, conditional on the baseline level of health. The impact, which is significant at the 5% level, is thus economically large. Turning to mental health in Model 2, we also find a strong effect. The coefficient implies that being retired generates a reduction in mental health by 1.23 SD, an effect that is significant at the 1% level. Finally, we find an even larger effect of on the health stock variable constructed from objective health indicators. The coefficient from Model 3 implies that being retired decreases a person's health stock by 1.65 SD, an effect that is significant at the 1% level.

Models 4-6 display the same regressions when using time spent in retirement as the main predictor. In Model 4, we find that a 100% increase in the time spent in retirement at the time of the first interview decreases a respondent' self-assessed health by 0.34 SD. Model 5 displays that the impact on mental health is 0.46 SD. Finally, Model 6 shows that the impact on the health stock variable amounts to 0.57 SD. The coefficient in Model 4 is significant at the 5% level, while it is significant at the 1% level in Models 5-6.

[INSERT TABLE 2]

Table 3 displays exactly the same regressions, but without the LDV. The estimates thus display the cumulative effects on health levels rather than the impact on changes in health between the first and second interview. Consequently, the coefficient is significantly larger. In Model 1, we observe that being retired at the time of the first interview decreases self-assessed health at the time of the second interview by 1.78 SD. Model 2 displays that the effect on mental health is 1.62 SD, and Model 3 displays a strong impact on the overall health stock of 2.18 SDs. Meanwhile, Models 4-6 display that a 100% increase in time spent in retirement generates a 0.58 SD lower self-assessed health; a 0.55 SD lower mental health; and a 0.76 SD lower health stock. All effects are significant at the 1% level.

[INSERT TABLE 3]

The first stage results, displayed under the results from each second-stage equation, show that both instruments are significant predictors of retirement status and time spent in retirement in the expected direction. The F-statistics display that they are also sufficiently strong, passing the conventional threshold, while the Hansen-J tests indicate that the instruments are valid with p-values between 0.17 and 0.95. Additionally, in all models, the Hausman test strongly indicates that IV estimates differ significantly from OLS estimates, indicating that we are correct at treating our main predictors as endogenous. The first estimates, therefore, indicate that retirement have negative effects on health.

Interestingly, the effect size when dealing with endogeneity is much larger than in unreported OLS estimates. However, it is often the case that IV estimates are larger. First, omitted variables may bias OLS estimates downwards. Second, measurement error in the main predictors can bias estimates downward due to attenuation bias more than reverse causality would bias estimates upwards. In our case, spouses' ages may decrease measurement error since it is reasonable to assume that retirement status and date are less often accurately reported and registered than respondents' date of birth, especially considering that the SHARE sample is explicitly selected based on age. Furthermore, IV models analyze a specific LATE rather than the ATE, meaning that different subpopulations may be affected differently. In the health literature, it is common that IV estimates are larger than OLS estimates, such as Behncke's (2012) findings regarding the impact of retirement on health in the UK, Bonsang, Adam, and Perelman's (2012) results regarding the impact of retirement on cognitive ability in the U.S., and Fonseca and Zheng's (2011) multi-country study on the effects of education on health.

We also note that being over, or having a spouse who is over, the early retirement age often has an independent positive effect on health. This may very well be due to the theoretical mechanisms discussed in Section 3.1, making it important to take precautions when using instruments constructed from the early retirement age. Respondents' and spouses' education levels are also often positively related to health as expected. Interestingly, however, there is no general negative health effect of being female, which prior studies have found. Only in the model explaining long-term differences do women have worse mental health than men. There is some evidence, in fact, that women have higher self-assessed health and better health stock.

6.2 Robustness Analysis

6.2.1 Controlling for Spouses' Health

Since our discussion in Section 5.2.1 indicates that spouses' health levels may make our instruments endogenous, Table 4 displays what happens when we include the index created in Equation (3). We observe that including the index together with the lagged dependent variable barely affects the coefficients and significance levels at all.[21] In fact, in the models explaining the health stock we observe that the spousal health index is significantly negative, which is solely due to the fact that we control for the lagged health stock, indicating that spouses' health does not have positive effects over and above respondents' objective health indicators.

[INSERT TABLE 4]

We also report findings when controlling for spouses' self-assessed health, which, as argued above, introduces further endogeneity and justification problems that might bias our findings. Table 5 displays the results. The impact of our main predictors remains significant across all specifications, although the effect size is reduced somewhat in the case of self-assessed health and the coefficient is only significant at the 10% level when we analyze the impact of being retired. The effect on mental health declines somewhat as well, but remains significant at the 5% level. The impact on the health stock variable, however, is barely affected and is significant at the 1% level. The estimates we obtain when using time spent in retirement as main independent variable are all very similar to previous results while being significant at the 5% level in the case of self-assessed health and 1% level in the case of mental health and health stock. Again, we find that the specification tests satisfy the conventional criteria.[22]

[INSERT TABLE 5]

---

[21] The unreported coefficients of the control variables were similar to those in Tables 2-3.

[22] We include spouses' self-assessed health measured at the same time as the dependent variable. This is because we in previous models control for spouses' lagged self-assessed health together with objective indicators. In unreported regressions, however, we instead included spouses' self-assessed health lagged and all effects were extremely similar. The same applies to the results in Table 6, the impact on physical health and cut-off points in Tables 7 and 8, and the impact when including age-country interactions in Table 9.

Table 6 displays the long-term equations, in which we merely include spouses' self-assessed health and exclude respondents' lagged indicators. The impact remains significant at least on the 5% level, but often stronger. The specification tests again indicate that our instruments are valid and that IV models should be used to deal with endogeneity. We therefore conclude that our estimates are very robust to including both our created index that include spouses' health indicators as well as controlling directly for spouses' self-assessed health. Thus far, therefore, we have no reason to alter the conclusions from the baseline estimates. All results indicate that being retired and spending longer time in retirement induce health deterioration.

[INSERT TABLE 6]

6.2.2   Physical Health Only

Although the health stock variable includes physical health, it might be the case that the results are solely driven by the impact on mental health. To investigate this, we analyze measures of physical health indicators discussed in Section 4. Table 7 displays the results. We find that all these measures are affected by retirement. Including the LDV, and controlling for respondents' lagged levels of subjective health as well as spouses' non-lagged subjective health, we find that being retired increases the number of diagnosed conditions by 1.21 SD (roughly one and a half condition) and the number of drugs taken by 0.97 SD (roughly one drug), while improving grip strength by 1.28 SD. Meanwhile, spending a 100% more time in retirement generates an increase in the number of diagnosed physical conditions and drugs for such conditions by 0.38 SD and 0.30 SD respectively, while decreasing grip strength by 0.38 SD. [23] We thus find statistically significant and economically large negative physical health effects of being retired and spending longer time in retirement.

[INSERT TABLE 7]

---

[23] Interestingly, the effect of lagged subjective health is negative, suggesting that better health is associated with weaker grip strength. Yet, this is entirely due to the inclusion of lagged grip strength, and the impact disappears entirely when excluding this variable.

6.2.3 Cut-Off Points

Having displayed the impact of retirement in old age on various health indices, we analyze relevant cut-off points indicating good health as discussed in Section 4. Since we analyze binary dependent variables, the IV-Probit model could be an alternative. However, there are several problems that arise by using this estimator. First, in regards to retirement status, we are dealing with an endogenous binary predictor for which regular IV-Probit models are not suitable. Furthermore, the regular 2SLS is generally preferable even if the dependent variable is dichotomous and the predictor is continuous (Angrist and Krueger 2001). We therefore opt for regular 2SLS—which in this case becomes a linear probability model—also for these estimations. It should be noted, however, that all effects remain when using IV-Probit models.

[INSERT TABLE 8]

Table 8 displays our findings. Being retired has a consistently negative and statistically significant impact on all indicators. It decreases the probability of being in good subjective health by 0.39 points. At the same time, it increases the probability of (1) suffering from clinical depression by 0.41 points, (2) having at least one diagnosed physical condition by 0.63 points; and (3) taking at least one drug against physical illnesses by 0.60 points. Meanwhile, spending 100% more time in retirement decreases the probability of (1) being in good health by 0.11 points; (2) not suffering from clinical depression by 0.17 points; (3) not having at least one diagnosed physical illness by 0.22 points; and (4) not taking at least one drug against physical illness by 0.19 points. All tests satisfy the conventional criteria apart from the Hansen-J test in Model 5, which indicates that the impact of the instruments differ significantly from each other. Nevertheless, overall, the results are robust to analyzing specific cut-off points despite the fact that much information in the variables disappears.

6.2.4 Including Age-Country Interactions

As a final robustness test in our first strategy, we also include interactions between the age terms and country dummies. This is supposed to take into account that there might be different effects of aging in different countries. Table 9 displays the results while controlling

for the LDV, lagged self-assessed health as well as spouses' self-assessed health.[24] The interactions are often jointly insignificant, and the results are very similar. Thus, there is no evidence that age-country interactions are biasing the above findings.[25]

[INSERT TABLE 9]

6.3   Strategy 2: IV Fixed-Effects with Instruments Constructed from Retirement Ages

Turning to our second strategy, we estimate Equations (8) and (9) in which we control for individual- and time-fixed effects and use the state pension ages as instruments. This is similar to previous multi-country research with two important exceptions: (1) we take into account potential lagged effects of our main predictor, and (2) take into account unobserved heterogeneity between individuals as well as common health shocks that apply equally to all individuals. As explained earlier, we can only analyze the impact on the self-assessed health index due to data availability.

Table 10 displays the results. Model 1 includes a dummy indicating whether or not the respondent is older than the regular retirement age as instrument, while controlling for the discontinuous impact of reaching the early retirement age, the LDV as well as marital status. We find that the impact of being retired is negative with a strong effect size of 0.95 SD, which is similar to our previous results. The specification tests indicate that the instrument is sufficiently strong, and that endogeneity is a problem. Similar to our first strategy, in unreported OLS estimates we found the impact to be insignificant, which again suggests that omitted variables and/or measurement error bias the coefficient towards zero.[26]

Model 2 displays that the positive impact remains when we also include the instrument constructed from the early retirement age, although the effect size drops somewhat and the Hansen-J test is only marginally insignificant. As argued in Section 3, it is more likely that

---

[24] Results are robust to including the index constructed from spouses' self-assessed health as well as their and individuals' objective health indicators.

[25] When also including age cube together with additional age cube-country interactions, the first stage always indicated a highly singular variance matrix. Furthermore, all age-variables, country dummies and age-country interactions turned insignificant with extreme p-values of 1.0 in the second stage. A couple of our results regarding the impact of being retired and time spent of retirement on self-assessed health did turn insignificant in this extreme case. The former, but not the latter, sometimes applied to mental health. This lack of precision is most likely caused by the fact that we add thirty highly correlated variables. But the impact on the health stock and all physical health measures always remained significant even then. The same problems and results apply to the *positive* impact of retirement on self-assessed health when using the second strategy, including age cube and age-country interactions, as well as excluding individual-level fixed effects (see Section 6.3 and footnote 28).

[26] Our instruments are based on thresholds set by governments and should thus reduce measurement error.

people retiring at the early retirement age differ systematically from other individuals compared to people retiring at the regular retirement age. We conclude, however, that our results are robust to using both the early and regular retirement ages as instruments.

[INSERT TABLE 10]

Models 1-2 suffer from Nickell bias due to the inclusion of the LDV. In Model 3, therefore, we exclude it. The effect is now 1.14 SD, which is again very similar to the first estimation strategy. In Model 4, we find that the impact is 1.04 SD when also including the instrument constructed from the early retirement age. Both effects are significant at the 1% level. The Hansen-J test now fails to reject the instruments as exogenous with a comfortable margin. Thus, excluding the LDV in individual-fixed effects regressions seems to make the early and regular retirement ages more similar as instruments.

What about the impact of time spent in retirement? Models 5-7 indicate that a 100% increase in time spent in retirement worsens health by 0.62-0.69 SD, depending on model specification. This impact is always significant at the 1% level. These results are stronger compared to the results using our first strategy in Tables 2-6. The impact withstands using both the early and regular retirement ages as instruments as well as including lagged levels of health, although the instrument constructed from the early retirement age is not a significant predictor of time spent in retirement. It is noticeable that the F-statistic displays that being over the regular retirement age is a strong instrument for time spent in retirement, and that it decreases significantly when also including the instrument based on the early retirement age. Again, therefore, we find no reason to alter our conclusions: the lagged impact of retirement on self-assessed health is negative and quite strong to boot.

In Table 11, we instead analyze the cut-off point for good health. We continue to find negative effects of being retired regardless of instruments used.[27] At the same time, the Hansen-J test is significant when including the early retirement age instrument, which again indicates the differences between the instruments as discussed. In the case of time spent in retirement, we also continue to observe a negative impact irrespective of model specification, while the Hansen-J test comfortably fails to reject the validity of the instruments. The impact is very similar to the results when using our first estimation strategy in the case of the impact of retiring, while it is stronger in the case of time spent in retirement.

---

[27] For these estimations, the IV-Probit model is not a viable alternative since it does not allow for the inclusion of individual-fixed effects.

[INSERT TABLE 11]

In addition, we also analyze the effect of prior changes in retirement status on *changes-in-changes* in subjective health. This means that we use the change in health as dependent variable in the IV fixed-effects model without including the LDV among the predictors. As argued in Section 5.3, this could minimize the potential bias of the instruments since we are focusing on changes in health trajectory rather than changes in health *per se*. Table 12 displays the results. We find a strong, negative effect of being retired and spending more time in retirement regardless of model specification, with the Hansen-J test failing to reject the null hypothesis. Retiring produces up to 1.84 SD worse health trajectory, while spending 100% longer time in retirement generates up to 0.96 SD worse health trajectory.

[INSERT TABLE 12]

Finally, we estimate the models while including the interactions between age terms and country dummies. Table 13 displays the results. The interactions are often jointly significant, but the general effects of being retired or time spent in retirement on self-assessed health are very similar, despite the fact that we add thirty interaction variables. The only qualitative difference is that the impact of being retired on the self-assessed cut-off point turns marginally insignificant when using both the early and regular retirement ages as instruments. Nevertheless, the impact of time spent in retirement is always significant also in these estimations. We thus conclude that our overall results are robust to including age-country interactions also in our second strategy, and that the coefficients in our main equations are not affected more than marginally. Apart from the fact that the Hansen-J test was significant and the Hausman test was insignificant in Model 12, all tests satisfy the conventional criteria.[28]

[INSERT TABLE 13]

These results display that our findings are robust even when using the instruments utilized in previous multi-country research, but with an improved research design as well as focusing on longer-term effects of retirement. Moreover, although we estimate different LATEs, the

---

[28] In unreported robustness tests, to ensure that country-specific common health shocks do not bias our findings, we included country-time interactions as well as the age interactions. Results were almost identical.

effect of being retired remains similar to our estimates using spouses' retirement status and age as instruments while the impact of time spent in retirement increases.

6.4   Estimation Strategy Matters

What explains the differences between our findings and the results in prior multi-country studies? Table 14 displays how controlling for individual-level and time-fixed effects are important for making the instruments based on retirement ages unbiased. To analyze this, we estimate exactly the same regressions as in Table 10, but exclude individual- and time-fixed effects. Instead, we include country-fixed effects while also controlling for education in addition to marital status. And, interestingly, we now find a highly significant, *positive* impact of being retired on self-assessed health in Model 1, which excludes the LDV. The Hansen-J test, however, rejects the instruments as exogenous. Since the Hansen-J test in the comparable Model 4 in Table 10 is strongly insignificant, this indicates that there are more significant differences between the early and regular retirement age when not including fixed effects. But in Model 2, where the LDV is included, the positive impact turns insignificant while the Hansen-J test continues to indicate that the instruments differ from each other.[29] Indeed, in unreported regressions, we used the retirement ages as instruments separately when controlling for the LDV. The impact then turned significantly positive in the case of the early retirement age, and insignificantly negative in the case of the regular retirement age. This is in line with our argument that individual-level fixed effects can be important for removing bias in the instruments.

[INSERT TABLE 14]

Similarly, we now find little evidence of a negative impact from spending longer time in retirement. Instead, in Model 3 in which the LDV is excluded, the impact is positive and only marginally insignificant (p=0.11). Also, unlike previous results, the early retirement instrument is now a significant predictor of time spent in retirement. This indicates that there

---

[29] These results remain when including time-fixed effects, suggesting that individual-fixed effects are driving the findings. We only include age and age squared since we again had problems with singularity when including age cube and excluding individual-level fixed effects. However, results in the second-stage equation are similar also when including age cube. The results are also robust to including age-country interactions when using age and age squared. Including age-country interactions together with age cube, however, produced the same problems as in our first strategy (see footnote 24) and the positive impact of being retired then turned insignificant.

are time-invariant omitted variables that are correlated with time spent in retirement and the early retirement age. Furthermore, the Hansen-J test continues to reject the instruments as exogenous, in sharp contrast to the comparable Models 6 and 8 in Table 10.[30] Again, in unreported regressions, we found that using the early retirement age as sole instrument produced significantly positive effects, while using the regular retirement age as sole instrument produced highly insignificant negative effects. Overall, therefore, we conclude that one key reason for why our results differ compared to prior multi-country research, using the same instruments, is that we partial out time-invariant omitted variables.

6.5   Do Short-Term and Medium- to Long-Term Effects Differ?

It is, however, also important to investigate whether the non-lagged and lagged impacts of being retired and spending longer time in retirement differ. The results from this exercise, again controlling for individual- and time-fixed effects, are displayed in Table 15. Using only the regular retirement age as instrument, we find the effect of being retired to be insignificantly positive. However, when including the early retirement age as instrument the positive impact actually turns significant at the 10% level, in sharp contrast to the lagged effects. The Hansen-J test is now also marginally insignificant. This may indicate that the immediate impact of retiring is positive, a finding that is wholly dependent on using the instrument constructed from the early retirement age. Given the Hansen-J test's marginal acceptance of the instruments, however, there are still questions regarding their validity. Note that, due to data availability, we cannot include the LDV or perform a similar analysis of health trajectory as above, which previous estimates suggest are important. Furthermore, we do not find any evidence that the non-lagged measure of time spent in retirement has a significantly positive impact regardless of specification, while the Hansen-J test now strongly rejects the exogeneity of the instruments. In this case, also, reaching the early retirement age has a discontinuously positive effect on health. Finally, the impact on the threshold for good health is also insignificant regardless of which instruments we include in the analysis. This indicates that the positive effect of the non-lagged retirement indicator is far from robust.[31]

---

[30] In this case, however, the F-statistics display figures marginally below the conventionally accepted threshold. But the results hold up when using LIML, which is supposed to be more robust to weaker instruments than 2SLS (Staiger and Stock 1997).

[31] The positive impact of retiring was also detected when analyzing the health stock, but not mental health. We also note, however, that time spent in retirement was strongly insignificant in these models. In addition, the Hansen-J tests were always significant, again suggesting bias in the early retirement age. When including two terms of age only, the positive impact of retiring was marginally significant in the case of mental health, while

[INSERT TABLE 15]

Nevertheless, these results, both significant and insignificant, differ considerably from our previous findings, which may very well lend support to our prediction that continuing employment in old age has different short- and longer-term consequences.[32] However, differential lagged and non-lagged impacts are not the only plausible explanation. It might also be the case that the instruments based on retirement ages are endogenous when not taking lagged effects into account. This is because the health impact of retirement planning and/or self-selection within the state pension system may subside relatively quickly following retirement. If so, endogeneity might be a more significant problem when examining instantaneous and short-term health impacts in comparison to medium- to long-term ones.

To shed more light on this issue, we again use our first strategy but presume that the impact does not operate with a lag while also controlling for spouses' self-assessed health. First, we want to compare individuals who recently retired with people who remain in employment, and we thus exclude all observations denoting people who have been retired for three years or longer. Models 1-6 in Table 16 display no significant effects of either of the retirement variables. This lends further support to our prediction that short- and longer-term effects differ substantially. Models 7-12 include all retirees but controls for the LDV. While this strategy does not separate short- from longer-term effects of retirement properly, it investigates whether the health-preserving impacts of remaining employed and spending shorter time in retirement operate non-lagged. The effects of the retirement variables are now consistently negative only in the case of the health stock variable. Overall, we thus find that assuming lagged effects is important also in the first strategy. However, we never find a significantly positive non-lagged impact as in the second strategy. Furthermore, the Hansen-J test sometimes turns significant, which indicates that the instruments are not necessarily jointly valid unless we assume lagged effects.

[INSERT TABLE 16]

---

the effect of time spent of retirement was marginally significant and positive in the case of health stock. This is probably because the early retirement age is a significant predictor of time spent in retirement only when excluding age cube, suggesting misspecification without the latter. And again, the Hansen-J test rejected the instruments as exogenous. While we cannot compare these results with models assuming lagged effects due to data availability, we nonetheless note that this supports the argument that the early and regular retirement ages differ significantly as instruments when not taking lagged effects into account.

[32] Interestingly, the marriage dummy is negative in Models 5-8, which contrasts to its lagged impact.

While the short-term health effects of being retired and spending longer time in retirement are somewhat uncertain, therefore, we nevertheless conclude that the negative longer-term relationships are very robust. Our overall results thus find strong support for the thesis that employment in old age has health-preserving effects in a longer perspective, and that, vice versa, spending longer time in retirement leads to a more rapid deterioration in health.

6.6 Are There Differential Gender Effects?

Finally, we also investigate potential differences between women and men. To do so, we construct interaction variables between the female dummy and our main predictors as well as with the instruments. We treat the interactions between the female dummy and the retirement indicators as endogenous, while using the female-instrument interactions as additional instruments. Table 17 displays the results. Using our first strategy, the interaction effect with retirement status indicates that women lose more from being retired than men in terms of the health stock and to a certain extent also mental health. However, the interaction effect is insignificant in the equation explaining self-assessed health, and when analyzing the impact of time spent outside the labor force, the interaction is never significant regardless of dependent variable.[33]

Using our second strategy to analyze self-assessed health, the interaction effect is also always insignificant. This further displays the similarity between the findings from our two distinct estimation strategies. We also find that the interaction between marital status and the female dummy is highly significant and positive, suggesting that women benefit from marriage. Men, however, do not and there is actually some evidence that marriage has a negative association with men's self-assessed health. Overall, these results thus indicate that the negative effects of retirement accrue to both men and women.[34]

[INSERT TABLE 17]

---

[33] In these estimations, we include the LDV and the index constructed from respondents' and spouses' objective health indicators as well as the latters' self-assessed health. When including spouses' self-assessed health directly instead, the results were similar. The only differences were that the main effect in Models 1 and 4 and the interaction effect in Model 2 turned insignificant. In other words, the overall conclusion that the negative health impact of retirement applies to both women and men remains.

[34] The results are robust to including age-country interactions. In unreported regressions, using the first strategy, we also analyzed physical health indicators as well as cut-off points using both strategies. Again, the retirement-female interaction was either insignificant or indicated that women gained slightly more from not retiring. This further suggests that the negative retirement impact on health applies to both men and women.

7   Conclusion

Following large pension reforms in developed countries, researchers have begun to evaluate the relationship between retirement and health. The best research displays significantly heterogeneous findings, which is likely due to different methodologies. Using two distinct identification strategies and panel data from eleven European countries, obtained from SHARE, we find relatively large negative, longer-term effects on health of being retired and spending longer time in retirement.

In our first strategy, which uses spouses' characteristics as instruments, we find that retirement has a significantly negative impact on self-assessed, general, mental, and physical health. These results are robust to a number of robustness checks. Conspicuously, we show that the results regarding self-assessed health hold also in our second strategy, which employs previously used instruments constructed from the official retirement ages, when controlling for unobserved heterogeneity between individuals and taking into account lagged effects. These two innovations also turn out to explain the differences between our findings and previous results from multi-country research in this respect. Due to data availability, we cannot carry out this analysis using other health indicators, but it nevertheless displays the robustness of our findings. Our analysis thus displays longer-term health gains from continuing paid work rather than retiring. This impact applies to both men and women.

The main differences compared to prior multi-country research are (1) that we take into account that short- and longer-term effects may differ substantially, (2) analyze both the impacts of being retired and time spent in retirement, and (3) use alternative research strategies. Prior multi-country research has only used official retirement ages as instruments without taking sufficient precautions to ensure that these are valid. We utilize both spouses' characteristics and official retirement ages as instruments, while also taking a number of steps to rid these of endogeneity. We note that our findings are in line with research showing that general practitioners note a drop in various health indicators as a result of early retirement, despite the fact that their patients often believe that retirement has positive effects on their health (Maes and Stammen 2011).

However, the fact that the non-lagged effect of being retired on self-assessed health turns positive when using the second strategy, while in one case becoming marginally significant, may indicate that the short-term impact differs from medium- to long-term effects. It could also mean that the instruments constructed from the official retirement ages are only valid when assuming lagged effects since the impact of planning for, and/or selection into,

retirement at those ages on health may disappear in a slightly longer-term perspective. Since we find that assuming lagged effects are important also in our first strategy, but that the non-lagged impact never turns significantly positive, both explanations are possible. Future country-specific research should clearly take both issues seriously.

It is important to point out that we are only able to estimate the causal impact of retirement on health among individuals who respond to their spouses' employment status and age—as well as to the financial incentives captured by the official retirement ages in each country—when deciding to retire. While the impact is often similar across the different subpopulations, future research should investigate how other subpopulations' health is affected by retirement. Furthermore, while we have shed more light on the health-preserving impact of working among older people, we know little about the actual mechanisms through which it operates—which would be of immense value for the purpose of countering the negative effects. This, therefore, is clearly also an important topic for future research to investigate.

Finally, despite these caveats, we can nevertheless draw significant policy conclusions from our findings. While the immediate health impact of retirement is somewhat uncertain, it seems clear that there are significant medium- to long-term health benefits of remaining in paid employment among older people. This, in turn, indicates that European politicians do not face a trade-off—at least in a longer-term perspective—between producing solvent pension systems, stronger economic performance and a healthier population. On the contrary, our findings lend credence to a win-win scenario in which all three go hand in hand.

References

Alavinia, Seyed M. and Alex Burdorf. 2008. "Unemployment and Retirement and Ill-Health: A Cross-Sectional Analysis across European Countries." *International Archives of Occupational and Environmental Health* 82:39-45.

Ambrasat, Jens, Jürgen Schupp, and Gert G. Wagner. 2011. "Comparing the Predictive Power of Subjective and Objective Health In- dicators: Changes in Handgrip Strength and Overall Satisfaction with Life as Predictors of Mortality." SOEPpapers on Multidisciplinary Panel Data Research No. 398-2011, German Socio-Economic Panel Study , German Institute for Economic Research, Berlin.

Angrist, Joshua D. and Alan B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* 15(4):69-85.

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

Asch, Beth, Steven J. Haider, and Julie Zissimopolous. 2005. "Financial Incentives and Retirement: Evidence from Federal Civil Service Workers." *Journal of Public Economics* 89(2-3):427-440.

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

Banks, James, Richard Blundell, and Maria R. Casanova. 2010. "The Dynamics of Retirement Behavior in Couples: Reduced-Form Evidence from England and the US." Unpublished Manuscript, University College London, London.

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

Bingley, Paul and Alessandro Martinello. 2011. "Retirement Improves Cognitive Performance." Working Paper 07:2011, Danish National Center for Social Research, Copenhagen.

Bingley, Paul and Peder J. Pedersen. 2011. "Pension Reform, Early Retirement, Health and Mortality: A 30 year follow-up study." Unpublished Manuscript, The Danish National Center for Social Research and School of Economics and Management, Aarhus University, Copenhagen/Aarhus.

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

Bonsang, Eric, Stéphane Adam, and Sergio Perelman. 2012. "Does Retirement Affect Cognitive Functioning?" *Journal of Health Economics* 31(3):490-501.

Bonsang, Eric and Tobias J. Klein. 2011. "Retirement and Subjective Well-Being." Discussion Paper No. 5536, Institute for the Study of Labor, Bonn.

Bound, John, Michael Schoenbaum, Todd R. Stinebrickner, and Timothy Waidmann. 1999. "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers." *Labour Economics* 6(1):179-202.

Bound, John and Timothy Waidmann. 2007. "Estimating the Health Effects of Retirement." Working Paper 2007-168, Michigan Retirement Research Center, University of Michigan , Ann Arbor, MI.

Brockman, Hilke, Rolf Müller, and Uwe Helmert. 2009. "Time to Retire - Time to Die? A Prospective Cohort Study of the Effects of Early Retirement on Long-Term Survival." *Social Science & Medicine* 69:160-164.

Cai, Linxin. 2010. "The Relationship Between Health and Labour Force Participation: Evidence from a Panel Data Simultaneous Equation Model." *Labour Economics* 17(1):77-90.

Calvo, Esteban, Natalia Sarkisian, and Christopher R. Tamborini. 2011. "Causal Effects of Retirement Timing on Subjective Well-being: The Role of Culture Norms and Institutional Policies." Working Paper No.17, Facultad de Economía y Empresa, Universidad Diego Portales, Santiago.

Casanova, María. 2010. "Happy Together: A Structural Model of Couples' Joint Retirement Choices." Unpublished Manuscript, Department of Economics, UCLA, Los Angeles.

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

Chung, Sukyung, Marisa E. Domino, and Sally C. Stearns. 2009. "The Effect of Retirement on Weight." *Journal of Gerontology: Social Sciences* 64B(5):656-665.

Chung, Sukyung, Marisa E. Domino, Sally C. Stearns, and Barry M. Popkin. 2009. "Retirement and Physical Activity: Analyses by Occupation and Wealth." *American Journal of Preventive Medicine* 36(5):422-428.

Coe, Norma B. and Maarten Lindeboom. 2008. "Does Retirement Kill You? Evidence from Early Retirement Windows." Discussion Paper No. 3817, Institute for the Study of Labor (IZA), Bonn.

Coe, Norma B., Hans-Martin von Gaudecker, Maarten Lindeboom, and Jürgen. 2012. "The Effect of Retirement on Cognitive Functioning." *Health Economics* 21(8):913-927.

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

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

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

de Griep, Andries, Arnaud Dupuy, Jelle Jolles, and Martin van Boxtel. 2012. "Retirement and Cognitive Development: Are the Retired Really Inactive?" Working Paper No. 2012/14, Maastricht School of Management, University of Maastricht, Maastricht.

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

d'Hombres, Béatrice, Lorenzo Rocco, Marc Suhrcke, and Martin McKee. 2010. "Does Social Capital Determine Health." *Health Economics* 19(1):56-74.

Euwals, Rob and Elisabetta Trevisan. 2011. "Early Retirement and Financial Incentives: Differences Between High and Low Wage Earners." Discussion Paper 12/2011-105, Network for Studies on Pensions, Aging and Retirement, Tilburg.

Fé, Eduardo and Bruce Hollingsworth. 2011. "Estimating the Effect of Retirement on Health via Panel Discontinuity Designs." Unpublished Manuscript, University of Manchester/University of Lancaster.

Fé, Eduardo. and Bruce Hollingsworth. 2012. "Estimating the Effect of Retirement on Mental Health via Panel Discontinuity Designs." Unpublished Manuscript, University of Manchester/Lancaster University.

Fonseca, Raquel and Yuhui Zheng. 2011. "The Effect of Education on Health: Cross-Country Evidence." Working Paper WR-864, RAND Labor and Population, Rand Corporation, Santa Monica, CA.

García-Gómez, Pilar, Andrew M. Jones, and Nigel Rice. 2010. "Health Effects on Labour Market Exits and Entries." *Labour Economics* 17(1):62-76.

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

Gruber, Jonathan and David Wise, eds. 2004. *Micro Modelling of Retirement Incentives in the World*. Chicago: NBER and Chicago University Press.

Hanel, Barbara. 2011. "The Effect of Disability Pension Incentives on Early Retirement Decisions." Melbourne Institute Working Paper No. 5/11 , Melbourne Institute of Applied Economic and Social Research, Melbourne.

Hernaes, Erik, Simen Markussen, John Piggot, and Ola L. Vestad. 2012. "Does Retirement Age Impact Mortality?" Working Paper, ARC Centre of Excellence in Population Ageing Research (CEPAR), University of New South Wales, Sydney.

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

Insler, Michael A. 2012. "The Health Consequences of Retirement." Unpublished Manuscript, Department of Economics, United States Naval Academy, Annapolis, MD.

Jang, Soong-Nang, Sung-Il Cho, Jiyeun Change, Kachung Boo, Hyun-Goo Shin, Hyejung Lee, and Lisa F. Berkman. 2009. "Employment Status and Depressive Symptoms in Koreans: Results From a Baseline Survey of the Korean Longitudinal Study of Aging." *Journal of Gerontology: Social Sciences* 64B(5):677-683.

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

Kajitani, Shinya. 2011. "Working in Old Age and Health Outcomes in Japan." *Japan and the World Economy* 23:153-162.

Kantarci, Tunga and Arthur van Soest. 2011. "Effects of Partial and No Retirement on Health in the United States." Unpublished Manuscript, Department of Econometrics, Tilburg University, Tilburg.

Kerkhhofs, Marcel and Maarten Lindeboom. 1997. "Age Related Health Dynamics and Changes in Labour Market Status." *Health Economics* 6:407-423.

Koball, Heather L., Emily Moiduddin, Jamila Henderson, Brian Goesling, and Melanie Besculides. 2010. "What Do We Know About the Link Between Marriage and Health?" *Journal of Family Issues* 31(8):1019-1040.

Kolodziej, Ingo. 2011. "The Relationship between Retirement and Mental Health: Investigating the Causal Relationship in Eleven European Countries using SHARE." MSc Thesis 2011-060, Network for Studies on Pension, Aging and Retirement, Tilburg.

Kuhn, Andreas, Jean-Philippe Wuellrich, and Josef Zweimüller. 2010. "Fatal Attraction? Access to Early Retirement and Mortality." Discussion Paper No. 5160, Institute for the Study of Labor, Bonn.

Kuvaja-Köllner, Virpi, Hannu Valtonen, Pirjo Komulainen, Maija Hassinen, and Rainer Rauramaa. 2012. "The Impact of Time Cost of Physical Exercise on Health Outcomes by Older Adults: the DR's EXTRA Study." *European Journal of Health Economics* Doi: 10.1007/s10198-012-0390-y.

Latif, Ehsan. 2012. "The Impact of Retirement on Health in Canada." *Canadian Public Policy* 38(1):15-29.

Lei, Xiaoyan, Li Tan, and Yaohui Zhao. 2011. "The Impact of Retirement on Health: Evidence from China." Unpublished Manuscript, China Center for Economic Research, Peking University, Peking.

Lindeboom, Marten and Henrik Lindegaard. 2010. "The Impact of Early Retirement on Health." Unpublished Manuscript, Free University/Danish Institute of Governmental Research, Amsterdam/Copenhagen.

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

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

Maes, Marjan and Benjamin Stammen. 2011. "The Impact of (Early) Retirement on the Subsequent Physical and Mental Health of the Retired: A Survey Among General Practitioners in Belgium." HUB Research Paper 2011/03, Hogeschool-Universiteit Brussel, Brussels.

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

Michaud, Pierre-Carl and Arthur van Soest. 2008. "Health and Wealth of Elderly Couples: Causality Tests Using Dynamic Panel Data Models." *Journal of Health Economics* 27(5):1312-1325.

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

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

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

Perelman, Sergio, Stéphane Adam, and Eric Bonsang. 2012. "Does Retirement Affect Cognitive Functioning?" *Journal of Health Economics* 31:490-501.

Petrou, Stavros and Emil Kupek. 2008. "Social Capital and Its Relationship with Measures of Health Status: Evidence from the Health Survey for England 2003." *Health Economics* 17(1):127-143.

Roberts, Beverly A., Rebecca Fuhrer, Michael Marmot, and Marcus Richards. 2011. "Does Retirement Influence Cognitive Performance? The Whitehall II Study." *Journal of Epidemiology & Community Health* 65:958-963.

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

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

Sasaki, Hideo, Fumiyoshi Kasagi, Michiko Yamada, and Fujita Schoichiro. 2007. "Grip Strength Predicts Cause-Specific Mortality in Middle-Aged and Elderly Persons." *American Journal of Medicine* 120(4):337-342.

Scheubel, Beatrice, Daniel Schunk, and Joachim Winter. 2009. "Don't Raise the Retirement Age! An Experiment on Opposition to Pension Reforms and East-West Differences in Germany." CESifo Working Paper No. 2752, Center for Economic Studies, Munich.

SSA and ISSA. 2010. *Social Security Programs Throughout the World: Europe, 2010*. Social Security Administration and International Social Security Association.

Staiger, Douglas and James H. Stock. 1997. "Instrumental Variables with Weak Instruments." *Econometrica* 65(3):557-586.

Stancanelli, Elena. 2012. "Spouses' Retirement and Hours Outcomes: Evidence from Twofold Regression Discontinuity with Differences-in-Differences." Discussion Paper No. 6791, Institute for the Study of Labor, Bonn.

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

Zucchinelli, Eugenio, Andrew M. Jones, Nigel Rice, and Anthony Harris. 2010. "The Effects of Health Shocks on Labour Market Exits: Evidence from the HILDA Survey." *Australian Journal of Labour Economics* 13(2):191-218.

Table 1: Descriptive Statistics

	First Strategy				Second Strategy			
Demographics	Average	Standard Deviation	Min	Max	Average	Standard Deviation	Min	Max
Age (50-69) at the first interview	59.84	5.62	50	69.92	59.69	5.64	50	69.92
Female	0.50	0.50	0	1	0.55	0.50	0	1
Education	2.73	1.50	0	6	2.75	1.52	0	6
Over full retirement age$_{t-1}$	0.31	0.46	0	1	0.36	0.48	0	1
Over early retirement age$_{t-1}$	0.57	0.50	0	1	0.60	0.49	0	1
Married$_{t-1}$	0.98	0.14	0	1	0.78	0.41	0	1
Health indicators								
Self-assessed health$_t$	3.10	1.03	1	5	3.02	1.04	1	5
ΔSelf-assessed health$_t$					-0.16	0.94	-4	4
Euro-D scale$_t$	10.11	2.01	1	12				
Health stock$_t$	3.21	0.49	0.52	3.90				
Physical conditions$_t$	10.84	1.22	4	12				
Drugs taken for physical conditions$_t$	8.98	1.18	1	10				
Grip strength$_t$	37.31	11.82	3	74				
Good self-assessed health$_t$	0.32	0.47	0	1	0.29	0.46	0	1
No depression$_t$	0.18	0.39	0	1				
No physical condition$_t$	0.37	0.48	0	1				
No drug against physical condition$_t$	0.43	0.50	0	1				
Retirement variables								
Retired$_{t-1}$	0.60	0.49	0	1	0.63	0.48	0	1
Time spent in retirement (log)$_{t-1}$	0.98	1.14	0	3.96	1.12	1.19	0	4.06

Table 2: First Strategy – Results Including Lagged Health

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Self-assessed health	Euro-D	Health stock	Self-assessed health	Euro-D	Health stock
Retired$_{t-1}$	-1.03**	-2.48***	-0.81***			
	(0.44)	(0.86)	(0.21)			
Time spent in retirement (log)$_{t-1}$				-0.35**	-0.93***	-0.28***
				(0.14)	(0.26)	(0.06)
Age$_t$	0.02	0.09	-0.01	-0.14*	-0.33**	-0.15***
	(0.04)	(0.10)	(0.03)	(0.08)	(0.16)	(0.04)
Age^2_t	0.001	-0.002	-0.002	0.01*	0.03**	0.01***
	(0.004)	(0.01)	(0.002)	(0.01)	(0.01)	(0.004)
Female	0.08	-0.03	0.06**	0.09	0.01	0.08***
	(0.06)	(0.12)	(0.03)	(0.06)	(0.12)	(0.03)
Education	0.02	-0.05	-0.01	0.03***	-0.02	0.002
	(0.02)	(0.03)	(0.01)	(0.01)	(0.02)	(0.01)
Spouse's education	0.02**	0.03*	0.01	0.02**	0.03	0.003
	(0.01)	(0.02)	(0.005)	(0.01)	(0.02)	(0.005)
Over early retirement age$_{t-1}$	0.18**	0.57***	0.17***	0.03	0.20**	0.05**
	(0.09)	(0.18)	(0.04)	(0.04)	(0.09)	(0.02)
Spouse over early retirement age$_{t-1}$	0.07	0.34***	0.06**	0.01	0.22***	0.02
	(0.05)	(0.09)	(0.02)	(0.03)	(0.07)	(0.02)
LDV	0.47***	0.44***	0.62***	0.48***	0.46***	0.65***
	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
First stage						
Retired spouse$_{t-1}$	0.05***	0.04***	0.04***	0.17***	0.17***	0.17***
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)
Spouse's age (log)$_{t-1}$	0.26***	0.31***	0.31***	0.66***	0.75***	0.75***
	(0.07)	(0.07)	(0.07)	(0.18)	(0.18)	(0.18)
Hausman test (p-value)	0.02	0.00	0.00	0.02	0.00	0.00
Hansen-J test (p-value)	0.66	0.17	0.47	0.95	0.34	0.91
F-statistic	14.08	15.80	14.77	24.21	25.35	23.42
N	6,751	6,631	6,140	6,239	6,130	5,699

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses. All regressions include country-fixed effects.

Table 3: First Strategy – Results Excluding Lagged Health

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Self-assessed health	Euro-D	Health stock	Self-assessed health	Euro-D	Health stock
Retired$_{t-1}$	-1.83***	-3.26***	-1.07***			
	(0.50)	(0.98)	(0.24)			
Time spent in retirement (log)$_{t-1}$				-0.60***	-1.11***	-0.37***
				(0.16)	(0.30)	(0.08)
Age$_t$	-0.04	0.12	-0.02	-0.31***	-0.36*	-0.20***
	(0.06)	(0.12)	(0.03)	(0.10)	(0.18)	(0.05)
Age^2_t	0.006	-0.003	0.004	0.03***	0.04**	0.02***
	(0.005)	(0.01)	(0.003)	(0.01)	(0.02)	(0.004)
Female	0.13*	-0.29**	0.01	0.12*	-0.27**	-0.01
	(0.08)	(0.15)	(0.04)	(0.07)	(0.13)	(0.03)
Education$_t$	0.04**	-0.03	0.01	0.07***	0.01	0.02***
	(0.02)	(0.04)	(0.01)	(0.01)	(0.03)	(0.01)
Spouse's education$_t$	0.05***	0.05**	0.02***	0.05***	0.04*	0.01**
	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)
Over early retirement age$_{t-1}$	0.36***	0.74***	0.23***	0.09*	0.23**	0.08***
	(0.11)	(0.21)	(0.05)	(0.05)	(0.10)	(0.03)
Spouse over early retirement age$_{t-1}$	0.12**	0.40***	0.06**	0.03	0.24***	0.01
	(0.05)	(0.11)	(0.03)	(0.04)	(0.08)	(0.02)
First stage						
Retired spouse$_{t-1}$	0.05***	0.04***	0.05***	0.18***	0.17***	0.17***
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)
Spouse's age (log)$_{t-1}$	0.30***	0.32***	0.33***	0.73***	0.75***	0.79***
	(0.07)	(0.07)	(0.07)	(0.18)	(0.18)	(0.18)
Hausman test (p-value)	0.00	0.00	0.00	0.00	0.00	0.00
Hansen-J test (p-value)	0.87	0.27	0.78	0.45	0.49	0.26
F-statistic	16.74	16.64	17.62	26.52	26.06	26.06
N	6,752	6,675	6,354	6,240	6,167	5,887

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses. All regressions include country-fixed effects

Table 4: First Strategy – Robustness Test 1 (Controlling for Lagged Health and Spouses' Health Stock)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Self-assessed health	Euro-D	Health stock	Self-assessed health	Euro-D	Health stock
Retired$_{t-1}$	-1.02**	-2.58***	-0.74***			
	(0.46)	(0.91)	(0.20)			
Time spent in retirement (log)$_{t-1}$				-0.33**	-0.92***	-0.25***
				(0.14)	(0.26)	(0.06)
Spouse health stock(1)$_{t-1}$	0.39***	0.23*	-0.36***	0.40***	0.26**	-0.33***
	(0.05)	(0.13)	(0.08)	(0.04)	(0.10)	(0.08)
LDV	0.36***	0.41***	0.98***	0.38***	0.42***	0.97***
	(0.02)	(0.02)	(0.08)	(0.02)	(0.02)	(0.08)
First stage						
Retired spouse$_{t-1}$	0.04***	0.04***	0.04***	0.18***	0.18***	0.17***
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)
Spouse's age (log)$_{t-1}$	0.27***	0.29***	0.31***	0.69***	0.72***	0.77***
	(0.07)	(0.07)	(0.07)	(0.18)	(0.18)	(0.19)
Hausman test (p-value)	0.02	0.00	0.00	0.02	0.00	0.00
Hansen-J test (p-value)	0.90	0.19	0.60	0.62	0.47	0.73
F-statistic	12.61	14.00	14.56	24.11	25.19	24.27
N	6,159	6,115	5,901	5,716	5,676	5,481

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses. All regressions include country-fixed effects as well as all the other control variables included in Tables 2 and 3.

Table 5: First Strategy – Robustness Test 2 (Controlling for Lagged Health and Spouses' Self Assessed Health)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Self-assessed health	Euro-D	Health stock	Self-assessed health	Euro-D	Health stock
Retired$_{t-1}$	-0.73*	-1.94**	-0.77***			
	(0.43)	(0.83)	(0.21)			
Time spent in retirement (log)$_{t-1}$				-0.27**	-0.75***	-0.27***
				(0.13)	(0.25)	(0.06)
Spouse self-assessed health$_t$	0.06***	0.11***	0.01	0.06***	0.12***	0.01
	(0.01)	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)
LDV	0.47***	0.44***	0.62***	0.48***	0.45***	0.65***
	(0.03)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
First stage						
Retired spouse$_{t-1}$	0.05***	0.04***	0.04***	0.17***	0.17***	0.17***
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)
Spouse's age (log)$_{t-1}$	0.26***	0.30***	0.31***	0.67***	0.74***	0.76***
	(0.07)	(0.07)	(0.07)	(0.18)	(0.18)	(0.18)
Hausman test (p-value)	0.10	0.02	0.00	0.08	0.01	0.00
Hansen-J test (p-value)	0.70	0.21	0.51	0.97	0.33	0.88
F-statistic	13.92	14.95	14.51	24.11	24.54	23.74
N	6,735	6,613	6,126	6,225	6,114	5,687

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses. All regressions include country-fixed effects as well as all other control variables included in Tables 2 and 3.

Table 6: First Strategy – Robustness Test 3 (Controlling for Spouses' Self Assessed Health)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Self-assessed health	Euro-D	Health stock	Self-assessed health	Euro-D	Health stock
Retired$_{t-1}$	-1.22**	-2.23**	-0.93***			
	(0.48)	(0.93)	(0.24)			
Time spent in retirement (log)$_{t-1}$				-0.40***	-0.77***	-0.30***
				(0.15)	(0.29)	(0.07)
Spouse self-assessed health$_t$	0.13***	0.21***	0.04***	0.13***	0.22***	0.05***
	(0.02)	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)
First stage						
Retired spouse$_{t-1}$	0.05***	0.04***	0.05***	0.17***	0.17***	0.17***
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)
Spouse's age (log)$_{t-1}$	0.29***	0.30***	0.32***	0.72***	0.74***	0.79***
	(0.07)	(0.07)	(0.07)	(0.18)	(0.18)	(0.18)
Hausman test (p-value)	0.03	0.03	0.00	0.04	0.02	0.00
Hansen-J test (p-value)	0.72	0.35	0.68	0.42	0.47	0.23
F-statistic	15.33	15.28	16.40	25.15	24.90	25.33
N	6,736	6,657	6,337	6,226	6,151	5,872

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses. All regressions include country-fixed effects as well as all other control variables included in Tables 2 and 3.

Table 7: First Strategy – Robustness Test 4 (Physical Health)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Physical conditions	Drugs against physical conditions	Grip strength	Physical conditions	Drugs against physical conditions	Grip strength
Retired$_{t-1}$	-1.48***	-1.14**	-15.10***			
	(0.52)	(0.46)	(4.81)			
Time spent in retirement (log)$_{t-1}$				-0.46***	-0.35**	-4.42***
				(0.15)	(0.14)	(1.32)
Self-assessed health$_{t-1}$	0.06**	0.05**	-0.72***	0.10***	0.08***	-0.36**
	(0.03)	(0.03)	(0.26)	(0.02)	(0.02)	(0.16)
Spouse self-assessed health$_t$	-0.004	-0.001	-0.19	-0.01	0.003	-0.14
	(0.01)	(0.01)	(0.12)	(0.01)	(0.01)	(0.12)
LDV$_{t-1}$	0.54***	0.64***	0.53***	0.52***	0.63***	0.55***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
First stage						
Retired spouse$_{t-1}$	0.05***	0.05***	0.04***	0.17***	0.17***	0.16***
	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.03)
Spouse's age (log)$_{t-1}$	0.26***	0.26***	0.28***	0.66***	0.66***	0.76***
	(0.07)	(0.07)	(0.07)	(0.18)	(0.18)	(0.18)
Hausman test (p-value)	0.00	0.01	0.00	0.00	0.02	0.00
Hansen-J test (p-value)	0.54	0.24	0.39	0.35	0.14	0.24
F-statistic	13.83	13.66	12.96	23.80	23.55	23.16
N	6,728	6,731	6,196	6,220	6,222	5,753

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses. All regressions include country-fixed effects as well as all other control variables included in Tables 2 and 3. Models 3 and 6 also include the Euro-D index non-lagged.

Table 8: First Strategy – Robustness Test 5 (Cut-off points)

	Model 1	Model 2	Model 3	Model 4
			No physical conditions	No drug against physical conditions
	Good health	No depression		
Retired$_{t-1}$	-0.39**	-0.41**	-0.63***	-0.60***
	(0.20)	(0.17)	(0.22)	(0.20)
First stage				
Retired spouse$_{t-1}$	0.05***	0.04***	0.05***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)
Spouse's age (log)$_{t-1}$	0.28***	0.29***	0.28***	0.28***
	(0.07)	(0.07)	(0.07)	(0.07)
Hausman test (p-value)	0.05	0.02	0.00	0.00
Hansen-J test (p-value)	0.12	0.26	0.94	0.41
F-statistic	15.43	14.91	15.23	15.11
N	6,735	6,612	6,728	6,731
	Model 5	Model 6	Model 7	Model 8
			No physical conditions	No drug against physical conditions
	Good health	No depression		
Time spent in retirement (log)$_{t-1}$	-0.11*	-0.17***	-0.22***	-0.19***
	(0.06)	(0.05)	(0.07)	(0.06)
First stage				
Retired spouse$_{t-1}$	0.18***	0.17***	0.18***	0.17***
	(0.03)	(0.03)	(0.03)	(0.03)
Spouse's age (log)$_{t-1}$	0.69***	0.74***	0.69***	0.70***
	(0.18)	(0.18)	(0.18)	(0.18)
Hausman test (p-value)	0.14	0.00	0.00	0.00
Hansen-J test (p-value)	0.03	0.40	0.62	0.19
F-statistic	25.11	24.90	24.89	24.85
N	6,225	6,114	6,221	6,223

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses. All regressions include country-fixed effects as well as all other control variables included in Tables 2 and 3. The LDV, the lagged self-assessed cut-off point, and spouses' self-assessed cut-off point are also included.

Table 9: First Strategy – Robustness Test 6 (Including Age-Country Interactions)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Self-assessed health	Euro-D	Health stock	Physical conditions	Drugs against physical conditions	Grip strength
Retired$_{t-1}$	-0.82*	-1.81*	-0.73***	-1.70***	-1.31**	-17.23***
	(0.49)	(0.92)	(0.22)	(0.61)	(0.54)	(5.58)
Interactions joint significance (p-value)	0.24	0.57	0.16	0.04	0.14	0.80
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Time spent in retirement (log)$_{t-1}$	-0.29**	-0.76***	-0.27***	-0.48***	-0.34**	-5.28***
	(0.14)	(0.27)	(0.07)	(0.16)	(0.14)	(1.58)
Interactions joint significance (p-value)	0.24	0.49	0.09	0.02	0.08	0.17
	Model 13	Model 14	Model 15	Model 16		
	Good health	No Depression	No physical condition	No drug against physical condition		
Retired$_{t-1}$	-0.52**	-0.40**	-0.74***	-0.70***		
	(0.23)	(0.19)	(0.25)	(0.23)		
Interactions joint significance (p-value)	0.17	0.63	0.34	0.28		
	Model 17	Model 18	Model 19	Model 20		
Interactions joint significance (p-value)	-0.11*	-0.17***	-0.24***	-0.20***		
	(0.07)	(0.06)	(0.07)	(0.07)		
Interactions joint significance (p-value)	0.23	0.46	0.32	0.18		

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Robust standard errors in parentheses. Models 1-3 and 7-9 include the control variables in Table 5, Models 4-6 and 10-12 include control variables in Table 7, and Models 13-16 include control variables in Table 8. All models include also include interactions between country dummies and age as well as age square.

Table 10: Second Strategy – Main Results (IV-Fixed Effects with Instruments Constructed from Retirement Ages)

	Model 1	Model 2	Model 2	Model 4
	Self-assessed health	Self-assessed health	Self-assessed health	Self-assessed health
Retired$_{t-1}$	-0.99**	-0.76**	-1.19***	-1.08***
	(0.38)	(0.33)	(0.44)	(0.40)
Married$_{t-1}$	0.06	0.06	0.13	0.13
	(0.08)	(0.08)	(0.09)	(0.09)
Over early retirement age$_{t-1}$	0.06		0.03	
	(0.04)		(0.05)	
LDV	-0.48***	-0.48***		
	(0.01)	(0.01)		
First stage				
Over regular retirement age$_{t-1}$	0.09***	0.09***	0.09***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)
Over early retirement age$_{t-1}$		0.05***		0.05***
		(0.01)		(0.01)
Hausman test (p-value)	0.01	0.02	0.00	0.00
Hansen-J test (p-value)		0.12		0.52
F-statistic	42.17	29.13	41.61	25.54
n	9,121	9,121	9,122	9,122
N	18,242	18,242	18,244	18,244
	Model 5	Model 6	Model 7	Model 8
	Self-assessed health	Self-assessed health	Self-assessed health	Self-assessed health
Time spent in retirement (log)$_{t-1}$	-0.64***	-0.65***	-0.72***	-0.71***
	(0.23)	(0.23)	(0.26)	(0.26)
Married$_{t-1}$	0.05	0.05	0.11	0.11
	(0.08)	(0.08)	(0.09)	(0.08)
Over early retirement age$_{t-1}$	0.01		-0.02	
	(0.04)		(0.04)	
LDV	-0.47***	-0.47***		
	(0.01)	(0.01)		
First stage				
Over regular retirement age$_{t-1}$	0.14***	0.14***	0.14***	0.14***
	(0.01)	(0.01)	(0.01)	(0.01)
Over early retirement age$_{t-1}$		0.01		0.01
		(0.01)		(0.01)
Hausman test (p-value)	0.00	0.00	0.01	0.01
Hansen-J test (p-value)		0.72		0.54
F-statistic	122.27	61.36	122.66	61.54
n	8,395	8,395	8,400	8,400
N	16,790	16,790	16,800	16,800

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses. All regressions include age, age squared and age cube (all three are significant at the 1% level in all regressions) as well as individual-fixed effects and time-fixed effects.

Table 11: Second Strategy – Robustness Test 1 (Cut-Off Points with IV-Fixed Effects using Instruments Constructed from Retirement Ages)

	Model 1	Model 2	Model 2	Model 4
	Good health	Good health	Good health	Good health
Retired$_{t-1}$	-0.43**	-0.28*	-0.46**	-0.36*
	(0.19)	(0.17)	(0.22)	(0.19)
Married$_{t-1}$	-0.01	-0.01	0.03	0.03
	(0.04)	(0.03)	(0.04)	(0.04)
Over early retirement age$_{t-1}$	0.04**		0.03	
	(0.02)		(0.02)	
LDV	-0.46***	-0.46***		
	(0.01)	(0.01)		
First stage				
Over regular retirement age$_{t-1}$	0.09***	0.09***	0.09***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)
Over early retirement age$_{t-1}$		0.05***		0.05***
		(0.01)		(0.01)
Hausman test (p-value)	0.02	0.09	0.02	0.22
Hansen-J test (p-value)		0.04		0.04
F-statistic	41.74	25.73	41.61	25.54
n	9,121	9,121	9,122	9,122
N	18,242	18,242	18,244	18,244
	Model 5	Model 6	Model 7	Model 8
	Good health	Good health	Good health	Good health
Time spent in retirement$_{t-1}$	-0.26**	-0.27**	-0.28**	-0.28**
	(0.12)	(0.12)	(0.13)	(0.13)
Married$_{t-1}$	-0.001	-0.001	0.03	0.03
	(0.04)	(0.04)	(0.04)	(0.04)
Over early retirement age$_{t-1}$	0.02		0.003	
	(0.02)		(0.02)	
LDV	-0.45***	-0.45***		
	(0.01)	(0.01)		
First stage				
Over regular retirement age$_{t-1}$	0.14***	0.14***	0.14***	0.14***
	(0.01)	(0.01)	(0.01)	(0.01)
Over early retirement age$_{t-1}$		0.01		0.01
		(0.01)		(0.01)
Hausman test (p-value)	0.02	0.01	0.03	0.03
Hansen-J test (p-value)		0.40		0.89
F-statistic	122.50	61.47	122.66	61.54
n	8,395	8,395	8,400	8,400
N	16,790	16,790	16,800	16,800

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses. All regressions include age, age squared and age cube (all three are significant at the 1% level in all regressions) as well as individual-fixed effects and time-fixed effects.

Table 12: Second Strategy – Robustness Test 2 (IV-Fixed Effects using Instruments Constructed from Retirement Ages – Changes in Health Trajectory)

	Model 1	Model 2	Model 3	Model 4
	ΔSelf-assessed health	ΔSelf-assessed health	ΔSelf-assessed health	ΔSelf-assessed health
Retired$_{t-1}$	-1.60**	-1.73***		
	(0.71)	(0.62)		
Time spent in retirement(log)$_{t-1}$			-0.90**	-0.86**
			(0.43)	(0.43)
Married$_{t-1}$	0.27*	0.27*	0.24	0.24
	(0.15)	(0.15)	(0.15)	(0.15)
Over early retirement age$_{t-1}$	-0.04		-0.11	
	(0.07)		(0.07)	
First stage				
Over regular retirement age$_{t-1}$	0.09***	0.09***	0.14***	0.14***
	(0.01)	(0.01)	(0.01)	(0.01)
Over early retirement age$_{t-1}$		0.05***		0.01
		(0.01)		(0.01)
Hausman test (p-value)	0.01	0.00	0.06	0.07
Hansen-J test (p-value)		0.63		0.12
F-statistic	41.59	25.52	122.76	61.58
n	9,121	9,121	8,395	8,395
N	18,242	18,242	16,790	16,790

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses. All regressions include age, age squared and age cube (all three are significant at the 5% level in all regressions) as well as individual-fixed effects and time-fixed effects.

Table 13: Second Strategy – Robustness Test 3 (Including Age-Country Interactions)

	Model 1	Model 2	Model 3	Model 4	Model 5
	Self-assessed health	Self-assessed health	Self-assessed health	Self-assessed health	ΔSelf-assessed health
Retired$_{t-1}$	-0.87**	-0.68**	-1.05**	-0.98**	-1.43**
	(0.39)	(0.34)	(0.44)	(0.40)	(0.72)
Joint significance of age-country interactions (p-value)	0.00	0.00	0.06	0.06	0.12
	Model 6	Model 7	Model 8	Model 9	Model 10
Time spent in retirement (log)$_{t-1}$	-0.69**	-0.69**	-0.79**	-0.79**	-1.01*
	(0.29)	(0.29)	(0.33)	(0.33)	(0.54)
Joint significance of age-country interactions (p-value)	0.00	0.00	0.04	0.04	0.15
	Good health	Good health	Good health	Good health	ΔSelf-assessed health
	Model 11	Model 12	Model 13	Model 14	Model 15
Retired$_{t-1}$	-0.36*	-0.23	-0.39*	-0.29	-1.60**
	(0.19)	(0.17)	(0.22)	(0.20)	(0.66)
Joint significance of age-country interactions (p-value)	0.00	0.00	0.00	0.00	0.11
	Model 16	Model 17	Model 18	Model 19	Model 20
Time spent in retirement (log)$_{t-1}$	-0.26*	-0.26*	-0.28*	-0.28*	-1.01*
	(0.15)	(0.15)	(0.16)	(0.16)	(0.54)
Joint significance of age-country interactions (p-value)	0.00	0.00	0.00	0.00	0.15

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses. All regressions include country-fixed effects and interactions between country dummies and age, age square and age cube. Model 1, 6, 11, and 16 include the regular retirement age as instrument and LDV. Model 2, 7, 12, and 17 include the early and regular retirement ages as instruments and the LDV. Models 3, 8, 13, and 18 include the regular retirement age as instrument without the LDV. Models 4, 9, 14, and 19 include the early and regular retirement ages as instruments without the LDV. Models 5 and 10 include the regular retirement age as instrument. Models 15 and 10 include both the early and regular retirement ages as instruments.

Table 14: Second Strategy – Robustness Test 4 (Instruments Constructed from Retirement Ages - Excluding Fixed Effects)

	Model 1	Model 2	Model 3	Model 4
	Self-assessed health	Self-assessed health	Self-assessed health	Self-assessed health
Retired$_{t-1}$	0.38***	0.11		
	(0.14)	(0.09)		
Time spent in retirement (log)$_{t-1}$			0.45	0.04
			(0.27)	(0.16)
Married$_{t-1}$	0.08***	0.05***	0.06	0.05***
	(0.02)	(0.01)	(0.04)	(0.02)
Education	0.14***	0.06***	0.16***	0.05***
	(0.01)	(0.01)	(0.02)	(0.01)
LDV		0.55***		0.55***
		(0.01)		(0.02)
First stage				
Over regular retirement age$_{t-1}$	0.13***	0.13***	0.09***	0.10***
	(0.01)	(0.01)	(0.03)	(0.03)
Over early retirement age$_{t-1}$	0.19***	0.20***	0.09***	0.19***
	(0.01)	(0.01)	(0.03)	(0.03)
Hausman test (p-value)	0.00	0.00	0.02	0.54
Hansen-J test (p-value)	0.02	0.03	0.04	0.05
F-statistic	192.12	211.70	7.78	9.23
n	9,065	9,065	8,341	8,341
N	18,130	18,130	16,682	16,682

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses.

Table 15: Non-Lagged Effects with Strategy 2

	Model 1	Model 2	Model 2	Model 4
	Self-assessed health	Self-assessed health	Self-assessed health	Self-assessed health
Retired$_t$	0.43	0.67*		
	(0.39)	(0.37)		
Time spent in retirement(log)$_t$			0.22	0.15
			(0.27)	(0.27)
Married$_t$	-0.14*	-0.14	-0.13	-0.13
	(0.09)	(0.09)	(0.09)	(0.09)
Over early retirement age$_t$	0.06		0.08**	
	(0.04)		(0.04)	
First stage				
Over regular retirement age$_t$	0.08***	0.08***	0.13***	0.13***
	(0.01)	(0.01)	(0.01)	(0.01)
Over early retirement age$_t$		0.05***		0.003
		(0.01)		(0.01)
Hausman test (p-value)	0.13	0.02	0.57	0.78
Hansen-J test (p-value)		0.12		0.03
F-statistic	38.66	23.37	94,92	48.11
n	9,121	9,121	8.395	8.395
N	18,242	18,242	16,790	16,790
	Model 5	Model 6	Model 7	Model 8
	Good health	Good health	Good health	Good health
Retired$_t$	0.05	0.14		
	(0.21)	(0.19)		
Time spent in retirement (log)$_t$			0.06	0.01
			(0.16)	(0.14)
Married$_t$	-0.09**	-0.09**	-0.08*	-0.08*
	(0.04)	(0.04)	(0.04)	(0.04)
Over early retirement age$_t$	0.02		0.02	
	(0.02)		(0.02)	
First stage				
Over regular retirement age$_t$	0.08***	0.08***	0.11***	0.13***
	(0.01)	(0.01)	(0.01)	(0.01)
Over early retirement age$_t$		0.05***		0.003
		(0.01)		(0.01)
Hausman test (p-value)	0.63	0.32	0.86	0.87
Hansen-J test (p-value)		0.25		0.19
F-statistic	38.66	23.37	77.01	48.11
n	9,121	9,121	8.395	8.395
N	18,242	18,242	16,790	16,790

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses. All regressions include age, age squared and age cube. All models also include individual-fixed effects and time-fixed effects.

Table 16: Non-Lagged Effects with Strategy 1

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Self-assessed health	Euro-D	Health Stock	Self-assessed health	Euro-D	Health Stock
Retired$_t$	0.03	0.71	-0.19			
	(0.46)	(0.91)	(0.21)			
Time spent in retirement (log)$_t$				0.06	0.58	-0.14
				(0.39)	(0.76)	(0.17)
Hausman test (p-value)	0.51	0.29	0.72	0.67	0.41	0.66
Hansen-J test (p-value)	0.02	0.81	0.04	0.02	0.77	0.04
F-statistic	15.86	14.71	14.10	27.10	25.99	25.04
n	3,817	3,804	3,736	3,817	3,804	3,736
N	6,737	6,671	6,454	6,737	6,671	6,454
	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
	Self-assessed health	Euro-D	Health Stock	Self-assessed health	Euro-D	Health Stock
Retired$_t$	-0.19	-0.77	-0.42***			
	(0.26)	(0.51)	(0.13)			
Time spent in retirement (log)$_t$				-0.08	-0.43*	-0.17***
				(0.13)	(0.24)	(0.05)
Hausman test (p-value)	0.95	0.27	0.02	0.85	0.16	0.01
Hansen-J test (p-value)	0.19	0.73	0.09	0.10	0.63	0.07
F-statistic	27.81	28.84	27.18	23.40	23.91	21.09
n/N	6,561	6,447	6,186	5,892	5,791	5,398

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses. All regressions include country-fixed effects as well as all other control variables included in Table 5 (although all lagged variables there are included here as non-lagged). Models 1-6 exclude all retirees who have been retired for three years or longer, while Models 7-12 include all retirees while controlling for the LDV.

Table 17: Gender-Specific Effects – Both Strategies

Strategy 1	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Self-assessed health	Euro-D	Health stock	Self-assessed health	Euro-D	Health stock
Retired$_{t-1}$	-0.54*	-1.69***	-0.40***			
	(0.28)	(0.56)	(0.12)			
Time spent in retirement (log)$_{t-1}$				-0.25**	-0.88***	-0.20***
				(0.12)	(0.24)	(0.05)
Interaction_female$_{t-1}$	-0.14	-0.46*	-0.17***	0.004	0.05	-0.01
	(0.12)	(0.27)	(0.06)	(0.04)	(0.09)	(0.02)
Hausman test (p-value)	0.18	0.00	0.00	0.12	0.00	0.00
Hansen-J test (p-value)	0.34	0.18	0.14	0.38	0.49	0.27
F-test (Retired)	34.72	35.79	35.07	14.76	15.44	14.55
F-test (Interaction)	289.15	284.31	271.20	185.59	182.67	175.42
N	6,159	6,115	5,901	5,716	5,676	5,481
Strategy 2	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
	Self-assessed health	Self-assessed health	Self-assessed health	Self-assessed health	Self-assessed health	Self-assessed health
Retired$_{t-1}$	-0.98***	-0.73**	-0.93***			
	(0.36)	(0.30)	(0.35)			
Time spent in retirement$_{t-1}$				-0.67***	-0.66***	-0.69**
				(0.23)	(0.23)	(0.28)
Interaction_female$_{t-1}$	-0.06	-0.09	-0.40	0.08	0.03	0.01
	(0.45)	(0.35)	(0.42)	(0.13)	(0.12)	(0.13)
Married$_{t-1}$	-0.24*	-0.22*	-0.18	-0.16	-0.16	-0.09
	(0.13)	(0.12)	(0.13)	(0.12)	(0.12)	(0.13)
Female*Married$_{t-1}$	0.46***	0.44***	0.48***	0.36**	0.36**	0.34**
	(0.16)	(0.16)	(0.17)	(0.15)	(0.15)	(0.17)
Hausman test (p-value)	0.01	0.04	0.01	0.02	0.02	0.07
Hansen-J test (p-value)		0.23	0.71		0.71	0.85
F-test (Retired)	23.25	14.75	14.53	62.93	31.72	27.67
F-test (Interaction)	64.45	39.12	39.37	1007.23	538.97	539.22
n	9,121	9,121	9,122	8,395	8,395	8,400
N	18,242	18,242	18,244	16,790	16,790	16,800

Note: Significance levels: *p<0.10; **p<0.05; ***p<0.01. Cluster-robust standard errors in parentheses. Models 1-6 include the same controls as in Table 4, and Models 6-12 include the same variables as in Table 10. The first stages in Models 7 and 10 only include the instrument constructed from the regular retirement age. Models 7 and 10 do not include the LDV.