Journal of Population Economics
https://doi.org/10.1007/s00148-019-00742-9

ORIGINAL PAPER

Health effects of retirement: evidence from survey and register data

Maja Weemes Grøtting[1] · Otto Sevaldson Lillebø[2]

Received: 9 September 2018 / Accepted: 5 July 2019 / Published online: 26 July 2019

Abstract

Using a local randomized experiment that arises from the statutory retirement age in Norway, we study the effect of retirement on health across gender and socioeconomic status. We apply data from administrative registers covering the entire population and from survey data of a random sample to investigate the effects of retirement on acute hospital admissions, mortality, and a composite physical health score. Our results show that retirement has a positive effect on physical health, especially for individuals with low socioeconomic status. We find no effects of retirement on acute hospitalizations or mortality in general. However, our results suggest that retirement leads to reduced likelihood of hospitalizations for individuals with low socioeconomic status. Finally, we show that the positive health effects are driven by reduced pain and reduced health limitations in conducting daily activities. Our findings highlight heterogeneity in the health effects across socioeconomic status and across subjective and objective measures of health.

Keywords Retirement · Health · Socioeconomic status · Gender · Regression discontinuity design

JEL Classification H75 · I14 · I18 · J26

1 Introduction

With increasing life expectancy, the number of retired individuals as a share of the total population is rising in most OECD countries. This has led to concerns about

Responsible editor: Alessandro Cigno

✉ Maja Weemes Grøtting
Maja.Grotting@oslomet.no

1 Norwegian Social Research, Oslo Metropolitan University, Stensberggata 26, 0170 Oslo, Norway

2 Department of Economics, University of Bergen, Fosswinckelsgate 14, 5007 Bergen, Norway

the fiscal sustainability of public pension systems and policy initiatives that aim to prolong working lives and increase retirement age. An important issue that seems to be overlooked in policy debates over these reforms is the impact that prolonged working lives has on health, and especially if there are heterogeneous retirement effects by socioeconomic status (SES).

Findings in the empirical literature regarding the health effects of retirement are mixed. Some studies report positive effects (Coe and Zamarro 2011; Eibich 2015), whereas others report negative effects (Behncke 2012; Mazzonna and Peracchi 2017) or no effects (Hernæs et al. 2013; Heller-Sahlgren 2017). Although some studies highlight the importance of SES in these health effects (Coe and Zamarro 2011; Eibich 2015), there is limited evidence from formal tests to suggest that the effects differ by SES. Another limitation in the literature is that most studies assess retirement effects in the early 60s, an age threshold that is substantially lower than proposed policies to postpone retirement toward age 70. Finally, most studies rely on survey data or administrative records (of sub-samples of the population), which often imply subjectivity in the health outcomes or small sample issues.

In this paper, we investigate the health effects of retirement across SES and gender in Norway by applying both survey and administrative data, where the latter cover the entire population. We assess the health effects of retirement at age 67, which was the statutory retirement age in 2007. This is a higher age threshold than what has previously been studied. To control for individuals self-selecting into retirement, we exploit that the statutory retirement age caused a discontinuous change in the likelihood of retiring at the exact timing of eligibility. This implies local randomization around the retirement eligibility age threshold and makes a regression discontinuity (RD) framework suitable. We compare the health outcomes for those right above the statutory retirement age threshold (i.e., the treatment group) to those right below (i.e., the control group). This allows for the identification of the causal short-term effects of retirement on health.

Most studies in this field rely on survey data with the well-known limitations related to non-response and recall bias. Furthermore, while measures of subjective health provide important insights into how individuals experience and rate their own health, such measures have been criticized for being contextual and can suffer from justification bias (see, e.g., McGarry (2004) for a thorough discussion). Another possible concern is that survey data of older adults are especially prone to health-related selection, as non-response or attrition is correlated with poor health.

The Norwegian administrative data are attractive with respect to overcoming some of these concerns. In particular, administrative data cover the entire population and record certain health conditions as truly objective. Still, measures of health from public registers are often extreme outcomes, such as mortality and acute hospital admissions, and hence unsuited for studying moderate health effects. In addition to records of mortality and acute hospital admissions from public registers, we therefore include a composite measure of self-assessed health from a representative sample of Norwegian older adults (The NorLAG Panel Survey (Slagsvold et al. 2012)). This measure is the short form-12 (SF-12) health survey (Ware et al. 1996). We assess both the overall physical score and the specific components that go into the SF-12.

We believe that our health measures, collectively, will provide important insight into the multidimensional effects of retirement on health. Moreover, both data sources (the administrative data and the NorLAG data) contain an exact birth month and retirement date from public registers, ruling out recollection bias. Finally, having monthly records allows for a more precise estimation of the effects of retirement on health, as it enables a more local estimation around the timing of retirement compared with analyses using data reported annually.[1]

As SES can be an important factor in analyses of retirement and health, we systematically assess how the health effects differ by common proxies for SES, such as education and occupation. A person's education or type of occupation is important in the analysis of health effects of retirement because they to a large degree determine the kind of work situation an individual retires from. Higher education and white collar jobs are often less physically demanding and associated with greater autonomy and control over the work situation, compared with manual jobs (Case and Deaton 2005; Mazzonna and Peracchi 2012). Moreover, Case and Deaton (2005) document that manual labor jobs, associated with low education and low income, are more "wear and tear" types of jobs, in which health deteriorates at a more rapid pace than in non-manual professions.

According to the Grossman (1972) model of health demand, individuals with low education or low financial capital (low SES) will have to rely more heavily on their health as an input in the labor market, compared with individuals with higher SES, as the different sources of capital are substitutes in the labor market. This is typically manifested through strenuous manual labor for the low SES groups. Moreover, individuals with higher education are assumed to be more efficient in promoting their 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 health inequalities between SES groups.

The RD application in this study identifies the short-term health effects of retirement. On the one hand, we can expect to see short-term effects on health as the relief from strenuous physical work or the relief from working in a stressful environment is an instantaneous change of circumstance. On the other hand, retirement may lead to a reduced sense of purpose before new routines are developed (Rohwedder and Willis 2010).

Our results show that retirement yields a sizeable and positive effect on physical health. This effect is especially strong for the low SES group, at about one standard deviation, whereas we find no effects for the high SES group. We find no effects on mortality or acute hospitalizations in general. However, for the low SES group, we find that retirement leads to a reduction in the likelihood of acute hospitalizations. Our results show that SES is important when studying the effect of retirement on health, but we find no gender differences. Moreover, we find that the reason why retirement leads to better physical health is due to reduced pain and a lower likelihood of reporting that physical health is a limitation in completing both "daily tasks" and

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

"specific tasks profoundly." The results for physical health and mortality are robust to a wide range of robustness and specification checks, whereas the checks regarding the results for hospital admissions are less robust and must therefore be interpreted with some caution.

A limitation of this study is that we only identify the short-term health effects of retirement. This is a consequence of the RD approach, in which using observations further away from the retirement threshold increases bias in the estimates. The choice of bandwidth is vital in an RD design and implies a trade-off between bias and precision, where moving away from the threshold implies including more data, thereby increasing precision, while also increasing bias (Lee and Lemieux 2010). In addition, the short-term focus can make detecting effects on more severe health outcomes, such as acute hospital admissions or mortality, difficult as the effect on these outcomes is less likely to be immediate. Another limitation is that the health effect estimated in this study is the average effect of retirement on health for those induced to retire by reaching the retirement threshold, i.e., the local average treatment effects (LATE). This implies that although the estimates in this study have strong internal validity, it is difficult to be explicit about external validity.

The paper proceeds as follows. Section 2 provides a review of previous research and describes the institutional structure of the Norwegian pension system. Section 3 describes our empirical strategy. In Section 4, we present the data, outcome variables, and some basic summary statistics. Our main results are presented in Section 5, and Section 6 concludes.

2 Earlier literature and institutional setting

2.1 Earlier literature

Our paper is related to a growing body of economic research about the effect of retirement on health. Given the important aspect of this issue and the vast amount of literature on the topic, there is a surprising lack of consensus across studies. One reason for this is that a large fraction of the existing evidence reports correlations rather than well-identified causal effects. Lately, there has been an increasing amount of well-identified studies, most of which use exogenous variation in retirement eligibility as sources of identification. As the majority of these studies use survey data or administrative records for sub-samples of the population, we contribute to the literature by providing objective health outcomes for the entire Norwegian population.

The strand of literature most relevant to this paper includes quasi-experimental evidence from analyses that focus on subjective and objective measures of general health, and is mainly from a European or US context. One of the most cited studies is Coe and Zamarro (2011). They study the extent to which retirement affects measures of self-reported health and a composite health index across several European countries using the Survey of Health, Ageing and Retirement in Europe (SHARE) data. They find that retirement reduces the likelihood of reporting bad self-rated health and leads to an improvement in a composite measure of subjective health.

From the US setting, Neuman (2008) uses age-specific retirement incentives as instruments for retirement. Applying data from the Health and Retirement Study (HRS), he provides evidence of retirement being both preserving and improving for self-rated health. He argues that since retirement removes the time constraint induced by labor market participation, more time can be devoted to activities that both preserve and enhance individuals' health. This is in line with Grossman's 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) uses data from HRS and applies workers' self- reported probabilities of working past ages 62 and 65 as instruments. He finds that retirees experience positive effects on a health index, which consists of both objective and subjective measures of health. Moreover, he finds that retirees tend to reduce smoking and to participate more in health-enhancing activities.

However, not all studies have shown retirement to have such a positive impact. Using data from the English Longitudinal Study of Aging (ELSA), Behncke (2012) reports that retirement actually increases the risk of being diagnosed with a cardiovascular disease[2] and cancer. Also contradictory to the findings of the aforementioned studies, she finds that retirement increases the probability of reporting poor health and the risk of being diagnosed with a chronic condition.

Another study using the ELSA is by Bound and Waidmann (2007). They find that retirement leads to a small but significant positive effect on physical health for men, where physical health entails self-assessed health, physical functioning, and biomarkers. Moreover, they show that these results are highly sensitive to job characteristics and differences in SES. As these differences arguably play an important role in determining the effect of retirement on health, there has recently been a growing interest in tackling these heterogeneity issues. To the best of our knowledge, only a small number of studies have investigated the presence of heterogeneity across SES or gender in the health effects of retirement.

Mazzonna and Peracchi (2017) stress the importance of heterogeneity in the health effects and argue that the previous literature failed to detect the potential heterogeneity. Using the SHARE data, they find that for people working in more physically demanding jobs, retirement has an immediate beneficial effect on both a health index of self-reported measures and cognition. For the rest of the workforce, however, retirement has negative long-term effects on health and cognition.

In the paper closest to our study, Eibich (2015) applies a RD framework to study the effect of retirement on several self-reported measures of health in Germany. The empirical evidence suggests effects that are heterogeneous by SES. Whereas he uncovered no effect of retirement on health for individuals with higher education, individuals who retire from strenuous jobs seem to experience a large and positive change in physical health.

Our study differs from the study by Eibich (2015) in three substantial ways. First, we use age measured in months, instead of in years, as our assignment variable. This

[2]Retirement is also found to have an impact on increased obesity (Godard 2016; Rohwedder and Willis 2010).

implies less bias in the estimates as this allows only using observations that are very close to the retirement threshold (Dong 2015). As mentioned in the introduction, using observations further away from the threshold introduces more bias. Second, we use data comprising the full population, which limits the issues related to sample selection in surveys of older adults. Third, we apply truly objective measures of health, such as acute hospitalizations and mortality, which are not self-reported.

From the Norwegian setting, Hernæs et al. (2013) employ a stepwise introduction of early retirement ages in Norway during the 1990s as instruments to assess whether retirement age matters for mortality.[3] They find no relationship between lowering early retirement age and mortality up through age 77 and question whether retirement has a causal impact on mortality.

Based on the relevant literature, it is unclear to what extent and in what direction retirement affects health. Previous findings are characterized by differences in methodology, be it an instrumental variable approach, RD approach, or difference-in-differences approach. Another aspect of the literature is the different health outcomes. While self-rated physical health is often found to be positively associated with retirement, several studies document that retirement is related to a decline in mental health and cognitive abilities. Although studies within the quasi-experimental literature on health effects of retirement are characterized by strong internal validity, the external validity can be poor or at least difficult to be precise about. Thus, the added-value of these studies seen as a whole is much greater than the sum of the individual contributions (Angrist and Pischke 2010).

2.2 Institutional setting in Norway

This section provides background information on the institutional setting in Norway in 2007/2008.[4] We start with a brief description of the pension system, as this is the focus of our study. An individual can start claiming retirement pension the first month after reaching the statutory retirement age of 67 and is, in our analysis, 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 labor market income. Replacement rates from annual earnings have been found to be around 72% on average (Røed and Haugen 2003).

In theory, the statutory retirement age did not force individuals to retire. However, most companies had contracted retirement upon reaching the statutory retirement, and the norm was that people retired once they hit this age threshold. Moreover, for most of the workforce, there was little economic incentive to prolong working

[3]Early retirement in Norway was introduced at age 65, but later reduced in a stepwise matter to age 62. The authors exploit this stepwise reduction as a source of exogenous variation.

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

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

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

The numbers are based on own calculations using the administrative data which covers the entire population of Norway (see Section 4 for a description). Work is defined as having earnings larger than zero. The states will not sum to unity because individuals can be in two states at the same time, e.g., by combining work and partial uptake of DI

life once eligible for the old age pension. There was a full earnings test in place for individuals aged between 67 and 69 for earnings above 2 basic amounts,[5] resulting in a 40% reduction of the old age pension for each dollar earned.[6]

Besides the statutory retirement age, there are two other commonly used exit routes from the labor market: disability insurance (DI) and the early retirement program. These are early exit routes that are temporarily available until the statutory retirement age. Eligibility for DI is based on health status and must be certified by a physician based on a permanently reduced ability to work. DI can also be graded in a way that allows individuals to combine work and DI. ER was available for all public and about half of the private sector workers from age 62.[7] At 67, recipients of DI and ER are automatically transferred to retirement pension.

Table 1 summarizes the labor market status for individuals aged 56-79 in 2007. This table shows the fraction of individuals who are either working, on ER and DI, or claiming retirement pension. The shares do not summarize to unity because it is possible for the same individual to be in two states, e.g., by combining partial uptake of DI and working.

Table 1 shows two important preconditions for our empirical analysis: the labor market participation rate remains relatively high for older workers in Norway and most individuals start claiming retirement pensions as soon as they reach age 67. Provided the strong link between retirement pension uptake and exit from employment, we argue that claiming retirement pension in practice means withdrawing from the labor market. Strictly speaking, in this analysis, we are estimating the intention-to-treat (ITT) effects of offering retirement pension at age 67. Because uptake of pension in practice means withdrawal from the workforce for the majority of the population, we assume that the health effects to a large degree will stem from the relief from work related tasks. We refer to claiming retirement pension as retirement in the remainder of this article.

[5] One basic amount is the lowest earnings required to accrue pension points.

[6] This was lifted in 2008 for 67-year olds.

[7] See Hernæs et al. (2013) or Kudrna (2017) for more details about the ER system.

3 Empirical strategy

3.1 Regression discontinuity design

We investigate the impact of retirement along several dimensions of health. Ideally, we seek to investigate the following linear relationship between health and retirement:

$$\text{Health}_i = \beta_0 + \beta_1 \text{ Retirement}_i + X_i^{'}\beta_2 + \varepsilon_i, \tag{1}$$

where Retirement$_i$ is a dummy variable equal to one if the individual has retired and zero otherwise and X_i is a vector of relevant covariates. If retirement were to be considered a random event, or if we could include all variables that correlate with retirement and health in X_i, Eq. (1) would provide us with an unbiased estimate of the effect of retirement on health. However, people typically decide themselves when to retire, and unobservable factors such as knowledge about own longevity or other unobservable factors that correlate with both health and the retirement decision remain unaccounted for in Eq. (1). This causes omitted variable bias in β_1. In addition, own health is likely to affect retirement, causing bias in β_1 due to reverse causation. In order to circumvent these issues in the OLS specification, we apply the RD design.

RD exploits institutional settings that determine access to a treatment. The idea is that treatment (retirement) is determined by a running variable (age), reaching a known threshold (the statutory retirement age). Units above the threshold receive the treatment, and units below the threshold do not receive the treatment. This means that we use age as an allocation mechanism that determines retirement, rather than using actual retirement behavior. The RD design relies on local identification by comparing individuals' right above and right below the retirement age cutoff. The discontinuity gap in health at this point identifies the treatment effect. Since the probability of retirement is discontinuous at the cutoff age 67, we assume that reaching this age limit causes individuals to retire. Importantly, this assumption only holds for individuals close to the cutoff on the age distribution.

As described in Section 2, the general rule was that individuals started claiming retirement pensions at the statutory retirement age of 67. However, about 16% of men and 13% of women within the eligible age groups chose to retire early through ER, and a small fraction retired later.[8] This is a setting of imperfect compliance. The fuzzy RD (FRD) design is therefore more appropriate. Unlike in the sharp RD, where all treated units are compliers, i.e., the likelihood of treatment goes from zero to one at the threshold, the FRD allows for a smaller discontinuity in the probability of retirement at the threshold.[9]

[8] DI provides another important channel out of the labor force. However, being granted DI is fundamentally different from retiring, as DI typically follows from a long period of sickness absence and must be certified by a physician and is granted through a thorough process.

[9] The difference between sharp and fuzzy RD is parallel to the difference between a randomized experiment with perfect compliance and a randomized experiment with imperfect compliance, when only the intention to treat is randomized.

3.2 Estimation

The FRD design resembles a setting with instrumental variables, with retirement coefficients consistently estimated by using two-stage least squares (2SLS) (Imbens and Lemieux 2008). The treatment effect is 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 (Hahn et al. 2001). In the setting of imperfect compliance with the treatment, the intention to treat (ITT) is as if randomized, which implies a causal interpretation of the estimated coefficients. The estimated effects are interpreted as the health effects of offering retirement pension at age 67.

Formally, we instrument for retirement using age equal to, or above the retirement threshold at 805 months, the month after which an individual turns 67 years of age. Specifically, we estimate the following two equations:

$$\text{Retirement}_i = \gamma_0 + \gamma_1 1[\text{Age}_i \geq c] + \gamma_2 \text{Age}_i^B + \gamma_3 \text{Age}_i^A + u_i, \tag{2}$$

where the endogenous regressor Retirement_i is a binary variable equal to one if the individual is retired, i.e., is claiming retirement pension. $1[\bullet]$ 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 (the first month after turning 67). 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 ages below the retirement threshold, and superscript A refers to ages above the threshold. The first stage in this 2SLS set-up is thus actual retirement predicted by age exceeding the threshold, controlled for the general effect of age on retirement.

The second stage is given by:

$$\text{Health}_i = \beta_0 + \tau \widehat{\text{Retirement}}_i + \beta_1 \text{Age}_i^B + \beta_2 \text{Age}_i^A + e_i, \tag{3}$$

where Health_i represents the different health measures for individual i. $\widehat{\text{Retirement}}_i$ is the predicted value of retirement from the first stage. Our parameter of interest is τ, and its estimate is the jump in the outcome variable at the threshold, divided by the fraction induced to take up the treatment at the threshold. This is the estimated treatment effect of retirement on health for individuals induced by the age threshold to retire.

As the health effects in the RD design are only identified close to the retirement threshold, the estimations are done locally around the threshold. We choose the optimal bandwidth, i.e., how many months on either side of the age cutoff to include in the estimation, in a cross-validation procedure suggested by Imbens and Kalyanaraman (2012).[10] This is designed to minimize the mean squared error and provides a trade-off between bias and variance. Based on this bandwidth selector, we choose a bandwidth of 10 months.[11] This means that only individuals in the age range

[10]Dong (2015) shows that using RD design calls for careful consideration of the unit of measurement when age is the forcing variable, as age in years, as opposed to age in months, might lead to inconsistent results.

[11]The optimal bandwidth suggested by Imbens and Kalyanaraman (2012) varies by SES group. The suggested bandwidth is in the range of 8–12 months for all the groups. For simplicity, we apply a bandwidth

795 months to 815 months (10 months before and 10 months after the retirement age threshold) are included in the estimations.[12] In the robustness analysis, we assess different bandwidths to check the sensitivity of the results with respect to the choice of bandwidth. In addition to assessing different bandwidths, we perform a range of robustness checks. Here, we follow the guide to practice by Imbens and Lemieux (2008) for robustness checks using the RD design. These results are presented in the Appendix, but we discuss them briefly in Section 5.

Finally, in the cross-sectional survey data, we follow Lee and Card (2008) and cluster at the age group level. As noted by Lee and Card (2008), for RD applications where the running variable is discrete, estimating a parametric function away from the discontinuity point can be seen as a form of random specification error. This implies a common component of variance for all the observations at any given value of the running variable. Thus, they suggest clustering at the age group level to account for this imperfect fit as clustering leads to wider confidence intervals. In the panel data from the administrative records, we cluster at the individual level to account for the within-person correlation in the error term. The structure of these data will be discussed in more detail in the next section.

4 Data and sample selection

4.1 Data

We use data from two separate sources in our analysis. The first is a survey carried out on a representative sample of Norwegian older adults, and the second is composed of administrative health and population registers covering the entire population. Unfortunately, individuals from the two sources cannot be connected as the first data source has been anonymized.

4.1.1 The NorLAG survey data

The first data source is a survey carried out on a representative sample of Norwegian older adults, the Norwegian Study on Life-Course, Aging and Generation (NorLAG) panel study.[13] The data were collected in 2002 and 2007. NorLAG contains individual data on a range of health outcomes, as well as information about SES. Data collection was carried out by Statistics Norway with computer-assisted telephone interviews (CATI).

of 10 months in all estimations. Choosing different bandwidths within this interval has little influence on the estimated effects. See the robustness checks in the Appendix for more on the sensitivity of the results with respect to the choice of bandwidth.

[12] Due to the small sample size left in the survey data when we apply the 10 months bandwidths, we also ran the entire analysis using a bandwidth of 20 months. This does not change the results from the survey data in any substantial way.

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

All respondents to the survey are merged with administrative registers for the period 2002-2012.[14] The registers contain information on year and month of birth and of retirement. Furthermore, the registers contain various sociodemographic background information such as labor income, social insurance take-up, and educational attainment. We are thus able to construct detailed information for each individual regarding attachment to the labor market, retirement status, and social security take-up, enabling identification of the exact timing of retirement, and whether the individual retired directly from the labor force or transitioned from disability insurance or other welfare programs.

Currently, the NorLAG panel consists of two waves. For the main analyses, we use the second wave as this contains a larger sample than the first wave.[15] However, for some specifications in the sensitivity analysis, we rely on data from the first wave to obtain information about past labor market performance. This is outlined in detail in Section 4.2.

Our health outcome from the NorLAG data is a composite measure of physical health, namely the physical component of the Short Form 12 (SF-12) scale (Ware et al. 1996). Self-rated health (SRH) is one of the components that go into the SF-12. 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 could not be performed due to health limitations, or whether pain limits daily activities.[16] The score is standardized on a scale from 0–100 with a mean of 50 and a standard deviation of 10 using the US population as a reference. SF-12 has been found to be a strong predictor of hospitalization, job loss due to health, future use of medical health services, and depression (Jenkinson and Layte 1997; Ware et al. 1996; Brazier and Roberts 2004).

Occupational status in the NorLAG data is coded in accordance with 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.[17] Professional workers are defined as high-skilled white-collar workers, the term "manual workers" refers to three categories: high- and low-skilled blue-collar workers and low-skilled white-collar workers. We apply this categorization of manual workers because the latter three groups are more similar based on observable characteristics.

4.1.2 Administrative data

Our second data source is comprised of administrative data that cover the entire Norwegian population. All residents are assigned to a unique personal identification number, which enables linking information from various administrative registers,

[14]The reason that the NorLAG survey responses cannot be connected to the administrative health registers applied in this study is that the NorLAG data were anonymized after they were connected to the registers.

[15]The first wave contains 5559 observations (response rate 67%), whereas the second wave contains 15,149 observations (response rate 60%).

[16]All SF-12 components are outlined in Section 5.2.2.

[17]Coded according to NACE Rev.1.1

such as health registers, income and social insurance registers, and population registers. These registers contain information on year and month of birth, death and retirement, as well as educational attainment, income, and social security uptake.

We apply two health outcomes from the administrative data. The first is a binary indicator of whether a person has been acutely hospitalized in a particular month. This information comes from the national patient register (NPR), which contains records of all inpatient and outpatient stays at Norwegian hospitals from 2008 through 2014. Admissions are coded by whether the hospitalization is a result of a planned or unplanned admission. The latter can be thought of as an acute admission in the sense that treatment has been deemed necessary, typically as a result of an accident, stroke, or severe heart condition.[18] The second health outcome is a binary indicator of whether a person passed away in a particular month. This information comes from the Norwegian cause-of-death registry which contains all recorded deaths in Norway from 1992 through 2014. Both outcomes thus yield the likelihood of the particular outcome at a specific age in months.

Importantly, these measures of health are not correlated with the time cost of consulting medical expertise. As individuals have more time at their disposal after retirement, the opportunity cost of seeking medical help is reduced for retirees compared to workers. It is therefore likely that the prevalence of a diagnosis or a medical treatment that is not acute increases after retirement, when the opportunity cost of seeing a physician has fallen. Applying health outcomes that are correlated with the opportunity cost of medical consultations can therefore erroneously lead to the conclusion that retirement caused the increased prevalence of such outcomes.

4.2 Sample selection

We restrict our attention to individuals aged 56–79 in 2007 and 2008 in both data sources. From the administrative records, we use data from 2008.[19] This leaves 4,619 individuals in the NorLAG sample and 892,908 individuals in the registered sample. The administrative data in our analysis is a panel data set, with monthly records of hospitalization, mortality, retirement, and age in months. As such, month by month, the treatment variable is determined according to age in months exceeding the retirement age threshold. Including fixed effects is unnecessary for identification in an RD design. Moreover, as the source of identification is a comparison between those just below and just above the threshold, which can be carried out with a single cross-section, imposing a specific dynamic structure introduces more restrictions without any gain in identification (Lee and Lemieux 2010). We therefore treat the sample from the register panel data as repeated cross-sections and pool all months together, treating each observation as an individual. This also makes the administrative data more comparable to the NorLAG data.

In order to maintain the intention to treat in the RD design and to ensure that we have enough data for inference, we place no further restrictions on the sample for the

[18] All admissions are coded in accordance with the International Statistical Classification of Diseases and Related Health Problems, ICD-10 (WHO 1992)).

[19] This is the earliest year in which data on hospitalizations are available.

main analysis. This means that our analytical sample will include individuals on DI or individuals who are not working for other reasons. Individuals on DI are automatically classified as retired once they hit the age threshold. In theory, we should expect no retirement effects for this group, as their work status remains unchanged when they retire. This would bias our results toward zero. However, the health outcomes in the survey data can suffer from justification bias (McGarry 2004). Being on disability insurance might make an individual, consciously or subconsciously, under-report their health in order to justify their status as disabled. The need for this justification is no longer present once they are transferred to retirement pension. In this case, the estimates would be biased upwards and we might worry that the positive effect on health was driven by these individuals. As a sensitivity analysis, we therefore run the whole analysis including only individuals who were gainfully employed or working until retirement.

Ideally, we want to compare individuals working up to the retirement age to individuals who retired from working. In the NorLAG data, this is done by adjusting the sample by two rules. The first rule implies including only individuals who had income from labor the previous year in the analysis; the second rule implies including only individuals who have stated that they are working or were working before they became retired. 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 labor income in 2006. 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. When 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 to maintain continuity across the retirement threshold. In the administrative data, we define individuals as working if they currently have positive income or if they had positive income before retirement. We find that these sensitivity analyses do not alter our conclusions.[20]

4.3 Descriptive statistics

Table 2 displays summary statistics for the NorLAG and administrative data. The first 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 (the first month after turning 67 years). These are the observations within the bandwidth used in the estimations of the short-term retirement effects. The third column presents a t test of balanced covariates. The t tests confirm that individuals on either side of the threshold are similar with respect to gender, education, living arrangements, and occupation for the NorLAG data and for gender and living arrangements for the administrative data. For the administrative data, the level of education is significantly higher below the threshold than above it, which is likely due to those above being marginally older than those below the threshold.[21]

[20] The results from the sensitivity analysis are shown in the Appendix.

[21] This is also true for the NorLAG data, but the differences in educational attainment is statistically significant in the administrative data due to a large number of observations.

Table 2 Descriptive statistics

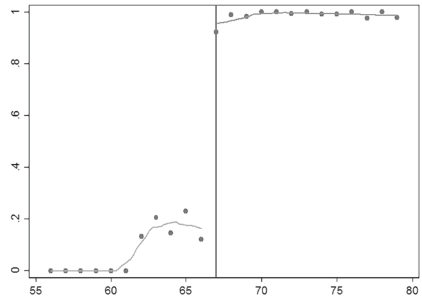
	Below threshold	Above threshold	*t* test
Source: NorLAG			
SF12	45.73	47.55	− 1.48
	(12.03)	(10.12)	
Age	66.15	67.00	59.84
	(0.36)	(0.0)	
Retired	0.18	0.96	− 34.86
Elementary education	0.25	0.25	− 0.07
High school degree	0.45	0.51	− 0.99
Any college	0.30	0.25	1.19
Professional	0.47	0.50	− 0.25
Manual	0.40	0.41	− 0.06
Living with partner	0.75	0.72	0.75
Female	0.47	0.50	− 0.71
Observations	190	200	
Source: Admin. Data			
Acute hospital admissions	0.140	0.141	−0.08
Mortality	0.017	0.018	− 1.69
Age	66.19	67.00	− 480
	(0.38)	(0.00)	
Retired	0.29	0.95	− 290
Elementary education	0.32	0.34	− 7.52
High school degree	0.46	0.45	2.34
Any college	0.23	0.21	5.17
Married	0.64	0.64	1.07
Female	0.51	0.51	− 0.38
Observations	31,751	33,752	

This table displays descriptive statistics for the two data sources, the NorLAG data (top panel) and the administrative data (bottom panel). Column 1 displays the means for the sub-samples aged 795–804 months (control group) and column 2 for those aged 805–814 months (treatment group). Column 3 presents the *t* statistics based on a *t* test between column 1 and column 2. Standard deviations in parenthesis

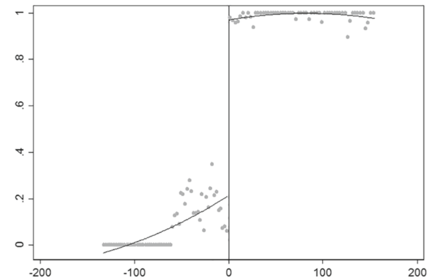
5 Results

5.1 Graphical results

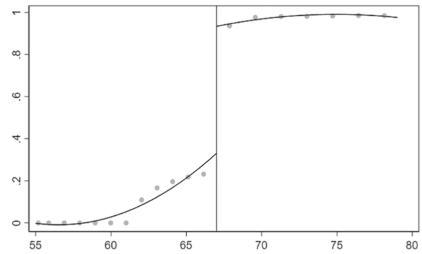
To motivate the use of the FRD design, Fig. 1 displays the share of retired individuals from age 55 until age 79. The two upper graphs are constructed using the survey



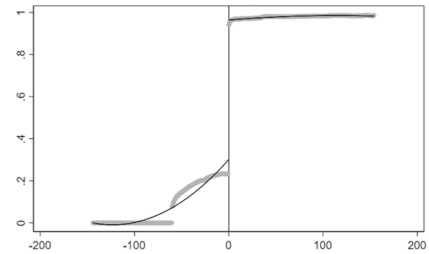
(a) NorLAG Data: Fraction Retired - Age in Years



(b) NorLAG Data: Fraction Retired - Age in Months



(c) Adm. Data: Fraction Retired - Age in Years



(d) Adm. Data: Fraction Retired - Age in Months

Fig. 1 Discontinuity in retirement at the retirement age threshold. The graphs show the fraction retired by age from the two datasets. The upper graphs are based on the survey data, whereas the two lower graphs are based on administrative data. All graphs depict the fraction retired across the age span of 55–79. The x-axis on the left two graphs depicts age in years, whereas the x-axis in the graphs to the right depicts age in months, relative to the retirement eligibility age in months (805 months)

data, whereas the two lower graphs are constructed using the administrative data. The age span in the four graphs is the same (55-79), but the x-axis on the two left graphs depicts age in years, whereas the x-axis on the two right graphs depicts age in months. The latter is to show that the discontinuity in retirement coincides with the first month after turning 67 (the first month of retirement eligibility).

In all four figures, the patterns are very similar.[22] There is a substantial discontinuity in the likelihood of being retired at age 67 (805 months). Since some workers chose to retire early, we also see a small discontinuity at age 62, the lowest eligibility age for ER. Only a negligible share of individuals chose to retire later than age 67. The graphical evidence thus provides evidence of a clear retirement response at the statutory retirement age. We build our empirical analysis on this discontinuity (age 805 months).

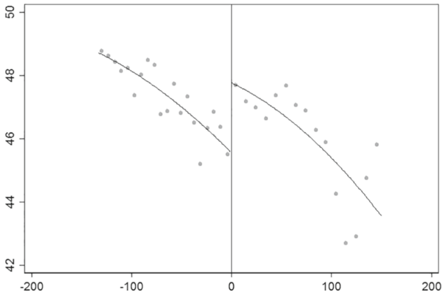
[22]In the graphs, retirement refers only to those who have actually retired, either through the early pension program or at the retirement age of 67. This means that individuals on DI are not considered retired. If we remove all individuals that are currently on DI or who were on DI before they retired, from our sample, the picture looks the same.

Figure 2 presents graphical evidence on the relationships between health and age for the three outcomes used in our study: physical health, acute hospitalizations, and mortality. The age range spans from 55 to 79 years, and the x-axes are depicted as age in months 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.

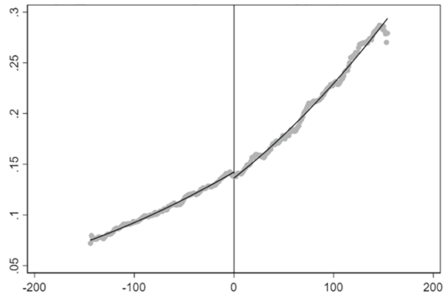
Graph (a) in Fig. 2 shows the observed health pattern for physical health for all individuals aged 56-79 in the NorLAG sample. In general, physical health declines with increasing age, but there is a substantial jump at the retirement threshold. At this threshold, the trajectory shifts up to a level of someone 80 months younger, which amounts to 6.5 years.

Graphs (b) and (c) depicts the incidence of acute hospitalizations and mortality respectively. We see that the incidence rates increase across the age span of 56–79, but there does not seem to be any substantial discontinuities at the retirement threshold in the outcomes reflected in the graphs. For acute hospitalizations, we see a small, possibly negligible, downward shift at the threshold.

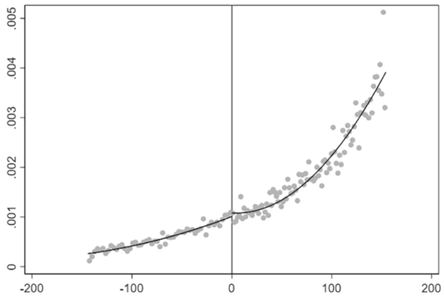
There is an ongoing debate as to whether the cumulative or the contemporaneous effects of retirement are larger (Coe and Zamarro 2011; Mazzonna and Peracchi 2017). As mentioned above, RD can only identify effects close to the threshold, so



(a) NorLAG Data: Physical Health (SF-12)



(b) Administrative Data: Acute Hospitalizations



(c) Administrative Data: Mortality

Fig. 2 Discontinuity in health at the retirement age threshold. The graphs present the age-health relationship for physical health, acute hospital admissions, and mortality. The scale for physical health is a point on the SF-12 scale, and the scale for acute hospital admissions and mortality corresponds to the incidence in the population. The x-axis displays age in months relative to the retirement age threshold of 805 months any prolonged retirement effects become mere speculation in this setting. However, visual inspection of the graph for physical health, a in Fig. 2, provides suggestive evidence of a prolonged effect of retirement on physical health, as retirement shifts individuals to a higher health trajectory, where they seem to stay as age increases.

5.2 Regression results

We present the 2SLS regression results for all three health dimensions in Tables 3 and 4. The effects are estimated using a bandwidth of +/− 10 months around the threshold, which is the optimal bandwidth using the selector suggested by Imbens and Kalyanaraman (2012). We estimate the effects for each gender and for the different SES groups separately. In Table 5, we present results for some of the different elements that go into constructing the physical health measure, and in Table 6, we present results from a formal test of heterogeneous retirement effects in which the instrument is interacted with indicators of gender and the different SES groups.

In Table 3, we present the first stage of the 2SLS regression results. This is the estimated effect of crossing the statutory retirement age on the probability of retirement, i.e., γ from Eq. (2). 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 presented in Fig. 1.

5.2.1 The effect on physical health

The first row in Table 4 displays the results of the short-term retirement effects on physical health. We find that retirement leads to a 5.7-point increase in physical

Table 3 First-stage regressions

	All	Men	Women
Source: NorLAG			
Retired	0.954***	0.941***	0.961***
	(0.0362)	(0.0587)	(0.0431)
Observations	371	190	181
Source: Admin. Data			
Retired	0.720***	0.683***	0.756***
	(0.00264)	(0.00389)	(0.0356)
Observations	825605	407386	418219

This table displays the first-stage regressions specified in Eq. (2). The reported coefficient is γ from Eq. (2). Estimation is done using a bandwidth of ten months. Standard errors in parentheses are clustered at the age in months level for the NorLAG data and at the individual level for the administrative data. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Table 4 Short-term retirement effects on health

	OLS	All	Men	Women	Manual	Professional	Lower	Higher
Physical health:								
Retired	4.015***	5.689***	8.036***	4.053	13.16***	− 0.333	8.358***	− 1.952
	(0.539)	(1.979)	(3.026)	(3.465)	(3.508)	(3.761)	(2.415)	(5.449)
Observations	4518	361	185	176	126	123	261	99
Hospitalizations:								
Retired	− 0.017***	− .00419	− .00417	− .00440			− .00589**	.00255
	(0.0016)	(.00258)	(.00395)	(.00339)			(.00292)	(.00535)
Observations	12612395	825605	407386	418219			643441	182164
Mortality:								
Retired	− 0.001***	− .000123	− .0000355	− .000266			− .0000895	− .000299
	(0.0000)	(.000204)	(.000343)	(.000236)			(.000233)	(.000399)
Observations	12852816	840239	416611	423628			655743	184496

This table displays the impact of retirement on the health outcomes. The first column (OLS) presents OLS results based on estimation of Eq. (1) on the full age span (56–79). The controls include age, educational level, occupation (only for physical health), gender, and marital/partner status, and the reported coefficient is β_1. The rest of the results are from 2SLS local estimations. The reported coefficient is τ from Eq. (3). Estimation is done using a bandwidth of ten months. All refers to the sample as a whole, Professional, and Manual to type of occupation and Lower and Higher to levels of education. Standard errors in parentheses are clustered at the age in month level for physical health and at the individual level for acute hospitalizations and mortality. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 5 Short-term retirement effects on health by SF12 components

	Absolute ability	Relative ability			
	Functional	Daily tasks	Specific tasks	Tasks profoundly	Pain
All	0.069	0.171**	0.229**	−0.075**	0.228***
	(0.0798)	(0.0776)	(0.0934)	(0.0377)	(0.0785)
Observations	369	368	369	368	368
Men	0.033	0.325***	0.401**	− 0.095	0.178*
	(0.1406)	(0.1234)	(0.1682)	(0.0803)	(0.1005)
Observations	189	189	189	188	189
Women	0.116	0.038	0.069	− 0.056	0.317***
	(0.1088)	(0.1280)	(0.1519)	(0.0419)	(0.1223)
Observations	180	179	180	180	179
Manual	0.258**	0.558***	0.551***	− 0.070	0.503***
	(0.1247)	(0.1277)	(0.1317)	(0.1231)	(0.1345)
Observations	127	126	127	127	127
Professional	0.011	− 0.094	− 0.016	0.046	− 0.034
	(0.1663)	(0.1233)	(0.1569)	(0.0486)	(0.1782)
Observations	123	123	123	123	123
Low educ.	0.150*	0.284***	0.292***	− 0.121***	0.308***
	(0.0903)	(0.1095)	(0.0971)	(0.0441)	(0.1055)
Observations	268	267	268	267	267
High educ.	0.124	− 0.104	0.060	0.021	− 0.018
	(0.2000)	(0.1557)	(0.1943)	(0.0723)	(0.1588)
Observations	100	100	100	100	100

This table displays the impact of retirement on selected components of the physical health outcome (SF-12). Positive coefficients indicate improved health. The reported coefficient is τ from Eq. (3). Estimation is done using a bandwidth of ten months. Standard errors in parentheses are clustered at the age in month level. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

health for the population as a whole (second column). This is slightly higher than the OLS estimate (first column). The effect is substantial given that the mean and standard deviation for this health outcome are 47 and 10 points respectively. We find a strong and positive effect for men (8 points), and a positive (4 points) but not statistically significant effect for women.

Based on the discussion in the introduction, we can expect different health effects of retirement depending on the level of education and type of occupation. The last four columns in Table 4 show the effects for the different SES groups. For manual workers and the low educated, the effects are large (13.2 and 8.4 points respectively) at about one standard deviation, and significant at the 1% level. For the high SES groups, we find no statistically significant effects and the coefficients are closer to zero.

Table 6 Formal test of differences by socioeconomic status

	Education	Gender	Occupation
Physical health			
Retired	4.975	3.696	6.858*
	(3.115)	(2.549)	(3.305)
Observations	361	361	249
Hospitalizations			
Retired	− .00294	.00224	
	(.00533)	(.00509)	
Observations	825605	825605	
Mortality			
Retired	.000066	.0000171	
	(.0000145)	(.000139)	
Observations	840239	840239	

This table displays the interaction between retirement eligibility and SES (education and occupation (only for the NorLAG)) and gender. The first row presents the results for physical health, and the second and third rows present the results for acute hospitalizations and mortality. The reported coefficient is γ from Eq. (4). Estimation is done using a bandwidth of ten months. Standard errors in parentheses are clustered at the age in month level for physical health and at the individual level for acute hospitalizations and mortality. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Power calculations show that a sample of at least 90 is needed to ensure a power of 0.8. Although well above this threshold, the sub-group samples are fairly small. 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). As a check, we ran the whole analysis using a bandwidth of 20 months. This about doubles the observations in each sub-group, but the sizes and significance levels of the estimates remain about the same. Figure 3 in the Appendix presents an overview of the point estimates and confidence intervals using a range of different bandwidths.

Our findings are in line with the evidence from Coe and Zamarro (2011) and Eibich (2015), who suggest that, in general, retirement leads to an increase in physical health in Europe. Although our results are estimates of the short-term effects, previous findings suggest that retirement also has a cumulative effect on physical health through increased physical activity (Eibich 2015). Our results on group heterogeneity 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 with individuals with low SES. Moreover, Insler (2014) suggests that wealthy people have more time to invest in their health while working. In an effort to say more about what aspects of health are improved by retirement, we look further into the physical health outcome (SF-12) in the next section.

5.2.2 Looking further into the effect on physical health

SF-12 is composed of survey responses to the following questions:[23]

1. Rate your health on a scale from 1-5 (self-rated health).
2. Is your health of such a character that it limits you in doing tasks like moving a table, vacuuming, hiking, or gardening?
3. Is your health of such a character that it limits you from climbing several flights of stairs?
   The following questions concern the last four weeks:
4. Has your physical health limited you in doing your daily tasks so that you have accomplished less than you wished for?
5. Has your physical health limited you from completing specific types of tasks?
6. Have psychological problems limited you from doing daily tasks so that you have accomplished less than you wished for?
7. Have psychological problems limited you from doing daily tasks as profoundly as you usually do?
8. Has pain limited you from doing your daily tasks?
9. Have you been feeling calm and harmonious?
10. Have you been feeling energized?
11. Have you been feeling downhearted and blue?
12. Have physical or mental health limited you from socializing as much as you wanted to?

Out of the 12 components that go into the SF-12, five were significantly impacted by retirement. These are questions 2, 4, 5, 7, and 8. Each question is coded as a binary variable, where one means that health or pain is not experienced as limiting, i.e., positive effects indicate improved health. In Table 5, we present the results for these five components.[24]

Retirement was found to reduce the experience that physical health is a limiting factor in accomplishing as much as one would like, and as a limiting factor in doing specific tasks. The former holds for both men and women, whereas the latter holds for men. We find particularly strong effects on reduced pain, especially for women. Furthermore, we find that, in general, having problems with completing tasks profoundly due to psychological health was reduced by retirement.

When we assess the effects of the different SES groups, we find that it is manual workers or lower-educated individuals who experience reduced pain and reduced limitations due to physical and mental health issues. We find no effects for the high SES groups. We also find that, for the low SES group, retirement improved functional health as measured by the ability to perform tasks such as vacuuming, moving a table, hiking, or gardening. These effects are statistically significant at the 5 percent level for manual workers and at the 10% level for the low-educated group. This implies that retirement not only affects relative abilities, i.e., reduces the burden of daily tasks, but also affects health in a more fundamental way.

[23] Translated from Norwegian by the authors.

[24] A table presenting the effects of all SF-12 components is presented in the Appendix (Table 7).

5.2.3 The effect on acute hospitalization

We now turn to our estimates from the administrative data. Acute hospitalization is based on a dummy for inpatient care in which treatment is deemed necessary. The results are presented in the second row in Table 4.

First, we explore how retirement affects acute unscheduled hospitalizations for the population on average and by gender. Although the OLS estimations show that retirement is associated with a statistically significant 1.7 percentage point reduction in acute hospitalizations, the 2SLS local estimations yield an effect size of about 0.4 percentage points that is insignificant. The same holds for both men and women separately. When we divide the population by SES, we find that retirement leads to a 0.6 percentage point reduction in the likelihood of acute hospitalizations for the low SES group. As the incidence of acute hospitalizations is 14%, this amounts to a 4% reduction in the likelihood of acute hospitalizations. The effect is significant at the 5% level. For the high SES group, we find an effect of 0.3, yet this is not significantly different from zero.

One way to think of these results is that retirement for the population in general leads to no short-term change in serious health conditions. Hallberg et al. (2015) studied a targeted early retirement offer to workers in the military at age 55 and find that the number of days in inpatient care is significantly reduced at the ages 61-70. One possible drawback with our method is that the RD design only captures the short-term effect of retirement, and the potential health gain of retirement is possibly not yet present in the months immediately after retirement. For instance, Hallberg et al. (2015) find a 4.7 days reduction in inpatient care 6–10 years after early retirement, whereas the estimated effect is 2 days in the first year after early retirement.

To some extent, the same intuition can be found in Behncke (2012), despite opposite direction of effects. She shows that retirement increases the risk of being diagnosed with a chronic condition in the subsequent years after retirement. However, assessments applying less acute diagnoses can be confounded for two reasons. First, the opportunity cost of seeking medical help is greatly reduced after retirement, hence increasing the likelihood of detecting such conditions. Second, the reason for seeking medical help can differ for individuals who are working and individuals who are retired. In Norway, for example, a sickness absence from work for longer than the self-certified absence period[25] must be certified by a physician, which means that retirees and employees most likely visit the doctor for different reasons.

5.2.4 The effect on mortality

The results described in the previous sections suggest that retirement leads to a short-term positive effect on subjective measures of health, whereas we find no or small effects on the number of acute hospitalizations. Given the latter findings, *a priori*, we expect to see little or no short-term effect on mortality. In the lower part of Table 4, we display the estimation results for mortality. Although retirement is associated

[25] A medical certificate is required for spells of absence of more than 3 days or 8 days, depending on whether the employer has signed the “IA-agreement” or not.

with a 0.1 percentage points reduction in mortality (OLS results), we find no short-term causal effect of retirement on mortality. Regardless of gender and sub-group, the estimates remain statistically indistinguishable from zero.

The question remains whether a short-term effect of retirement on relatively serious outcomes, such as mortality, is implausible in the short run. Hallberg et al. (2015) use Cox regression models to form hazard ratios and find that early retirement at age 55 reduces the risk of dying at age 70 by around 26%. Studying the first 5 years after an early retirement window, Bloemen et al. (2017) find a drop in the probability of dying of around 2.6%. The same effect is found in Blake and Garrouste (2013) and Kuhn et al. (2010), albeit the latter only for male blue-collar workers. However, studying the introduction of early retirement in Norway, Hernæs et al. (2013) find no effect of early retirement on mortality. They follow workers until a maximum of 77 years of age, with eligibility for early retirement varying between 62 and 65 years of age. They conclude that early retirement in itself has no effect on mortality.

Taken together, our results show that in general there are no effects of retirement on objective health outcomes. However, as this and several other studies show, retirement affects subjective health. Moreover, our results are clear in that retirement affects subjective health for the low SES groups, but not the high SES groups. The gender dimension does not appear to be important as we find little evidence of gender differences in the estimated effects. Despite large inter group differences by SES group and small differences by gender, a formal test of effect heterogeneity is needed to be precise about these differences. In the next section, we formally test whether there are significant differences in the retirement effects on health by gender or SES group.

5.2.5 A formal test of effect heterogeneity

Table 6 presents the results from the formal test of heterogeneity. These are the results of a reduced form of Eq. 3, where the instrument is interacted with SES groups and gender. We estimate the following:

$$\begin{aligned}\text{Health}_i &= \beta_0 + \gamma 1[\text{Age}_i \geq c] \times \text{SES}_i/\text{Woman}_i + \beta_1 1[\text{Age}_i \\ &\geq c] + \beta_2 \text{Age}_i^B + \beta_3 \text{Age}_i^A + e_i, \end{aligned} \tag{4}$$

where γ is the coefficient of interest and $1[\text{Age}_i \geq c]$ is the instrument indicating whether age in months is equal to or exceeds the retirement age threshold. SES is a binary indicator of either manual workers or low education, and gender differences are identified by the binary indicator: Woman. We apply the same +/- 10 months bandwidth in these estimations.

We see that the effects of retirement are statistically different from each other when SES is measured by occupation. Although the estimated effects differ quite substantially by the educational group as shown in Table 4, 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. Hence, we show that accounting for differences by SES can be important in analyses of retirement effects on health.

5.3 Robustness checks and sensitivity analysis

The results from the robustness checks are presented in the Appendix, but we provide a brief overview here. First, we show that our results on physical health and mortality are robust to different bandwidths. For acute hospitalizations, however, we find that increasing the bandwidth from 10 to 15 months yields significant (negative) effects. The effects are small, ranging from 0.7 to 1 percentage points. However, due to the small incidence of this outcome, this implies that retirement leads to a 5–7% reduction in the likelihood of acute hospitalization. One reason for this finding can be that including more post-retirement months increases the likelihood of finding significant effects as it may take some time for retirement to affect severe health issues such as stroke and acute heart conditions. Another reason can be that increasing the bandwidth increases the likelihood of factors, other than retirement, affecting acute hospital admissions, i.e., it increases bias. It is also worth noting that when we run the entire analysis on acute hospitalizations using a bandwidth of 20, we find larger (and negative) significant effects for the population as a whole, for men, and for the low-educated group. Effects sizes range from 1 to 1.5 percentage points and are significant at the 5% level. Using this bandwidth, we still find no significant effects of retirement on mortality.

Next, we look for discontinuities at the retirement age threshold in a covariate that should not be affected by the treatment, in this case: marital status. Although retirement can affect the likelihood of being married, it is highly unlikely in the immediate aftermath (within 10 months) of retirement. We find no retirement effects on the likelihood of living with a partner or spouse (the NorLAG data) or on being married (the administrative data).

Third, we perform placebo tests by checking for discontinuities in the health outcomes at values of the forcing variable (age in months) where there should be no discontinuities. We find no discontinuities in the health outcomes at the placebo thresholds of age 61 and 73 for physical health or mortality, but we find some inconsistencies at these thresholds for acute hospitalizations. The effects are smaller than at the retirement threshold, yet significant. Thus, we might worry that this outcome is prone to be discontinuous at arbitrary age thresholds.

We then test for discontinuities in the conditional density of the forcing variable (age in months) to avoid self-selection or sorting into the treatment or control groups. The RD design may be invalid if individuals just above the threshold are more likely to participate in a survey than those just below the threshold, i.e., violating the RD assumption that the running variable is continuous at the threshold. In the Appendix, we provide histograms that display the distribution of age in months in the NorLAG data. There is no apparent discontinuity at the threshold in these histograms. Moreover, we applied the local polynomial density estimator for testing the null of continuous density of the forcing variable at the threshold proposed by Cattaneo et al. (2016). The p value under this test is 0.3251.

Finally, the results for physical health and mortality are robust to the different subsamples that are conditional on working or working until retirement, as described in Section 4. For acute hospitalizations, we find the same results as in the main analysis for all sub-groups except for the lower SES group, where the negative impact of retirement on the likelihood of acute hospitalization is no longer found when we condition on working or working until the retirement age.

6 Conclusion

Whether retirement has a causal effect on health is a difficult question to answer because of selection into retirement. In this paper, we study the short-term health effect of retirement using the statutory retirement age at 67 in a FRD design. We exploit the fact that once individuals reach the statutory retirement age, the probability of claiming retirement pension drastically increases. We apply both subjective measures of health from survey data and objective health outcomes from administrative data, where the latter covers the entire Norwegian population.

We find that, on average, in the population, retirement has a positive effect on self-assessed physical health, but no effects on the objective measures of health: acute hospitalizations and mortality. When we assess the effects on different SES groups, we find that retirement has a large, positive effect on physical health and reduces the likelihood of acute hospitalizations among the low SES groups. We find no significant effects for the high SES groups for any of these outcomes. For mortality, we find no significant effects of retirement for any group.

We thus confirm what has been found in several studies, namely that retirement has a positive effect on health for subjective health outcomes. How this manifests to objective outcomes is less clear as there exist little evidence using objective health outcomes, especially on the full population. In general, we find no effects on the objective outcomes, besides suggestive evidence of a retirement effect on reduced likelihood of hospitalizations for the low SES group. However, this result does not pass the robustness tests, and must therefore be interpreted with caution.

Based on our findings, we can conclude that retirement mainly affects subjective outcomes and not objective ones. However, both acute hospitalizations and mortality are extreme outcomes in the sense that they represent very poor health. Retirement effects on these outcomes are potentially not detectable in the short-term. When we assess the factors that go into the physical health outcome, SF-12, we find that the positive health effect was driven by a few different factors. On the one hand, finding that retirement reduces the likelihood of reporting that health is limiting in managing in daily chores and in conducting specific chores profoundly can be due to the fact that work (possibly a health consuming chore) is no longer present, so health feels less limiting. This implies that the underlying health has not changed, but the presence of health consuming activities has. On the other hand, we also found that retirement reduced the presence of pain and reduced the likelihood of reporting difficulties with activities such as vacuuming, moving a table, hiking, or gardening. This indicates that retirement affects health in a more fundamental way. Future research should thus assess objective health outcomes that are less extreme. In doing so, it is key to recognize that retirement necessarily coincides with reduced opportunity cost of time.[26]

[26] See Jürges (2007) for a more general discussion of subjective versus objective measures of health.

This study accentuates the importance of assessing the potential heterogeneity in the health effects for individuals in different circumstances. Occupation, more than education, determines the socioeconomic differences in the effects of retirement on health. Our findings indicate that retirement reforms aimed at prolonging working life can be socially distortive due to the differential effects based on SES. We find that retirement at age 67 has positive effects on health for the low SES groups, but we find no effects for the high SES groups. A formal test of these differences confirms that occupation matters for the health effects of retirement.

Finally, our study contributes to generalizing the positive physical health effect of retirement found in the literature across a larger age span. The current literature has mainly assessed retirement ages from the late 50s to about 65. Here, we confirm that the positive effects still hold for individuals retiring at age 67. Two factors might be of concern for the external validity of our results. First, workers who retire at age 67 are workers who are healthy enough to remain in the labor force until this high age threshold. Second, the labor force participation of older workers is relatively high in Norway compared with other European countries. However, assessments of higher age thresholds are valuable for policymakers as current retirement reforms typically aim at increasing labor force participation among older workers. These reforms will likely affect relatively healthy individuals, i.e., workers who can remain employed until these higher retirement ages.

Acknowledgments We would like to thank Arild Aakvik, Ragnhild Balsvik, Dorly Deeg, Elisabeth Fevang, Astrid Grasdal, Thomas Hansen, Thomas Hofmarcher, Christian Monden, Matthew Neidell, Viggo Nordvik, Miikka Rokkinen, Ingrid Hoem Sjursen, Eirik Strø mland, Kristin Årland, and seminar participants at Columbia University, Norwegian Social Research, 23rd Nordic Congress of Gerontology, Norwegian School of Economics, University of Bergen, The Frisch Centre, and three anonymous referees for helpful comments and suggestions.

Funding Information This study was funded by The Research Council of Norway (NFR), grant number 237011 and grant number 50813, and NordForsk, grant number 74637.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest

Appendix - Sensitivity and Robustness

A.1 Disabled individuals: past labor income and self-reported work status

People on DI are mechanically transferred from disability pension to retirement pension at age 805 months. We need to make sure that the positive physical health effects found in the main analysis are not driven by these individuals. Initially, there is no reason to believe that there should be a health effect for people on DI, as they were not working before retirement, and should therefore have no change in circumstances. However, as the physical health measure contains elements of self-assessed health, it is possible that someone who is disabled may need to justify their disability status, consciously or subconsciously, by under-reporting health. In this case, health

Table 7 All SF-12 components

	All	Men	Women	Manual	Professional
Absolute ability:					
Functional	0.069	0.033	0.116	0.258**	0.011
	(0.0798)	(0.1406)	(0.1088)	(0.1247)	(0.1663)
Climbing stairs	0.111	0.120	0.135	0.313	−0.045
	(0.0850)	(0.1091)	(0.1638)	(0.2080)	(0.1104)
Observations	369	189	180	127	123
Relative ability:					
Poor Health	0.068	0.189	−0.070	−0.027	0.075
	(0.0859)	(0.1421)	(0.1312)	(0.1693)	(0.1546)
Daily tasks physical	0.171**	0.325***	0.038	0.558***	−0.094
	(0.0776)	(0.1234)	(0.1280)	(0.1277)	(0.1233)
Specific tasks physical	0.229**	0.401**	0.069	0.551***	−0.016
	(0.0934)	(0.1682)	(0.1519)	(0.1317)	(0.1569)
Daily tasks mental	−0.041	−0.063	−0.013	−0.016	0.065
	(0.0389)	(0.0491)	(0.0618)	(0.0687)	(0.0604)
Tasks profoundly mental	−0.075**	−0.095	−0.056	−0.070	0.046
	(0.0377)	(0.0803)	(0.0419)	(0.1231)	(0.0486)
Pain	0.228***	0.178*	0.317***	0.503***	−0.034
	(0.0785)	(0.1005)	(0.1223)	(0.1345)	(0.1782)
Feeling Harmonious	0.094	0.229***	−0.048	0.168	0.109
	(0.0818)	(0.0691)	(0.1264)	(0.1222)	(0.1244)
Feeling Energized	0.037	0.041	0.035	0.036	0.170
	(0.0784)	(0.0735)	(0.1281)	(0.1688)	(0.1595)
Feeling blue	0.043	0.084	0.020	−0.126	0.048
	(0.1040)	(0.1131)	(0.1370)	(0.1672)	(0.1735)
Limited socializing	−0.003	0.073	−0.061	0.077	−0.072
	(0.0843)	(0.1135)	(0.0880)	(0.1472)	(0.1369)
Observations	368	188	180	127	123

Note: This table displays the effect on all components of the physical health outcome (SF-12). Positive coefficients indicate improved health. The reported coefficient is τ from Eq. 3. Estimation is done using a bandwidth of ten months. Standard errors in parentheses are clustered at the age in month level. *p<0.10, **p<0.05, ***p<0.01

prior to retirement may be under-reported. Post retirement, when they are no longer in a situation where poor health is defining their labor market status, they might feel healthier, or they no longer have the need to report poor health. Provided this is a plausible scenario; We need to rule out that the results found in Section 5 are driven by justification bias.

The first two rows of Table 8 displays the results on two sub-samples of the survey data (labeled “Working” and “Income”), each aimed at running the analysis only on the sub-sample that was recorded as working until the statutory retirement age. The working sub-samples are defined in Section 4.2. Finding coefficients of the same sign and magnitude, especially for the rule based on self-assessed work status, ensures us that these effects are not driven by the disability justification hypothesis. The estimations based on the income-rule yields large and insignificant coefficients, both a consequence of the small sample sizes. Yet, the direction of the effects is similar to what was found in the main analysis.

For the outcomes from the administrative data, we should expect that individuals who retire formally at 67 but without any actual change in circumstances should

Table 8 Robustness checks survey data: physical health

	All	Men	Women
Conditional on income	16.42***	15.83	−1.553
	(2.966)	(10.88)	(7.264)
Observations	82	53	39
Conditional on working	6.274***	9.741***	2.523
	(2.089)	(3.758)	(7.312)
Observations	247	142	105
Bandwidth 7	9.472***	14.69***	2.623
	(2.019)	(5.206)	(4.245)
Observations	275	142	133
Bandwidth 15	5.801***	9.391***	2.623
	(2.130)	(3.109)	(6.628)
Observations	540	278	262
Placebo at 61	−1.441	.971	−5.752
	(3.665)	(4.220)	(6.628)
Observations	454	242	212
Placebo at 73	−1.111	-1.264	.628
	(1.685)	(4.786)	(2.213)
Observations	251	127	124
Living with a partner	−0.106	-0.0413	−0.162
	(0.0931)	(0.108)	(0.176)
Observations	371	190	181

Note: This table displays the various robustness checks described in the Appendix, for the physical health outcome and the NorLAG data. 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 administrative data: acute hospitalizations

	All	Men	Women	Low Educ.	High Educ.
Conditional on working	0.00212	0.00266	0.00133	0.000309	0.00687
	(0.00295)	(0.00426)	(0.00397)	(0.00345)	(0.00566)
Observations	362,857	203,212	159,645	259,427	103,430
Bandwidth 7	−0.00246	−0.00117	−0.00374	−0.00343	0.00231
	(0.00227)	(0.00343)	(0.00303)	(0.00258)	(0.00459)
Observations	583,686	287,791	295,895	455,797	127,889
Bandwidth 15	−0.00722*	−0.00977*	−0.00520	−0.0101**	0.00321
	(0.00377)	(0.00584)	(0.00487)	(0.00427)	(0.00773)
Observations	1,241,687	612,603	629,084	965,278	276,409
s Placebo at 61	−0.000587	−0.0000883	−0.00105*	−0.00155**	0.00199*
	(0.000510)	(0.000610)	(0.000541)	(0.000705)	(0.00104)
Observations	1,311,705	667,661	644,044	962,159	349,546
Placebo at 73	0.00106	0.00284*	−0.000557	0.00194**	−0.00321**
	(0.000676)	(0.00162)	(0.000861)	(0.000699)	(0.00153)
Observations	634,319	294,672	339,647	527,740	106,579

Note: This table displays the various robustness checks described in the Appendix, for acute hospital admissions. Standard errors in parentheses are clustered at the individual level. *p<0.10, **p<0.05, ***p<0.01

water down the effects, as the health measures from this data source are not subject justification bias. We can therefore expect that this assessment can uncover significant effect, not detected in the gross sample. The first rows of Tables 9 and 10 present the estimations restricted to “workers” for acute hospitalizations and mortality, respectively. Here, we find no significant results for any of the sub-groups. The significant result on hospitalizations found for men with low education in the main analysis is no longer present.

A.2 Robustness checks and validity of 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 affected by the threshold; and, we check for discontinuities in the outcomes of interest at points in the age distribution where there should be no discontinuities. This robustness assessment follows the suggestions in Imbens and Lemieux (2008) closely.

Table 10 Robustness checks administrative data: mortality

	All	Men	Women	Low Educ.	High Educ.
Conditional on working	0.0000837	0.000214**	−0.0000777	0.000105	0.0000131
	(0.0000754)	(0.0000981)	(0.000117)	(0.0000848)	(0.000164)
Observations	363,123	203,383	159,740	259,628	103,495
Bandwidth 7	−0.000204	0.0000123	−0.000396	−0.000206	−0.000201
	(0.000249)	(0.000421)	(0.000288)	(0.000285)	(0.000485)
Observations	593,966	294,309	299,657	464,453	129,513
Bandwidth 15	−0.000197	−0.000229	−0.000173	−0.000196	−0.000248
	(0.000164)	(0.000277)	(0.000188)	(0.000187)	(0.000322)
Observations	1,263,829	626,544	637,285	983,902	279,927
s Placebo at 61	−0.0000369	−0.00000279	−0.0000721	0.0000256	−0.000210
	(0.0000838)	(0.000143)	(0.000116)	(0.000108)	(0.000125)
Observations	1,324,398	675,316	649,082	972,873	351,525
Placebo at 73	−0.0000758	−0.000128	−0.0000321	−0.0000559	−0.000167
	(0.000128)	(0.000191)	(0.000192)	(0.000144)	(0.000335)
Observations	653,875	306,487	347,388	545,003	108,872

Note: This table displays the various robustness checks described in the Appendix, for mortality. Standard errors in parentheses are clustered at the individual level. *p<0.10, **p<0.05, ***p<0.01

A.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 using bandwidths of 7 and 15 months for physical health. Using a bandwidth of 7 months does not alter the results, whereas increasing the bandwidths to 15 months somewhat reduces the effects. This is not surprising given the downward slope of the health trajectory across age and the upward shift in this trajectory at the retirement eligibility threshold.

The results for hospitalizations and mortality are displayed in Tables 9 and 10. For acute hospital admissions, we find that increasing the bandwidth to 15 months yields significant, negative effects. The effects are still small ranging from 0.7 to 1 percentage points. As the incidence is 14%, this entails a 5–7% reduction in the likelihood of acute hospitalization. Increasing the bandwidth increases the likelihood of factors, other than retirement, affecting acute hospital admissions. Another explanation can be that it takes some time for retirement to take effect on health issues such as stroke and acute heart conditions, thus including more post-retirement months

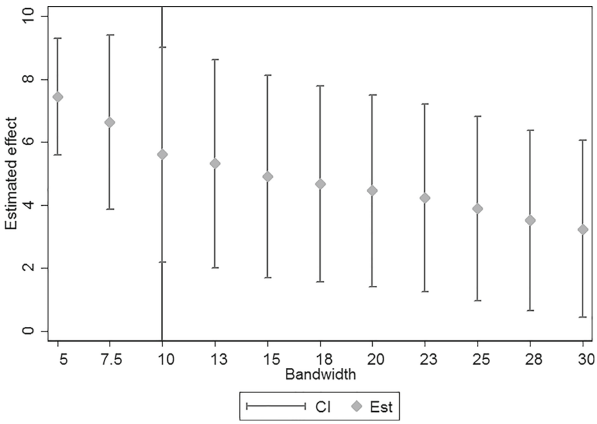


Fig. 3 Graphical Display of Different Bandwidths The graph displays point estimates and confidence intervals of the effect of retirement on physical health using different bandwidths. Bandwidth is measured as age in months from the retirement threshold

increase the likelihood of finding significant effects. As in the main analysis, we find no effects of retirement on mortality at any of these bandwidths.

A.2.2 Continuity of the forcing variable

Vital to any RD application is the individual's incapability of manipulating the forcing variable. In this case, the forcing variable is age (reported by public registers), which individuals cannot manipulate 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 4 shows two histograms of age in months assessing potential bunching at the threshold. There is no evidence of any discontinuity in

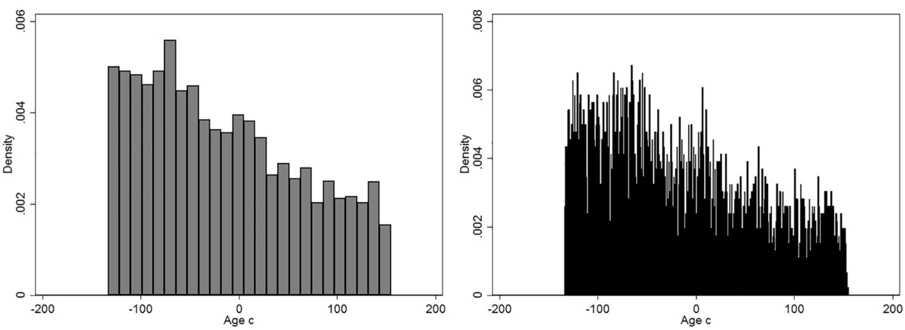


Fig. 4 Discontinuity of the Forcing Variable Note: The histograms show the distribution of age in months for the age range of 56-79 using the bin-width suggested by STATA (left histogram) and using one bin for each age in months (right histogram)

the forcing variable at the threshold. We also performed a more formal test proposed by Cattaneo et al. (2016). This implies testing the null of the continuous density of the forcing variable at the threshold using a local polynomial density estimator. The p-value under this test is 0.3251. For the population level data, this holds by construction, as people cannot manipulate their age and as all individuals in the population are represented in the data.

A.2.3 Placebo tests

The placebo tests entail testing for discontinuities in the three health outcomes at points in the age distribution where there should be no discontinuities. A common practice is to conduct placebo tests at the median age of the two sub-samples below and above the actual cut-off. In this case, the median age below the threshold is age 62. However, some individuals can retire at this age, thus making is an unsuited placebo threshold. Consequently, we use age 61 for the lower placebo. For the upper placebo, we use age 73. No discontinuities or significant effects were found at these placebo thresholds for physical health (Table 8). For acute hospital admissions (Table 9), we find significant effects for both the upper and lower placebo. For the lower placebo, this could be due to some occupations having special age-limits for retirement at 61. However, we find no explanations for why the upper placebo yields significant, and even positive effects. This finding reduces the credibility of the effects found in the main analysis for this outcome. The placebo results for mortality is presented in Table 10. There are no significant effects and the coefficients are close to zero for all sub-group at both placebo thresholds.

A.2.4 Discontinuity in Other Outcomes

Finally, we look for discontinuities in an outcome that should not be affected by retirement, at least not in the short-term. Here, we assess the likelihood of living with a partner or spouse (NorLAG) or being married (administrative data). The regression results shown in Tables 8 and 11 confirm that there are no retirement effect on these outcomes.

Table 11 Robustness checks administrative data: discontinuity in marital status

	All	Men	Women	Low Educ.	High Educ.
Married	0.00233	0.00324	0.00104	0.00156	0.00654
	(0.00229)	(0.00336)	(0.00311)	(0.00249)	(0.00576)
Observations	825,605	407,386	418,219	643,441	182,164

Note: This table displays the impact of retirement on the likelihood of being married. The reported coefficient is τ from Eq. (3). Estimation is done using a bandwidth of ten months. Standard errors in parentheses are clustered at the age in month level. *p<0.10, **p<0.05, ***p<0.01

References

Angrist JD, Pischke J-S (2010) The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. J Econ Perspect 24(2):3–30

Behncke S (2012) Does retirement trigger ill health? Health Econ 21(3):282–300

Blake H, Garrouste C (2013) Killing me softly: work and mortality among French seniors. Technical report HEDG, Department of Economics, University of York

Bloemen H, Hochguertel S, Zweerink J (2017) The causal effect of retirement on mortality: evidence from targeted incentives to retire early. Health Econ 26(12):204–218

Bound J, Waidmann T (2007) Estimating the health effects of retirement. Michigan Retirement Research Center Research Paper No. UM WP 2007-168

Brazier JE, Roberts J (2004) The estimation of a preference-based measure of health from the SF-12. Med Care 42(9):851–859

Case A, Deaton AS (2005) Broken down by work and sex: how our health declines. In: Analyses in the Economics of Aging. University of Chicago Press, pp 185–212

Cattaneo MD, Jansson M, Ma X (2016) Simple local regression distribution estimators with an application to manipulation testing. Unpublished Working Paper, University of Michigan, and University of California Berkeley

Coe NB, Zamarro G (2011) Retirement effects on health in Europe. J Health Econ 30(1):77–86

Dong Y (2015) Regression discontinuity applications with rounding errors in the running variable. J Appl Econom 30(3):422–446

Eibich P (2015) Understanding the effect of retirement on health: mechanisms and heterogeneity. J Health Econ 43:1–12

Godard M (2016) Gaining weight through retirement? Results from the SHARE survey. J Health Econ 45:27–46

Grossman M (1972) On the concept of health capital and the demand for health. J Polit Econ 80(2):223–255

Hahn J, Todd P, Van der Klaauw W (2001) Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica 69(1):201–209

Hallberg D, Johansson P, Josephson M (2015) Is an early retirement offer good for your health? Quasi-experimental evidence from the army. J Health Econ 44:274–285

Heller-Sahlgren G (2017) Retirement blues. J Health Econ 54:66–78

Hernæs E, Markussen S, Piggott J, Vestad OL (2013) Does retirement age impact mortality? J Health Econ 32(3):586–598

Imbens GW, Lemieux T (2008) Regression discontinuity designs: a guide to practice. J Econ 142(2):615–635

Imbens G, Kalyanaraman K (2012) Optimal bandwidth choice for the regression discontinuity estimator. Rev Econ Stud 79(3):933–959

Insler M (2014) The health consequences of retirement. J Hum Resour 49(1):195–233

Jenkinson C, Layte R (1997) Development and testing of the UK SF-12. J Health Serv Res 2(1):14–18

Jürges H (2007) True health vs response styles: exploring cross-country differences in self-reported health. Health Econ 16(2):163–178

Kudrna G (2017) The Norwegian pension reform: an external perspective. CEPAR Working Paper 2017/07, CEPAR

Kuhn A, Wuellrich J-P, Zweimüller J (2010) Fatal attraction? Access to early retirement and mortality. IZA discussion paper No. 5160

Lee DS, Card D (2008) Regression discontinuity inference with specification error. J Econ 142(2):655–674

Lee DS, Lemieux T (2010) Regression discontinuity designs in economics. J Econ Lit 48(2):281–355

Mazzonna F, Peracchi F (2012) Ageing, cognitive abilities and retirement. Eur Econ Rev 56(4):691–710

Mazzonna F, Peracchi F (2017) Unhealthy retirement? J Hum Resour 52(1):128–151

McGarry K (2004) Health and retirement do changes in health affect retirement expectations? J Hum Resour 39(3):624–648

Neuman K (2008) Quit your job and get healthier? The effect of retirement on health. J Labor Res 29(2):177–201

Røed K, Haugen F (2003) Early retirement and economic incentives: evidence from a quasi-natural experiment. Labour 17(2):203–228

Rohwedder S, Willis RJ (2010) Mental retirement. J Econ Perspect 24(1):119–138
Slagsvold B, Veenstra M, Daatland SO, Hagestad G, Hansen T, Herlofson K, Koløen K, Solem PE (2012) Life-course, ageing and generations in Norway: The NorLAG study. Norsk Epidemiol 22(2)
Ware JE Jr, Kosinski M, Keller SD (1996) A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 34(3):220–233
WHO (1992) The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines. World Health Organization, Geneva