WORKING PAPERS IN ECONOMICS

No. 2/17

MAJA WEEMES GRØTTING AND
OTTO SEVALDSON LILLEBØ

HEALTH EFFECTS OF RETIREMENT. EVIDENCE FROM NORWEGIAN SURVEY AND REGISTER DATA

Department of Economics
UNIVERSITY OF BERGEN

Health Effects of Retirement

Evidence from Norwegian Survey and Register Data*

Maja Weemes Grøtting[†‡]   Otto Sevaldson Lillebø[§]

April 7, 2017

Abstract: We investigate the relationship between retirement and health, using comprehensive measures from survey- and register data with detailed information on socioeconomic status. Using regression discontinuity at the statutory retirement age of 67, we study health in terms of depression, physical health and acute hospital admissions, with the latter observed for the entire Norwegian population. Our findings suggest that retirement leads to better physical health outcomes for individuals with low socioeconomic status, both for subjective- and objective outcomes, and to increased symptoms of depression for individuals with high socioeconomic status. Our findings highlight the importance of heterogeneity in the health effects.

Keywords: Retirement; Health; Socioeconomic Status; Gender; Regression Discontinuity Design
JEL codes: I12; I14; I18; I38; J26

*We would like to thank Arild Aakvik, Ragnhild Balsvik, Dorly Deeg, Elisabeth Fevang, Astrid Grasdal, Thomas Hansen, Christian Monden, Matthew Neidell, Viggo Nordvik, Miikka Rokkinen, Ingrid Hoem Sjursen, Eirik Strømland, Kristin Årland and seminar participants at Columbia University, Norwegian Social Research, 23rd Nordic Congress of Gerontology, Norwegian School of Economics, University of Bergen and The Frisch Centre for very usefull comments. We would also like to thank Statistics Norway for providing access to the different data sources. We are grateful for financial support from the Norwegian Research Council.

†NOVA Norwegian Social Research, Oslo, Norway

‡Corresponding Author: Oslo and Akershus University College of Applied Sciences, PO box 4 St. Olavs plass, N-0130 Oslo, Norway. E-mail adress at: Maja.Grotting@nova.hioa.no (M.Grøtting)

§Department of Economics, University of Bergen

1 Introduction

Life expectancy steadily increases and the number of individuals retired as a share of the total population rises in most OECD countries. At the same time, concerns about the fiscal sustainability of the public pension system has led policy makers in the same countries to introduce policies aimed at prolonging individuals' working lives. An important question that seems to be overlooked in these policy debates is the importance of health with regard to retirement. Findings in the empirical literature regarding the health effect of retirement are mixed; individuals in strenuous jobs might have a beneficial change in health after retirement, whereas retirement may be harmful for individuals with a strong attachment to their working environment (see, e.g. Insler (2014) and references therein). Importantly, a few recent studies show that endogeneity and heterogeneity play an important role in these conflicting results (see Eibich (2015); Coe and Zamarro (2011)).

In this paper, we build on recent literature and investigate the health effects of retirement across differences in socioeconomic status and gender. We exploit an institutional setting in Norway that causes a discontinuous change in the likelihood of retiring at the age of 67, making a regression discontinuity (RD) framework suitable for our setting. This allows for identification of the causal short term effects of retirement on health.

We contribute to the literature by introducing both subjective and objective measures of health. This is an important extension compared to previous literature for two reasons; first, survey data is important, as it provides important insight into how individuals experience their own health, which is important for their well-being. However, self-reported measures of health have been criticized for being contextual and can suffer from justification bias. Moreover, samples of older adults might be prone to selective attrition. These issues are absent in register data, as they are reported from administrative registers. In particular, administrative data covers the entire population and records certain health conditions as truly objective. One drawback is that these measures are often extreme outcomes of health, and hence unsuited for studying moderate health effects. Taken together, however, these two sources will help us gain more insight into the effect of retirement on several dimensions of health.

In this paper, we assess the health effects of retirement at age 67, which was the statutory retirement age in Norway in 2007. Although most individuals chose to retire at the statutory retirement age, the entire public sector and half of the private sector have access to early retirement plans. Furthermore, a substantial fraction of the older working age population is on disability insurance (DI) which provides for a full or partial withdrawal from the labour market. DI is granted based on a physician's assessment that an individual has a permanently reduced capability of working. In this study, we define retirement as claiming retirement pension. Importantly, individuals on DI are transferred to retirement pension once they reach the statutory retirement age. We therefore assess the intention-to-treat effect of retirement on health.

From a theoretical perspective, we can expect the retirement effects to differ across socioeconomic status (SES) groups. According to Grossman's (1972) model, individuals with low education or low financial capital (low SES) will have to rely more heavily on their health as an input in the labour market, compared to individuals with high education or wealth, assuming that the different sources of capital are substitutes in the labour market. This is typically manifested through strenuous manual labour for the low SES groups. Moreover, individuals with higher education are assumed to be more efficient in promoting own health. In sum, the two mechanisms make it more costly for low SES groups to continue working. Retirement can therefore be seen as a mechanism that levels the increasing health inequalities between groups. Unless the difference in health deterioration is compensated for by investing in health-promoting activities for the low SES, the health stock at retirement will differ between these two groups.

Finally, institutional factors concurring with retirement can lead to reduced health inequalities. Examples are lower income dispersion among pensions compared to income from labour, or the provision of certain health-related benefits such as Medicare in the US. The theoretical and institutional mechanisms outlaid above are likely not instantaneous health consequences of retirement. However, the predictions from Grossman partly apply in this setting. The relief from strenuous manual work can be experienced as an instantaneous health improvement. For individuals working in a stressful environment, the relief can also be instantaneous. By contrast, retirement may lead to a reduced sense of purpose. This may have a negative effect on mental health. In order to tackle these theoretical issues, we systematically assess how the health effects differ both by socioeconomic status and gender, as well as by the three health outcomes: physical health, depression and acute hospitalization. We argue that this provides a more comprehensive picture of the different dimensions of health and retirement.

The health outcomes are taken from two different data sources. The first is a nationally representative sample of older adults with vast information on individual physical and mental health. The second data source contains individual level administrative records of acute hospital admissions for the entire Norwegian adult population. Both data sources contain exact birth month and retirement date from public registers. This rules out recollection bias, and together, the level of details allows for a more precise estimation of the retirement effects, as it allows for a more local estimation around the timing of retirement compared to analyses using data on the year level.[1]

A salient aspect of our study is that we assess the health effect of retirement at a higher age threshold than most studies in the field, which typically apply retirement age limits in the age range 60–65. As the retirement reforms that are being implemented are aimed at prolonging individual's working life, especially those who stay in the labour force until the statutory retirement age, it is crucial to assess the effects of retirement on health at these higher age thresholds. Our results show that socioeconomic status is important when studying the effect of retirement on health. The RD results indicate that retirement on average yields a positive effect of retirement on physical health for low SES groups. This effect is statistically significant and fairly sizeable when considering this subjective measure of health. We also find that retirement reduces the likelihood of being hospitalized, in general and for men with low education. Interestingly, we find that retirement increases the presence of depressive symptoms for individuals with high SES, but we find no effects on depression for the low socioeconomic status. In general, our study suggests that health effects on average can be canceled out by differences in socioeconomic status.

The paper proceeds as follows. In Section 2, we provide a discussion of earlier literature. We describe the institutional details for the Norwegian pension system and link the institutional setting to the empirical strategy in Section 3. In Section 4 we provide information about the data, outcome variables and some basic summary statistics. Our main results are presented in Section 5, and Section 6 concludes.

2 Earlier literature

Our paper is related to a large body of economic research about the effect of retirement on health. Given the important aspect of this issue and the vast amount of literature, there is a surprising lack of consensus across studies. One stand in the literature is related to the effect of retirement on cognitive functions. Exploiting variation in retirement ages between and within countries as instruments, both Mazzonna and Peracchi

[1]See e.g Lee and Card (2008) and Dong (2015) for a discussion on why age in years might yield inconsistent results unless properly accounted for

(2012) and Rohwedder and Willis (2010) found retirement to have a significantly negative impact on cognitive abilities. The rationale behind this finding is that retirement arguably removes the incentive to maintain cognitive functions required in the labour market.

A loss in cognitive abilities has further been linked with a loss of social interactions (see, e.g. Börsch-Supan and Schuth (2013); Mazzonna and Peracchi (2016)). The removal of work-related social interactions may have an impact on increased obesity (Godard (2016); Rohwedder and Willis (2010)), which in turn increases the risk of being diagnosed with a severe cardiovascular disease, reduced daily functioning and reduced self-rated health (Behncke (2012)). These findings also show that the outcomes in question are highly interrelated; as a reduced level of cognitive abilities might be triggered from a large set of unobserved chain reactions related to retirement.

From the US setting, Neuman (2008) provides evidence that retirement both preserves and improves health. Because retirement removes the time constraint induced by labour market participation, more time can be devoted to activities that both preserve and enhance individuals' health. This is in line with Grossman's (1972) model of health demand, where it can be shown that especially time-intensive workouts may be more attractive after retirement, when the opportunity cost of participating in such activities drops. Insler (2014) investigates several health-related outcomes and finds that retirees tend to reduce smoking and participate in health-enhancing activities. Bound and Waidmann (2007) find that retirement leads to a small, but significant positive effect on health for men. However, the same authors show that these results are highly sensitive to job characteristics and hinge on differences in socioeconomic status. As these differences plays an important role in determining the effect of retirement on health, there has recently been a growing interest in tackling these heterogeneity issues. Differences in socioeconomic status may be induced by income, education or endowments stemming from investments at an early stage of individual's working life.

To the best of our knowledge, only a small number of studies have investigated the presence of heterogeneity of retirement on health. Coe and Zamarro (2011) studied the extent to which retirement affects measures of self-reported health and depression across several European countries. They find that retirement reduces the likelihood of reporting bad health. Furthermore; the evidence suggests an improvement in a composite measure of self-reported health. In order to tackle the potential problem of effect-heterogeneity, they include controls for socioeconomic status, proxied by education and household income. Individuals with high education are less likely to report bad measures of self-reported health, given that they have retired, compared to individuals with no higher education.

Using a regression discontinuity framework, Eibich (2015) studied the effect of retirement on several subjective measures of health in Germany. The empirical evidence suggests the presence of effect-heterogeneity; whereas he uncovered no effect of retirement on health for individuals with a high education, individuals who retire from strenuous jobs seem to experience a large and positive change in physical health. We have yet to find a study that assesses both measures of subjective health and measures of objective health taken from register data.

Dong (2015) shows that using regression discontinuity design calls for careful consideration of the unit of measurement when age is the forcing variable. He shows that age in years, as opposed to age in months, might lead to inconsistent results. It is our impression, that in general the relevant regression discontinuity analyses of retirement and health do not take this into account.

3 Institutional Setting and Empirical Strategy

3.1 Institutional Settings in Norway

This section provides background information for the institutional setting in Norway in 2007/2008, which is important for our empirical analysis.[2] We start with a brief description of the pension system, as this is of the main focus in our study. An individual can start claiming retirement pension the month after reaching the statutory retirement age of 67, and is thus considered retired once this claim is made. The main provider of retirement pension is the mandatory public National Insurance System (NIS). This is a pay-as-you-go defined benefit system, and all individuals with a minimum number of years of residence are covered. Once retired, the pension consists of a mix between fixed earnings-independent basic pension and pension contributions based on previous labour market income.[3] Replacement rates from annual earnings have been found to be on average around 72% (Røed and Haugen (2003)).

Besides the statutory retirement age, there are two other commonly used exit routes from the labour market: disability insurance (DI) and the Early Retirement Program (ER). These are early exits routes that are only temporarily available until the statutory

[2] The system was reformed in 2011, but none of the new rules was in place throughout our study-period.

[3] For more details, see Brinch et al. (2017)

retirement age, after which they are all equally considered as retired. Eligibility for DI is based on health status and must be certified by a physician based on a permanent reduced ability to work. Despite being granted 100% DI, individuals are still able to earn income up to 1 basic amount (G).[4] Furthermore, DI can also be graded in a way that allows individuals to combine work and DI. This combination cannot exceed the income prior to the date in which they were granted DI.

ER grants the possibility of retiring in the age span 62–66. The entire public sector and nearly 50 percent of the private sector had access to such a scheme in 2007 and 2008. The scheme is equal regardless of whether an individual works in the private or public sector, but average replacement rates differ between the private and public sectors due to differences in earning levels (see Bratberg et al. (2004)) for a detailed description). Uptake of ER does not lead to curtailment of future pensions. In contrast to both old age pension and DI, ER is conditional on full labour market withdrawal.

Table 1 summarizes the labour market status for individuals aged 56–79 in 2007. This table shows the fraction of individuals who are either working, on ER, DI or retired. The shares do not summarize to one, as the same individual can be in two states, e.g. working part time and on graded DI.

Table 1. Labor market participation for individuals aged 56-79 in 2007.

Age Group	Working		Retired		ER		DI	
	Men	Women	Men	Women	Men	Women	Men	Women
56 - 61	79%	72%	-	-	-	-	19%	28%
62 - 66	59%	49%	-	-	16%	13%	31%	41%
67 - 69	17%	9%	89%	92%	-	-	-	-
70-79	18%	2%	98%	98%	-	-	-	-

Notes: This table is calculated based on register data from Statistics Norway covering the entire population of Norway (See chapter 4 for a description). We define any positive income as work as it depicts some relations to the labor market. This table will not sum to unity due to individuals being in other states, such as unemployed, recently emigrated to Norway or on other social security schemes.

Table 1 shows two important preconditions for our empirical analysis: labour market participation rate remains relatively high for older workers, and most individuals have started claiming old age pensions as soon as they reach the age of 67. While the statutory retirement age by no means forces individuals to retire, there was for most of the workforce no economic incentive to prolong working life once eligible for old age pension; and most firms have contracted that individuals have to retire once they

[4] In 2007, 1G represented around 11.100 US dollars using the average exchange rate NOK/\$ = 6.

reach the age of 67. Moreover, the norm was that people retired once they reached the statutory retirement age.

The lack of economic incentive to remain in work is due to how the supplementary pension was calculated, which was based on the number of years with positive pension points and the average number of pension points over the best 20 years. Since the number of years with positive pension points was limited to 40 years, the supplementary pension for individuals who already had reached this threshold was highly unlikely to change with one year extra of work. Therefore, only individuals with fewer than 40 years of positive pension points had any incentive not to retire once they have reached the statutory retirement age. Furthermore, a full earnings test was in place for individuals aged between 67 and 69 for earnings above 2G, resulting in a 40% reduction of the old age pension for each dollar earned. Taken together, all aforementioned factors provides strong incentives for individuals to retire once they reach the statutory retirement age.

3.2 Empirical Design

3.2.1 General idea

We investigate the impact of retirement on several dimensions of health. A cause of concern when estimating how retirement affects such measures is endogeneity. The relationship between health and retirement can be described by this linear equation:

$$Y_i = \beta_0 + \beta_1 R_i + \beta_2 X_i + \varepsilon_i \tag{1}$$

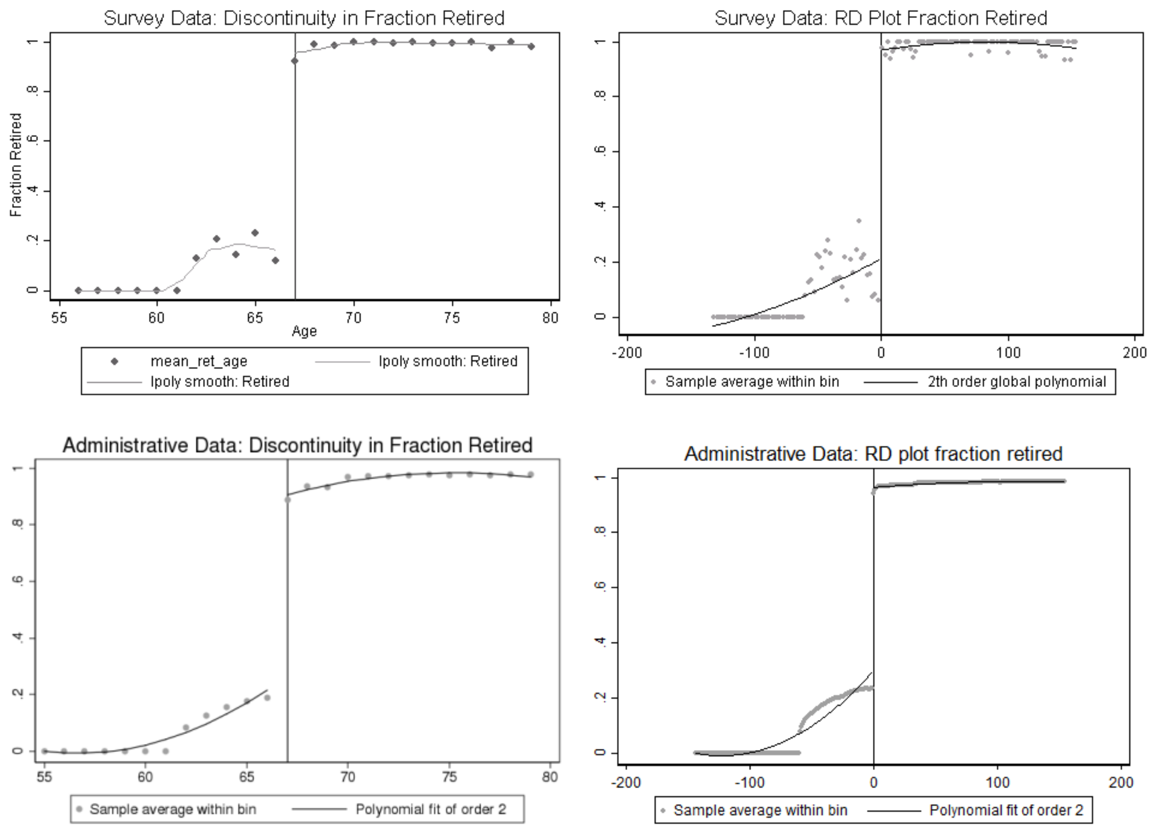
where Y_i represents health, R_i is a dummy variable equal to one if the individual has retired, and where X_i is a vector of relevant covariates. If retirement could be considered as a random event, equation (1) would provide us with an unbiased estimate of the short-term effect of retirement on health. However, people typically decide when to retire, and unobservable factors such as knowledge about own longevity or other factors that correlate with both health and the retirement decision remain unaccounted for, hence causing omitted variable bias. Importantly, own health is likely to affect retirement, causing bias in β_1 due to reverse causation. In order to circumvent these endogeneity issues of retirement in the OLS specification, we apply regression discontinuity design (RD).

Regression Discontinuity Design (RD) exploits institutional settings that determine access to a treatment. The general idea is that treatment (retirement) is determined by a running variable, in this case age, reaching a known threshold. Units above the threshold receive the treatment and units below the threshold do not receive the treatment, i.e. we use age as an allocation mechanism that determines retirement rather than using actual retirement behavior.

As laid out in the previous section, the institutional setting in Norway[5] provided incentives for people to retire as soon as they reach the eligible retirement age at 67. In Figure 1, we show the share of retired individuals from age 55 until age 79 to motivate the use of RD. The two upper panels are constructed from the survey data, whereas the two lower panes are constructed using the register data. We show the retirement timing both by age in years and by age in months to show that people generally do retire exactly at the timing of eligibility.

[5]As it was before the retirement reform

Figure 1. Fraction retired by age in years and distribution of age



Notes: The figure shows fraction retired by age in years and distribution of age from the two datasets. The two top figures stems from the survey data whereas the two bottom graphs stems from our administrative data. Points represent means among people in each month-year of birth cell.

In all of the four figures, the patterns are very similar.[6] The likelihood of retirement increases at age 62 when the possibility for early retirement becomes available to most workers. Then, there is a substantial jump in the fraction retired at age 67, after which only a negligible share of individuals remains in the labour force. The graphical evidence thus shows a clear response in terms of retirement at the statutory retirement age. We build our empirical analysis on this discontinuity in the probability of retirement that is caused by crossing the age limit for retirement pension eligibility.

[6] In the graphs above, retirement refers only to those who have actually retired, either through the early pension program or at the retirement age of 67. This means that individuals on DI are not considered retired, i.e retirement before the age of 67 refers only to those who have actually retired through ER. If we take out all individuals that are currently on DI or who were on DI before they retired, the picture looks the same.

Formally, identification in the RD framework relies on discontinuities in the probability of treatment at a (known) threshold of the running variable, and on continuity in the outcome variable in the absence of treatment at the point of discontinuity. Applied to our setting, we need the likelihood of retirement to change discontinuously at a known point in the age distribution, and that absent retirement, there would be no changes in health at this age threshold. Age determines treatment, but individuals cannot manipulate their age. Treatment is therefore randomized very close to the threshold as a direct consequence of the running variable.

The RD relies on local identification by comparing individuals right above and below the retirement age cut-off. The discontinuity gap in health at this point identifies the treatment effect. Since the probability of retirement is discontinuous at the cut-off of age 67, we can assume that reaching this age limit is what causes individuals to retire. Importantly, this assumption only holds for individuals close to the cut-off on the age distribution.

Despite having imperfect compliance, the intention to treat is as if randomized, which implies a causal interpretation of the estimated effects. The case of imperfect compliance is referred to as fuzzy RD, and implies that not all individuals chose to retire at the statutory retirement age of 67. There must still be a difference in the probability of treatment right above and right below the threshold. The jump in the outcome variable at the cut-off is divided by the fraction induced to take up treatment at the cut-off, which resembles a setting with instrumental variables. The effect identified by the fuzzy RD is the average effect of treatment for units at the threshold and who are compliers, meaning only people who take the treatment at the threshold. Applied to our setting, we need the likelihood of retirement to change discontinuously at a known point in the age distribution, and that absent retirement, there would be no changes in health at this age threshold. Age determines treatment, but individuals cannot manipulate their age. Treatment is therefore randomized very close to the threshold as a direct consequence of the assignment variable. The RD relies on local identification by comparing individuals right above and below the retirement age cut-off. The discontinuity gap in health at this point identifies the treatment effect.

The graphs in Figure 1 are indicative of a strong first stage in an instrumental variable approach, and motivate the use of fuzzy RD. As the probability does not jump from zero to one at the threshold, there is not perfect compliance. There is a small jump in the probability of retirement at age 62, the lowest age eligible for early retirement, and not all individuals chose to retire at 67. Importantly, there is still a strong first stage with the probability shifting from 0.2 to 0.92 at the threshold in both data sets.

Despite the fact that age determines eligibility and that it cannot be manipulated, the design may be invalid if individuals just above the threshold are more likely to answer a survey than those just below the threshold. This would clearly violate the RD assumption that the running variable is continuous at the threshold. We assess whether this is the case in the robustness section with a graph that displays age-in-months-distribution. There is no apparent discontinuity at the threshold. Importantly, there is universal access to health care from cradle to grave, i.e. there are no discontinuities in access to health care at any points along the age distribution.

3.2.2 Estimation and validity of the method

The raw statistics in the previous section show that the fuzzy RD framework is well suited to answering our research question. As described by Imbens and Lemieux (2008), the fuzzy RD resembles a setting with instrument variables in which the retirement coefficient can be consistently estimated using two stage least squares (2SLS). As noted by Hahn et al. (2001), the treatment effect will therefore be interpreted as a local average treatment effect (LATE), i.e. the estimated treatment effect of retirement on health for individuals induced by the age threshold to retire. Formally, we instrument for retirement using age equal to or above the retirement threshold at 805 months.

The first stage is given by

$$retirement_i = \gamma_0 + \gamma_1 1[age_i \geq c] + \gamma_2 age_i^B + \gamma_3 age_i^A + u_i \tag{2}$$

where the endogenous regressor $retirement_i$ is a binary variable equal to one if the individual is retired. $1[\bullet]$ is an indicator function taking the value one if the condition inside the brackets is true, and zero otherwise. c represents the retirement eligibility threshold at 805 months. age is measured in months, and we include continuous age controls. These are allowed to have different slopes at either side of the threshold. Superscript B refers to below cut-off and superscript A refers to above the threshold.

The first stage in this 2SLS set-up is simply actual retirement predicted by age exceeding the threshold, controlling for the general effect of age on health.

We apply retirement as predicted in the first stage, and the second stage is given by:

$$health_i = \beta_0 + \tau \widehat{retirement} + \beta_1 age_i^B + \beta_2 age_i^A + e_i \tag{3}$$

Here, $health_i$ represents the different health measures, all discussed in the next section.

In all of our estimates, we cluster at the age group level accounting for the imperfect fit of regression function, as we are using observations away from the discontinuity point (see Lee and Card (2008)).

As discussed earlier, the RD relies on local identification by comparing health outcomes for individuals right above c to individuals right below c. Estimation of the main results are done locally around the threshold. Local estimation implies that we have to choose the area around the threshold which we use in our regressions, i.e. what bandwidth to choose. We use the approach proposed by Imbens and Kalyanaraman (2011) for choosing a bandwidth. This is designed to minimize the mean squared error, and provides a trade-off between bias and variance. Based on the bandwidth selector, we chose a bandwidth of 10 months.[7] This means that in the estimations we will use the age range 795 months to 815 months, i.e. 10 months before and after the retirement age threshold. In the sensitivity analysis, we assess different bandwidths to check the sensitivity of the results with respect to different bandwidths.

For the 2SLS to provide consistent estimates of the retirement effect on health, we need to assume monotonicity and that the exclusion restriction holds, besides random assignment of the treatment and the existence of a first stage as outlined above. For monotonicity to hold, we must assume that reaching the statutory retirement age does not make people less likely to retire. We believe that it is highly unlikely that becoming eligible for retirement pension makes someone less likely to retire. The exclusion restriction implies that the only effect on health from crossing the age threshold for the statutory retirement age, is through becoming retired. This ensures that there are no events that exactly concur with the timing of crossing the age threshold that also affect health, thereby contaminating the estimated effects of retirement on health. We have no reason to believe that there should be any changes to circumstances that concur perfectly with aging one month from age 804 to 805 months. As mentioned above, access to health care remains unchanged throughout the age distribution. We test for discontinuities in other outcomes in the robustness section.

3.2.3 Robustness Checks

Vital to any RD analysis are thorough robustness checks. We follow Imbens and

[7] The optimal bandwidth suggested by Imbens and Kalyanaraman (2011) varies by SES-group. The suggested bandwidth is in the range 8-12 months for all the groups. For simplicity, we apply a bandwidth of 10 months in all estimations. Choosing different bandwidths within this interval has no influence on the estimated effects. See the robustness checks in the appendix for more on sensitivity of bandwidths.

Lemieux (2008) for robustness checks in RD. These tests include looking for discontinuities in the value of covariates that are not affected by the treatment at the cut-off test for discontinuities in the conditional density in the forcing variable to avoid self-selection or sorting into treatment or control; checking for discontinuities in average outcomes at other values of the forcing variable; and applying different values of the bandwidth.

The results from the robustness checks are shown in the appendix. Moreover, as will be discussed in more detail in the data section, we include a sensitivity analysis where we exclude individuals that are disabled from the analysis. As shown in Figure 1 above, there is also a discontinuity at age 62, when the early retirement eligibility comes into play. We will not use this cut-off in the analysis because poor health is an important predictor of choosing early retirement. It has also been shown that the introduction of early retirement reduced the incidence of disability insurance (see Bratberg et al. (2004)).

4 Data

We use data from two separate sources in our analysis. The first is a survey carried out on a representative sample of Norwegian adults collected in 2007 and 2008. This is the second wave of the Norwegian Study on Life-Course, Aging and Generation (NorLAG) panel study.[8] NorLAG contains individual data on a wide range of both physical and mental health outcomes, as well as information about socioeconomic status. Data collection was carried out by Statistics Norway with computer-assisted telephone interviews (CATI), and all respondents are merged with administrative registers from Statistics Norway for the time period 2002–2012. The registers contain information on year and month of birth and the exact month of retirement. Furthermore, these registers contain various sociodemographic background information such as labour income, social insurance payouts (such as DI) and educational attainment. We are thus able to construct detailed information for each individual regarding attachment to the labour market, retirement status and social security take-up, enabling identification of exact timing of retirement and whether the individual retired directly from the labour force or transitioned from disability insurance or other welfare programs.

Currently, the panel consists of two waves. The first was collected in 2002. We use the second wave for most all specifications in this analysis, as this contains a larger

[8]See Slagsvold et al. (2012) for a thorough description

sample than the first wave.[9] However, for some specifications in the sensitivity analysis we rely on data from the first wave to obtain information about past labour market performance. In these sensitivity analyses, we thus use the balanced panel which consists of a total of 3,774 individuals.

To measure health, we apply one physical and one mental health indicator. Physical health is measured by the Short Form 12 (SF12) scale (Ware et al. (1998)). Self-rated health (SRH) is one of the components that go into this composite measure of health. Other factors are measures of the degree to which an individual is able to perform tasks like vacuuming, moving a table or climbing stairs, whether there are certain tasks that he could not perform, or whether he had pain that limited his daily activities. The score is standardized on a scale from 0–100 with a mean of 50 and standard deviation of 10 using the US population as a reference. SF12 have been found to be a strong predictor of hospitalization, job loss due to health, future use of medical health services and depression (see e.g. Ware et al. (1998); Brazier and Roberts (2004) or Jenkinson and Layte (1997)).

Mental health is measured by the 20-item Center for Epidemiological Studies Depression (CES-D) scale (Radloff (1977)). Respondents were asked to indicate on a 4-point scale (1 = rarely or none of the time, 4 = all of the time) how often they felt sad, depressed, "that my sleep was restless", "that my life has been a failure," etc., during the previous week. The scores are added to comprise a 0–60 scale of depressive tendencies, and a score of 16 or higher indicates the presence of significant depressive symptoms (ibid.). It was designed to identify depression among the general population and is currently the most widely used instrument to measure depressive symptoms and to estimate prevalence rates in population surveys (see, e.g. Shafer (2006)).

Occupational status in the NorLag/LOGG data is coded by the ISCO-88 scale. This has been re-coded into two occupational groups, manual and professional workers. Following the classical division into blue and white collar workers of higher and lower skills respectively,[10] *Manual Workers* contains the three categories high and low skilled blue collar workers and low skilled white collar workers. *Professionals* are defined as high skilled white collar workers. This division is made because the first three are more similar based on observable characteristics.

In the second part of the empirical analysis, we use data that cover the entire

[9] The first wave contains 5,559 observations (response rate 67%), whereas the second wave contains 15,149 (response rate 60 %)

[10] http://www.eurofound.europa.eu/surveys/ewcs/2005/classification

population. Our outcome of interest is a binary indicator of whether someone suffered an unscheduled and acute hospitalization. Information on hospitalization comes from the National Patient Register (NPR). The hospitalization is acute in the sense that treatment is deemed necessary and thus not possible to postpone. Importantly, this measure of health is not correlated with the time cost to consult medical expertise. After retirement, individuals have more leisure time and the opportunity cost of seeking medical help is thus reduced, compared to individuals still in the workforce. It is therefore likely that the prevalence of diagnoses or medical treatments that are not acute increase after retirement, when the time cost of seeing a physician has fallen.

As with the NorLAG-data, we also have detailed information for each individual on social security take-up, date of retirement, month and year of birth as well as educational attainment and labour income.

4.1 Sample Restrictions and Individuals Outside the Labor Force

We restrict our attention to individuals aged 56–79 in 2007 and 2008 in both data sets. In the register sample, we use 2008 only, as this is as far back as the NPR data span. This leaves 4,619 individuals in the NorLAG sample and 892,908 individuals in the register sample. The register data is a panel of monthly hospitalizations on the individual level. Here we follow Imbens and Lemieux (2008) and treat the data as repeated cross sections and pool all months together, treating each observation as an individual. This also makes the register data more comparable to the NorLAG data. We account for within person correlation by clustering the errors at the individual level in the estimations.

For the main analysis, we leave all respondents in the sample. This includes individuals in the working age population that for some reason are outside the labour force. This ensures that the intention to treat is maintained in our design. One potential issue with this approach is that individuals who receive disability insurance are also retired at 67. We need to exclude the possibility that these individuals are driving the results. One concern is that a mechanism exists whereby disabled individuals, consciously or subconsciously, justify their status as disabled while they are of working age, and therefore under-report their health. As soon as they are eligible for retirement pension, they no longer need to justify their health status for receiving disability pension. If this is the case, one can expect to see an improvement in health for this group. We conduct a sensitivity analysis where we exclude all individuals that were not working until retirement age from the analyses.

In these sensitivity analyses, we run the whole analysis including only those that were gainfully employed or working until retirement. Ideally, we want to compare individuals working up to retirement age to individuals who retired from working (at the threshold). In the NorLAG data, this is done by adjusting the sample by two rules. The first implies including only individuals who had income from labour in 2006 in the analysis, the second by including only individuals who have stated that they are working or were working before they became retired in the analysis. Some caveats are worth mentioning; the first rule results in a substantial reduction in the sample size, as we need to use the balanced panel from both waves of the NorLAG study to identify the labour income in 2006. In identifying the sub-samples for the sensitivity analyses, it is crucial that we apply exactly the same selection rule on either side of the threshold, i.e. we cannot use different "rules" on either side of the cut-off to define those currently in the labour market and those who were in the labour market prior to retirement. A potential concern with the second rule is that the formulation of the question to the working and retired part of the population differs slightly in the NorLAG data. In the register data, we define individuals as working if they currently have or had positive income before retirement. The results from the sensitivity analysis are shown in the Appendix.

4.2 Descriptive statistics

Table 2 displays summary statistics for the sample from both the NorLAG data and the register data. These are men and women aged 56 - 79 in 2007 and 2008 respectively.

Table 2. Descriptive statistics

	Entire Sample		Below Threshold		Above Threshold	
Source: NorLAG	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Age	65.34	[6.58]	66.15	[0.36]	67.00	[0.00]
Female	0.49	[0.50]	0.47	[0.50]	0.50	[0.51]
Living with partner	0.71	[0.45]	0.75	[0.43]	0.72	[0.45]
Less than high school degree	0.23	[0.42]	0.25	[0.44]	0.25	[0.43]
High school degree	0.51	[0.50]	0.45	[0.50]	0.51	[0.50]
Any college	0.27	[0.44]	0.30	[0.46]	0.25	[0.43]
PCS12	46.93	[45.73]	45.73	[12.03]	47.55	[10.12]
Depression	8.72	[7.53]	8.03	[7.15]	8.75	[7.72]
Professional	0.48	[0.50]	0.47	[0.50]	0.50	[0.50]
Manual	0.43	[0.50]	0.40	[0.49]	0.41	[0.49]
Retired	0.44	[0.50]	0.18	[0.39]	0.96	[0.18]
N	4619		190		200	
Source: Population level Data						
Age	64.92	[6.67]	66.19	[0.38]	67.00	[0.00]
Female	0.51	[0.50]	0.50	[0.50]	0.50	[0.50]
Married	0.63	[0.48]	0.64	[0.48]	0.64	[0.48]
Less than high school degree	0.31	[0.46]	0.32	[0.47]	0.34	[0.47]
High school degree	0.45	[0.50]	0.46	[0.50]	0.45	[0.50]
Any college	0.24	[0.42]	0.23	[0.42]	0.21	[0.41]
Acute Hospital Admission	0.01	[0.10]	0.01	[0.10]	0.01	[0.10]
Retired	0.40	[0.49]	0.29	[0.43]	0.95	[0.21]
N	1,071,068		31,751		33,752	

Notes: Standard deviations in square brackets. This table displays descriptive statistics from our two data sources (See start of section 4 for further details). The NorLag data consits of a representative sample from 2007/08 whereas the administrative data is measured in 2008. The descriptive statistics for our population level data are measured as a mean for the year of 2008. The dependent variables from the NorLAG survey are measures of depression and physical health (PCS12). The dependent variable from our population data is measured as acute and unscheduled hospitalization.

The first two columns are summary statistics for the whole sample whereas the next two columns show the summary statistics for those within the bandwidth of 10 months below and 10 months above the retirement threshold of 805 months. These are the observations within the bandwidth used for estimating the short term retirement effects in the regression analysis. It is important that the two groups are balanced with respects to the covariates. In this case, individuals on either side of the threshold are similar with respect to education,[11] living arrangements and occupation.

5 Results

We begin this section by discussing the first-stage results. The first stage is specified in equation (2), and we show the estimates of γ in Table 3.

Table 3. First-stage regressions

	Entire Sample	**Men**	**Women**
Source: NorLAG			
Retired	0.954***	0.941***	0.961***
	(0.0362)	(0.0587)	(0.0431)
N	371	190	181
Source: Register Data			
Retired	0.706***	0.665***	0.746***
	(0.00264)	(0.00390)	(0.0359)
N	840,239	416,611	423,628

Note: This table shows the first-stage regressions as specified in equation 2. Standard errors (in parentheses) are clustered at the age in month level for the NorLAG data and at the individual level for the register data. *=p<0.10, **=p<0.05, ***=p<0.01.

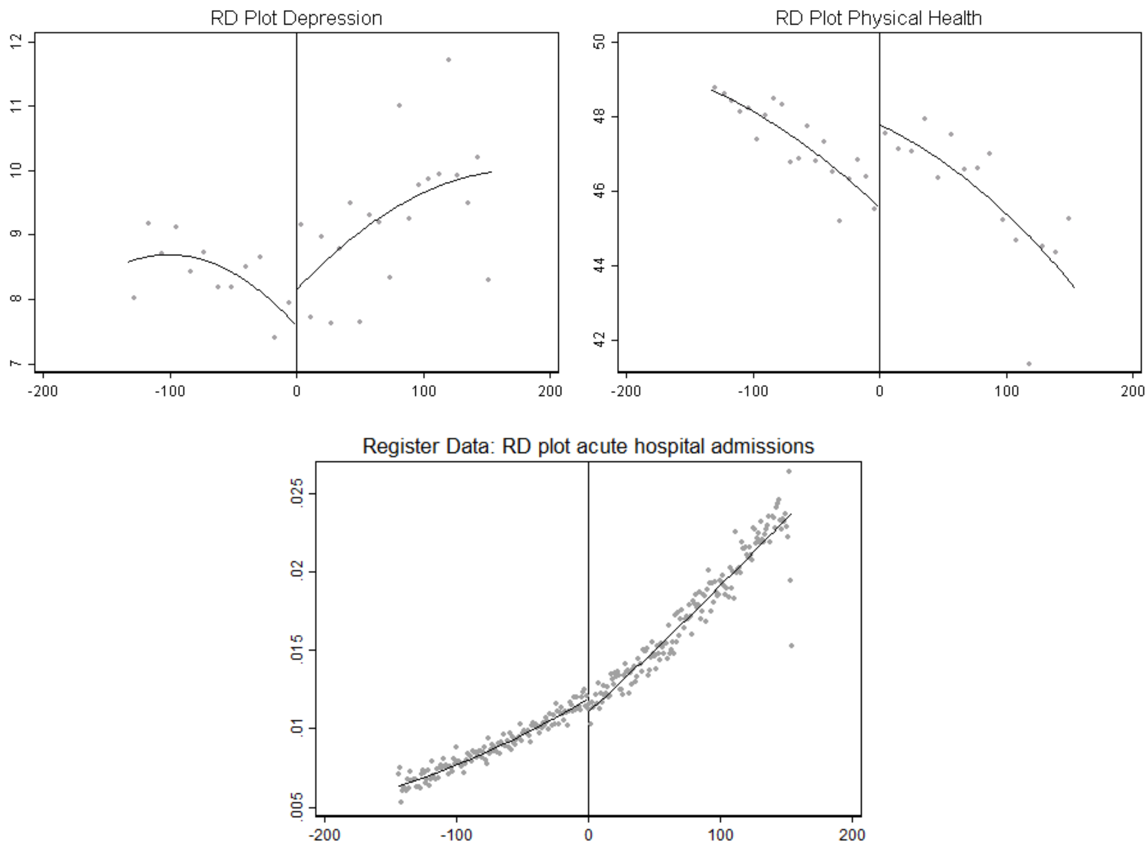
The results in Table 3 show that crossing the statutory retirement age significantly increases the probability of retirement, thus indicating a strong first stage. These results are in line with the graphical results that we presented in Figure 1.

5.1 Graphical Results

Figure 2 presents graphical evidence on the relationships between health and age for the three outcomes used in our study: physical health, depression and acute hospital admissions. The age range spans from 56 to 79 years, and all outcomes are displayed relative to the retirement age threshold at 805 months, normalized to zero. The lines are fitted on either side of the threshold using a second order polynomial global fit.

[11] T-tests confirm that individuals on either side of the threshold are statistically similar with respect to education, living arrangements and occupation

Figure 2. Discontinuity in health at threshold



Notes: The three figures presents the age-health relationship for depression, physical health and acute hospital admissions. The two top figures stems from the survey data, whereas the bottom graphs stems from our administrative data. Points represent local means for the health outcomes based on bins among people in each month-year of birth cell. Depressive symptons corresponds to higher values on the y-scale whereas the scale of acute hospital admissions corresponds to fraction of population within the bin. The x-axis displays months relative to the retirement age threshold of 805 months, which is equal to 67 years and one month of age.

The upper right graph in Figure 2 shows the observed health pattern for physical health for all individuals aged 56–79 in the NorLAG sample. Physical health declines as individuals age, but there is a substantial jump at the threshold. At this threshold, the trajectory shifts up to a level of someone 80 months younger, which amounts to 6.5 years. For depression, the evidence of a discontinuity is not as strong. Rather than observing a substantial jump, there seems to be more of a change in the direction of the trend. Higher values are associated with more depressive symptoms, so depressive symptoms are falling towards retirement, and increasing after retirement. Finally, acute hospital admissions are increasing across the age span 56–79, but we observe a slight downward shift in the upward trend at the threshold. This shift suggests that reaching the retirement threshold leads to reduced risk of acute hospital admissions.

As mentioned above, the effects estimated using RD are only identified close to the threshold. There is an ongoing debate as to whether it is the cumulative or contemporaneous effects of retirement that are the largest (see Coe and Zamarro (2011) and Mazzonna and Peracchi (2016)). By visual inspection of the graphs in Figure 2, there is suggestive evidence of a prolonged effect of retirement on both subjective and objective measures of health. For physical health, retirement shifts individuals to a higher health trajectory and they seem to stay on that higher health trajectory. We observe a modest reduction in acute hospitalizations, and retirees seem to stay on that lower trajectory. For depression, the trajectories before and after retirement are completely different. Retirement can be a factor that tilts the trajectory from downward to upward sloping.

5.2 Regression results

We present the 2SLS regression results for all three health dimensions in Tables 4, 5 and 6. The effects are estimated using a bandwidth of +/- 10 months around the threshold. We used the bandwidth selector suggested by Imbens and Kalyanaraman (2011).[12] We estimate the effects for each gender and for the different SES-groups separately.

In Table 7 we present results from a formal test of heterogeneous retirement effects in which the instrument is interacted with the different SES-groups.

5.2.1 Results, Physical health and Depression

Table 4 displays the results of the short-term retirement effects on physical health and depression. We find that retirement leads to a 5.7 point increase in physical health for the population as a whole. This is a substantial effect given that the mean and standard deviation for this health outcome is 47 and 10 points, respectively. We also find a strong and positive effect for men (8 points), and we find a positive (4 points), but not significant effect for women. Our findings are in line with evidence from Coe and Zamarro (2011) and Eibich (2015). In general, their findings suggests that retirement leads to an increase in physical health in both Germany and the USA. Although our estimates are short-term effects, previous findings suggest that retirement also has a cumulative effect on physical health through increased physical activity (Coe and

[12]This bandwith selector finds the optimal distance in age in months from the threshold by minimizing a trade-off between bias and variance.

Zamarro (2011); Eibich (2015); Kämpfen and Maurer (2016)). Moreover, Wolin et al. (2008) point towards "lack of time" as a common barrier to more physical activity. In line with empirical findings, the Grossman (1972) model shows that retirement may correspond to increased investments in physical activity, as retirees have a lower alternative cost of engaging in health-promoting activities.

Table 4. The Short Term Retirement Effects on Health

	Physical Health			Depression		
	All	**Men**	**Women**	**All**	**Men**	**Women**
Retired	5.689***	8.036**	4.053	3.437**	2.032	7.816**
	(1.979)	(3.026)	(3.465)	(1.707)	(2.517)	(3.750)
Observations	361	185	176	291	154	137

Note: This table displays the impact of retirement on measures of physical health and depression for the sample as a whole and for men and women separately. The scale for depression is reversed, meaning that a positive coefficient equals that people show more depressive symptoms by retiring. Standard errors (in parentheses) are clustered at the age in month level. *=p<0.10, **=p<0.05, ***=p<0.01.

Depression is measured by the sum score CES-D. This scale is reversed so that higher values imply more depressive symptoms. A positive coefficient implies increased depressive symptoms. To better fit the data, second order age polynomials for the age trends both below and above the threshold are included in equations 2 and 4.

We find that in general, people get more depressed when they retire. The effect size is 3.5 points, which corresponds to one-half of a standard deviation. For women, the effect is 7 points and significant at the 5 % level. For men, the effect is positive, but not significant. Our findings provide further empirical evidence to the literature on retirement and depression. Mazzonna and Peracchi (2016) argue that the shift in environment caused by retirement can be linked to depressive symptoms, but that the direction of depressive symptoms can go either way. Coe and Zamarro (2011) find no causal effects of retirement on depression, whereas Charles (2002) and Eibich (2015) find that retirement leads to better subjective well-being and mental health, respectively. Börsch-Supan and Schuth (2013) show that a loss of social interactions related to retirement may have a negative effect on both depressive symptoms and mental health.

Based on the discussion in the introduction, we should expect different health effects of retirement depending on education and occupation. Table 5shows the results by gender and SES.

Table 5. Short Term Retirement Effects on Health by SES and Gender

	Measure		Measure	
	Physical Health	**Depression**	**Physical Health**	**Depression**
Panel A	**Male Manual Workers**		**Female Manual Workers**	
Retired	10.93***	8.236	13.77*	0.664
	(2.067)	(5.224)	(7.329)	(1.408)
Observations	67	51	59	44
Panel B	**Male Professional Workers**		**Female Professional Workers**	
Retired	-0.595	9.509	1.062	9.408*
	(4.575)	(6.854)	(3.804)	(5.301)
Observations	61	54	62	52
Panel C	**Male Low Education**		**Female Low Education**	
Retired	9.314***	-1.170	7.122	5.329
	(2.467)	(4.562)	(4.817)	(3.326)
Observations	130	108	131	98
Panel D	**Male High Education**		**Female High Education**	
Retired	4.799	10.64***	-6.610	13.62*
	(5.925)	(2.649)	(6.846)	(7.175)
Observations	54	45	45	39

Note: This table displays the impact of retirement on physical health (SF-12) and depression (CES-D) estimated within groups of different SES and gender. The scale for depression is reversed, meaning that a positive coefficient equals that people are more depressed. Standard errors (in parentheses) are clustered at the age in month level. *=p<0.10, **=p<0.05, ***=p<0.01.

For men in the manual sector (Panel A) and men with low education (Panel C), we find large and positive effects of retirement on physical health. The effects of 9 to 11 points are over one standard deviation and significant at the 1 % level for both measures of SES. For men with higher SES (Panels B and D), we find small and insignificant effects, and for men in professional occupations (Panel B), the effect is even negative.

For female manual workers (Panel A), we find a large effect of almost 14 points (significant at the 10 % level) on physical health. This group contains only 59 individuals, which makes it difficult to precisely estimate the effect. For women in the low education group (Panel C) the group size is larger and the coefficient is positive, yet the estimate is insignificant. For women in the professional sector (Panel B), we find no effect on physical health, and this effect is actually zero for women with higher education (Panel D).

Our results are in line with the findings of Eibich (2015). He shows that highly educated individuals benefit less from retirement in terms of self-reported health, compared to individuals with low SES. Insler (2014)) suggests that wealthy people have more time to invest in their health while working. When we assess depressive symptoms, we find that men in the high education group (Panel D) and women with high SES (Panels B and D) get significantly more depressed by retiring. These effects are large and significant at the 1 % and 10 % level, respectively. Mazzonna and Peracchi (2012) and Charles (2002)) argue that highly educated individuals are more likely to have a job where retirement may have a negative effect on measures such as depressive symptoms.

Power calculations show that a sample of at least 90 is needed to ensure a power of 0.8. Hence, some of these groups could be too small to precisely estimate the effect. It could be argued that this should lead to the application of wider bandwidths. However, wider bandwidths also imply more bias (Lee and Lemieux (2010)). Moreover, as is the case for physical health, an upward shift in a downward sloping curve implies underestimating the shift as you start moving away from the threshold.

To sum up, the results are clear in that retirement leads to better physical health given the low SES for men, but for women the results are too noisy. We cannot say whether the coefficients are insignificant due to small sample sizes or that there are no effects for women. However, the pattern in health effects of retirement remains fairly consistent in that low SES groups have large, positive coefficients in the analysis of physical health for both men and women, whereas we find large, positive coefficients on depression for the high SES for both men and women. Based on this analysis, there do not seem to be substantial differences by gender, but this will be formally assessed in Section5.2.3.

5.2.2 Results Emergency Hospital Admissions

Table 6 displays the results on acute hospital admissions from the register data.

Table 6. Short term retirement effects on acute hospitalization

	Acute Admissions		
	All	Men	Women
Panel A		**Difference in Gender**	
Retired	-0.00128*	-0.00148	-0.00112
	(0.000687)	(0.00108)	(0.000874)
Observations	825,605	407,386	418,219
Panel B		**Low Education**	
Retired	-	-0.000941	-0.00164*
		(0.00126)	(0.00944)
Observations		302,376	341,065
Panel C		**High Education**	
Retired	-	-0.00342	0.0196
		(0.00208)	(0.00233)
Observations		105,010	77,154

Note: This table displays the impact of retirement on acute unscheduled hospitalizations. We include all acute hospitalizations resulting in an inpatient stay. Standard errors (in parantheses) are clustered at the individual level. *=p<0.10, **=p<0.05, ***=p<0.01.

Panel A explores how retirement affects acute unscheduled hospitalizations for the population on average and by gender. In general, retirement leads to a small, but significant reduction in the incidence of acute hospitalizations. The effect is rather small with a change in incidence equal to -0.00128 and only significant at the 10 % level. It is worth noting that the overall incidence of acute hospital admission in the population is low. The effect size of 0.00128 amounts to about 10 % of the total incidences in the population. When we estimate the effect by gender, we still find negative effects, yet these are insignificant.

When we assess the effect by gender and SES, we find that retirement leads to reduced acute hospitalizations for women with low education (Panel B). The effect is slightly higher than in the population and significant at the 10 % level. Regardless of the educational level (Panels B and C), we find negative, but insignificant effects of retirement for men. For women with high education (Panel C) the effects are in fact positive, but also insignificant. In the sensitivity analysis, we show the retirement effect when we assess only individuals who are working and who were working until retirement.

To some extent, our results contradict findings from Behncke (2012). He shows that retirement increases the risk of being diagnosed with a chronic condition. However, the opportunity cost of seeking medical help is greatly reduced after retirement, and it is likely that individuals with such conditions seek medical help and hence are screened, thereby increasing the likelihood of diagnosing such conditions.

We also assessed the effect of retirement on mortality, but the results were marginal and insignificant for any of the groups in question. This is largely in line with evidence from Hernaes et al. (2013). They find that a reduction in the retirement age in Norway had no effect on mortality.

5.2.3 A Formal Test of Heterogeneity

Table 7 presents the results from the formal test of the heterogeneity. These are the results of a reduced form of equation (3), where the instrument is interacted with SES groups and gender. We estimate the following:

$$health_i = \beta_0 + \gamma 1[age_i \geq c] * SES_i + \beta_1 1[age_i \geq c] + \beta_2 age_i^B + \beta_3 age_i^A + e_i \qquad (4)$$

where γ is the coefficient of interest and $1[age_i \geq c]$ is the instrument indicating whether age in months is equal to or exceeds the threshold. We apply the same +/- 10 months bandwidth in these regressions.

Table 7. Formal test of differences by SES

	Measure		
	PCS12	**Depression**	**Acute Admissions**
Panel A		**Low education**	
Retired	4.975	2.809	-0.000608
	(3.12)	(1.94)	(0.000618)
Observations	361	291	840,239
Panel B		**Gender**	
Retired	3.696	-1.556	-0.000468
	(2.55)	(2.08)	(0.000576)
Observations	361	291	840,239
Panel C		**Manual Workers**	
Retired	6.858*	3.160	-
	(3.30)	(1.94)	-
Observations	249	201	-

Note: This table displays the interactions between retirement and SES groups. Standard errors are clustered at the age in months-group. *=p<0.10, **=p<0.05, ***=p<0.01.

We see that the retirement effects are statistically different from each other when SES is measured by occupation. Although the estimated effects differ quite substantially by educational group, the differences are not statistically significant when SES is proxied by education. Moreover, there are no statistically significant differences in the retirement effect by gender.

6 Conclusion

In this paper, we study the health effect of retirement using the statutory retirement age at 67 in a regression discontinuity design. We apply a combination of survey and register data – with measures of acute hospitalization, symptoms of depression and physical health – which allows a more comprehensive approach to study health effects of retirement.

We find that on average retirement has a positive effect on physical health and a negative effect on mental health. When we assess the effects by different SES groups, we find that retirement has a positive health effect for individuals with low socioeconomic status. This holds for both physical health as measured by the subjective SF-12, and the objective acute hospital admissions. We find no significant effects on these health outcomes for the high SES groups. For depression, the pattern is somewhat reversed; we find no effect for the low SES group, but retirement leads to increased depressive symptoms for the high SES groups.

Robustness checks confirm that there are no discontinuities at the threshold in the distribution of the forcing variable, or in covariates that should not be affected by retirement. Additionally, we do not find discontinuities in the dependent variable at other points along the age distribution. Finally, the results are robust to different bandwidths.

The results for depression should be handled with some care. The graphic evidence did not show a clear discontinuity at the cut-off, and the regression results are somewhat sensitive to the polynomial specification applied. For acute hospital admissions, we found some inconsistencies in that there is a small significant effect at the lower placebo level. In general, our findings on physical health are in line with the previous literature. However, this is not the case for depressive symptoms. Furthermore, the retirement age applied in earlier assessments, ranging from late 50s to about 65, are lower than in our study. We use a retirement age of 67. The higher age threshold applied in this study is valuable for policymakers. Current retirement reforms typically aim at increasing working lives, thereby mainly affecting workers who stay in the labour force until these higher retirement ages.

There is a paucity of analyses that apply objective health outcomes in the literature. Importantly, this study confirms that it is not just the experience of health that improves, but the actual risk of hospital admissions that is reduced by retirement. We can thus confirm that it is not only that individuals feel better about their health after retirement, they are in fact healthier in terms of reduced likelihood of hospital admissions.

This study accentuates the importance of assessing heterogeneous effects for individuals in different circumstances. Our findings indicate that the retirement reforms aimed at prolonging working life can be socially distortive due to the differential effects based on SES. Retirement at 67 leads to reduced likelihood of emergency hospitalization and increased physical health for low SES groups, while the higher SES groups might benefit from working longer, as they experience more depressive symptoms due to retirement at 67.

References

Behncke, Stefanie, "Does retirement trigger ill health?," *Health economics*, 2012, *21* (3), 282–300.

Börsch-Supan, Axel and Morten Schuth, "29 Early retirement, mental health and social networks," *Active ageing and solidarity between generations in Europe: First results from SHARE after the economic crisis*, 2013, p. 337.

Bound, John and Timothy Waidmann, "Estimating the health effects of retirement," 2007.

Bratberg, Espen, Tor Helge Holmås, and Øystein Thøgersen, "Assessing the effects of an early retirement program," *Journal of Population Economics*, 2004, *17* (3), 387–408.

Brazier, John E and Jennifer Roberts, "The estimation of a preference-based measure of health from the SF-12," *Medical care*, 2004, *42* (9), 851–859.

Brinch, Christian N., Erik Hernaes, and Zhiyang Jia, "Salience and Social Security Benefits," *Journal of Labor Economics*, 2017, *35* (1), 265–297.

Charles, Kerwin Kofi, "Is Retirement Depressing?: Labor Force Inactivity and Psychological Well-Being in Later Life," Technical Report, National Bureau of Economic Research 2002.

Coe, Norma B and Gema Zamarro, "Retirement effects on health in Europe," *Journal of health economics*, 2011, *30* (1), 77–86.

Dong, Yingying, "Regression discontinuity applications with rounding errors in the running variable," *Journal of Applied Econometrics*, 2015, *30* (3), 422–446.

Eibich, Peter, "Understanding the effect of retirement on health: mechanisms and heterogeneity," *Journal of health economics*, 2015, *43*, 1–12.

Godard, Mathilde, "Gaining weight through retirement? Results from the SHARE survey," *Journal of health economics*, 2016, *45*, 27–46.

Grossman, Michael, "On the concept of health capital and the demand for health," *Journal of Political economy*, 1972, *80* (2), 223–255.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, "Identification and estimation of treatment effects with a regression-discontinuity design," *Econometrica*, 2001, *69* (1), 201–209.

Hernaes, Erik, Simen Markussen, John Piggott, and Ola L Vestad, "Does retirement age impact mortality?," *Journal of health economics*, 2013, *32* (3), 586–598.

Imbens, Guido and Karthik Kalyanaraman, "Optimal bandwidth choice for the regression discontinuity estimator," *The Review of economic studies*, 2011, p. rdr043.

Imbens, Guido W and Thomas Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of econometrics*, 2008, *142* (2), 615–635.

Insler, Michael, "The health consequences of retirement," *Journal of Human Resources*, 2014, *49* (1), 195–233.

Jenkinson, Crispin and Richard Layte, "Development and testing of the UK SF-12," *Journal of Health Services Research*, 1997, *2* (1), 14–18.

Kämpfen, Fabrice and Jürgen Maurer, "Time to burn (calories)? The impact of retirement on physical activity among mature Americans," *Journal of health economics*, 2016, *45*, 91–102.

Lee, David S and David Card, "Regression discontinuity inference with specification error," *Journal of Econometrics*, 2008, *142* (2), 655–674.

_ **and Thomas Lemieux**, "Regression discontinuity designs in economics," *Journal of economic literature*, 2010, *48* (2), 281–355.

Mazzonna, Fabrizio and Franco Peracchi, "Ageing, cognitive abilities and retirement," *European Economic Review*, 2012, *56* (4), 691–710.

_ **and** _ , "Unhealthy Retirement?," *Journal of Human Resources*, 2016.

Neuman, Kevin, "Quit your job and get healthier? The effect of retirement on health," *Journal of Labor Research*, 2008, *29* (2), 177–201.

Radloff, Lenore Sawyer, "The CES-D scale a self-report depression scale for research in the general population," *Applied psychological measurement*, 1977, *1* (3), 385–401.

Røed, Knut and Fredrik Haugen, "Early Retirement and Economic Incentives: Evidence from a Quasi-natural Experiment," *Labour*, 2003, *17* (2), 203–228.

Rohwedder, Susann and Robert J Willis, "Mental retirement," *The Journal of Economic Perspectives*, 2010, *24* (1), 119–138.

Shafer, Alan B, "Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung," *Journal of clinical psychology*, 2006, *62* (1), 123–146.

Slagsvold, Britt, Marijke Veenstra, Svein Olav Daatland, Gunhild Hagestad, Thomas Hansen, Katharina Herlofson, Kristine Koløen, and Per Erik Solem, "Life-course, ageing and generations in Norway: the NorLAG study," *Norsk Epidemiologi*, 2012, *22* (2).

Ware, John E, Susan D Keller, and Mark Kosinski, *Sf-12: How to Score the Sf-12 Physcial and Mental Health Summary Scales*, QualityMetric Incorporated, 1998.

Wolin, Kathleen Y, Gary G Bennett, Lorna H McNeill, Glorian Sorensen, and Karen M Emmons, "Low discretionary time as a barrier to physical activity and intervention uptake," *American journal of health behavior*, 2008, *32* (6), 563–569.

7 Appendix

7.1 Disabled Individuals - Past Labor Income and Self-Reported Work Status

As stated above, individuals on disability insurance are mechanically transferred from disability pension to retirement pension at age 805 months. We need to make sure that the positive effect we found on the SF-12 is not driven by these individuals. Initially there is no reason to believe that there should be an effect for these individuals as they were not working before retirement, i.e. they should have no change in circumstances. However, as the health measure contains elements of self-assessed health, one could imagine that someone who is disabled need to justify their status as disabled, consciously or subconsciously. In this case health prior to the statutory retirement age would be underreported. When they are no longer in a situation where poor health is defining their labor market status, they feel healthier, or no longer have the need to report poor health. If this scenario is plausible, we need to rule out that the results found in Section 5 are driven by this group.

Table 8 displays the results from two sub-samples, each aimed at eliminating the disabled from the analysis. The non-disabled subsamples are defined in Section 3. Finding coefficients of the same sign and magnitude, albeit not statically significant for the second rule, can ensure us that these effects are not driven by the disability justification hypothesis.

In Table 9 we assess this with the population level data, the results are amplified in both magnitude and statistical precision. The results are still insignificant for all groups except men in the low SES group, thereby confirming that individuals outside the labor force are not driving the results. Rather, for this outcome, as this is not subject to the potential justification bias, we should expect that individuals who retire formally at 67, but without any actual change in circumstances, should water down the effects we see in the population in general.

7.2 Robustness Checks in the Regression Discontinuity Design

Below we assess the sensitivity of the results for different bandwidth selections; we check for discontinuities in the forcing variable, age, at the cutoff; we test for discontinuities in other outcomes that should not have been effected by the threshold; and we check for discontinuities in the outcome of interest at points in the age distribution where there should not be any discontinuities. This follows the suggestions in Imbens and Lemieux

(2008) closely.

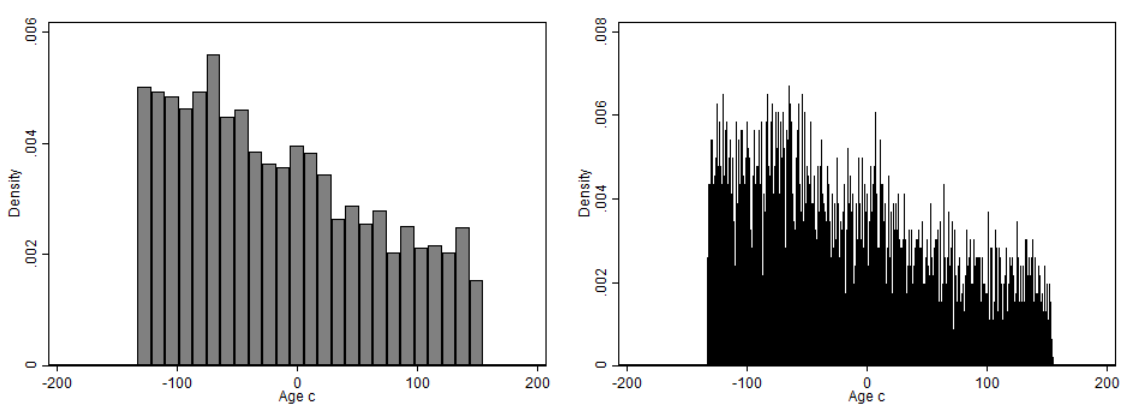
7.2.1 Bandwidth Selection

The worry in an RD application is that using a bandwidth that is too wide, allows for other things than the intervention of interest to drive differences in outcomes for those right above compared to those right below the threshold. In Table 8 we display the results for physical health and depression, and for the groups as a whole and for men and women separately using a bandwidth of 7 months and 15 months. Using a bandwidth of 7 months does not alter the results, increasing the to 15 months bandwidth somewhat reduce the effects

7.2.2 Continuity of the Forcing Variable

Vital to any RD application is that individuals are unable of manipulating the forcing variable. In this case, the forcing variable is age, which individuals cannot change in any way. It could however be the case that retired individuals are more likely to respond to the survey due to the reduced opportunity cost of time. Figure 3 shows two histograms of age in months to look for bunching at the threshold.

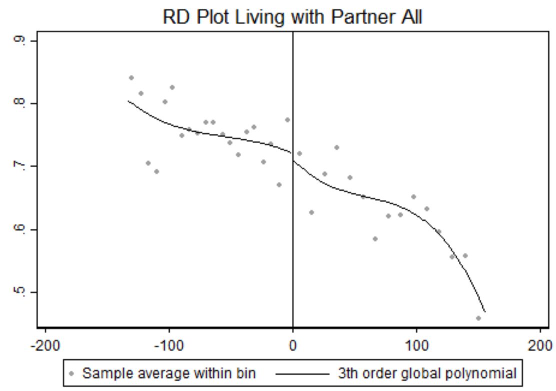
Figure 3. Discontinuity of the Forcing Variable



There is no evidence of any discontinuity in the forcing variable at the threshold.[13] For the population level data, this holds by construction, as all individuals are represented in the data, and people cannot manipulate their age.

[13] We also did a more formal test proposed by Cattaneo, Jansson, and Ma (2015a). a local polynomial density estimator for testing the null of continuous density of the forcing variable at the threshold. The p-value under this test is 0.3251.

Figure 4. Discontinuity in other variables



7.2.3 Placebo Tests

In Table 8 we also test for jumps in the two health outcomes at points in the age distribution where there should be no discontinuity. A common practice is to conduct placebo tests at the median of the two sub-samples below and above the actual cutoff. In this case, the median age at the distribution below the threshold is at age 62. This cannot be used for this purpose, as some individuals can retire at this age. We therefore use age 59 for the lower placebo. For the upper placebo, we use the cutoff median which is age 72. No jumps or significant effects was found at these placebo thresholds. We apply the same placebo thresholds for acute hospital admissions, and the results are displayed in Table 9. Here we find no effects at the higher placebo, but we do find some effects at the lower.

7.2.4 Discontinuity in Other Outcomes

To ensure that it is retirement that is driving the effect and not something else, we look for discontinuities in an outcome that should not be affected by retirement, at least not in the short term. Figure 4 checks for discontinuities for the likelihood of living with a partner.

This graph does not indicate that are any discontinuities in the likelihood of living with a partner at the threshold. this variable at the threshold. Moreover, the 2SLS estimation of equation 2 and 3 confirm this. For the register data, we see no significant difference in being married in either side of the threshold.

Table 8. Robustness Checks for the NorLAG-data.

	All PCS12	**Men** PCS12	**Women** PCS12	**All** Depression	**Men** Depression	**Women** Depression
Conditional on income	16.42***	15.83	-1.553	-14.733***	-20.37**	-30.29
	(2.966)	(10.88)	(7.264)	(3.419)	(8.364)	(43.69)
Observations	92	53	39	76	46	30
Conditional on working	6.274***	9.741***	2.523	1.686	2.698	3.821
	(2.089)	(3.758)	(7.312)	(1.649)	(2.672)	(3.608)
Observations	247	142	105	203	117	86
Bandwidth 7	9.472***	14.69***	4.497	1.047	.215	2.544
	(2.019)	(5.206)	(4.394)	(1.373)	2.436)	(1.841)
Observations	275	142	133	221	118	103
Bandwidth 15	5.801***	9.391***	2.623	2.865**	3.451	3.546
	(2.130)	(3.109)	(4.245)	(1.455)	(2.697)	(3.201)
Observations	92	53	39	76	46	30
Placebo at 59	-1.441	.971	-5.752	-4.442	-4.4381	-5.315
	(3.665)	(4.220)	(6.628)	(3.109)	(2.689)	(4.556)
Observations	454	242	212	351	177	174
Placebo at 72	-1.111	-1.264	.628	2.190	2.253	2.754
	(1.685)	(4.786)	(2.213)	(3.571)	(5.302)	(3.689)
Observations	251	127	124	186	97	89
Living with a partner	-.0411					
	(.124)					
	371					

Note: This table displays different robustness checks on the impact of retirement on physical health and depression. The scale for depression is reversed, meaning that a positive coefficient equals that people show more depressive symptoms by retiring. Standard errors (in parentheses) are clustered at the age in month level. *=p<0.10, **=p<0.05, ***=p<0.01.

Table 9. Robustness Checks for the Population level data.

	All	**Men**	**Women**	**Men**		**Women**	
				Low Education	High Education	Low Education	High Education
Conditional on working	-.00103	-.000883	-.00123	-.000677	-.00492**	-.00118	-.00135
	(.000877)	(.00129)	(.00114)	(.00152)	(.00243)	(.00126)	(.00273)
Observations	362,857	203,212	159,645	138,119	65,093	121,308	38.337
Bandwidth 7	-.00128*	-.00148	-.00112	-.000941	-.00342	-.00164*	.00196
	(.000687)	(.00108)	(.000874)	(.00126)	(.00208)	(.000944)	(.00233)
Observations	825,605	407,386	418,219	302,376	105,010	341,065	77,154
Bandwidth 15	-000996*	-.00135	-.000705	-.00150	-.00105	-.00125	-.00231
	(.00599)	(.000937)	(.000767)	(.00110)	(.00174)	(.000830)	(.00200)
Observations	1,241,687	612,603	629,084	453,761	158,842	511,517	117,567
Placebo at 59	.000193	.0000006	.000395	.000565	-.00126***	.000333	.000556
	(.000309)	(.000601)	(.000471)	(.000764)	(.000403)	(.000530)	(.000666)
Observations	1,269,311	649,906	619,405	449,981	199,925	451,397	168,008
Placebo at 72	-.000980*	-.000609	-.000130*	-.000169	-.00243*	-.00158*	.000742
	(.000401)	(.000912)	(.000732)	(.000906)	(.000132)	(.000878)	(.000147)
Observations	609,661	282,319	327,342	225,474	56,845	286,818	40,524

Note: This table displays different robustness checks on the impact of retirement on acute hospital admissions. Standard errors (in parentheses) are clustered at the age in month level. *=p<0.10, **=p<0.05, ***=p<0.01.

Institutt for økonomi
Universitetet i Bergen
Postboks 7800
5020 Bergen
Besøksadresse: Fosswinckels gate 14
Telefon: +47 5558 9200
Fax: +47 5558 9210
www.uib.no/econ/