Title: How does health change after retirement, and for who? New evidence using objective and subjective health measures

Ludovico Carrino[1], Merel Schuring[2], Alex Burdorf[2], Mauricio Avendano[1,3]

Affiliations :
1 Department of Global Health & Social Medicine, King's College London, UK
2 Erasmus MC, Department of Public Health, Rotterdam, The Netherlands
3. Department of Social and Behavioural Sciences, Harvard University, Boston, USA

Funding: Ludovico Carrino and Mauricio Avendano are supported by the Economic and Social Research Council (ESRC) and the Joint Programme Initiative 'More Years, Better Lives' (ES/P000487/1)

Acknowledgements:
The paper makes use of data from Understanding Society, an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by the National Centre for Social Research and TNS BMRB. The research data are distributed by the UK Data Service.

Abstract
This paper provides new insights on the effects of retirement on various subjective and objective health outcomes. We look at both short- and long-term effects of disengagement from paid work, by accounting for the distance from retirement. We use comprehensive longitudinal data for UK women close to retirement age, using a sample of 34,722 observations for 7,684 women. Our identification comes from an instrumental variable panel estimation methods, and exploits a recent substantial reform of female pension rules which increased incentives to stay at work longer. We explore multiple mechanisms that can explain the retirement-health link, and explore heterogeneity by levels of job-strain. Preliminary findings indicate a positive causal impact of retirement on mental health (lower depression) as measured by the GHQ and MCS indices and their sub-components. Crucially, the effect is driven by women who were employed in jobs with high levels of physical or psychosocial burden.

1. Introduction

As ageing is one of the great contemporary social and economic challenges for most Welfare Systems in developed economies, most OECD countries have promoted reforms to increase the labor market engagement of older people, enhance sustainability of pension systems and promote healthy ageing. In this light, it is thus crucial to understand what is the net effect of retirement on health. Theoretically, retirement can lead to an improvement in health, through reduced stress and increased enjoyment of life, or to negative effects, e.g., due to a reduced sense of purpose and the loss of social interactions (Hessel, 2016; Mazzonna and Peracchi, 2017; Zhu, 2016). Empirical evidence is ambiguous, with studies highlighting positive effects on mental (Belloni *et al.*, 2016) or physical health (Bertoni *et al.*, 2017), or negative effects (Behncke, 2012; Bonsang *et al.*, 2012). Findings are sensitive to the choice of country, empirical strategy and health outcome (Avendano and Berkman (2014)). The endogenous relation between retirement and health means that the causal impact of retirement on health cannot be easily established empirically (Zhu, 2016).

In this paper, we provide new evidence on the effect of retirement for women living in the UK.
We contribute to the existing literature in several ways. First, we differentiate from other studies who exploit only cross-country variation in pension eligibility ages, and exploit a major change in rules for UK women, which substantially altered the incentives to retire raising the State Pension Age from 60 to 65 between 2010 and 2018. We implement a panel instrumental variable approach (instrumenting the retirement status and retirement duration with pension eligibility status and the duration of the eligibility status), for a powerful identification of the effect of retirement.

Second, we investigate both the short- and the long-term effects of retirement. Indeed, retirement may imply a positive or negative health shock in the initial period, e.g., the so called “honeymoon phase” where individuals engage in different activities that were previously foregone because of work-leisure conflicts. On the other hand, it may take time for people to adjust to a new lifestyle after disengaging from the labor market, and such adjustments may translate progressively into health changes, so the effects may not be instantaneous (Bonsang *et al.*, 2012). Moreover, retirement can be viewed as a cumulative process of exposure to being out of the labor force (see the theoretical model in the work of Mazzonna and Peracchi (2017)), so that health deterioration rates may depend on the duration of retirement, which we estimate by the distance from individuals’ age at interview and their age when left their last job.
Third, we are able to assess a wide range of health measures, from self-reported health (including quality of sleep), to reports of diagnosed conditions, to biomarkers. These measures are complementary in our view. Subjective measures have strong power in predicting morbidity, but may be subject to misreporting; diagnosed conditions are more objective, but may be biased as some conditions may be undiagnosed; biomarkers are free from contamination, but sensitive to specific dimensions of health, and often of limited availability (Behncke, 2012).
Fourth, we explore several mechanisms and heterogeneity effects that can help disentangling multiple factors connecting health and employment status. First, it has been shown that the retirement may be more protective of people having lower socioeconomic status and burdensome jobs (Belloni *et al.*, 2016; Bertoni *et al.*, 2017; Coe *et al.*, 2012; Mazzonna and Peracchi, 2017; Westerlund *et al.*, 2009). We exploit an externally validated index of job-strain (involving physical and psychosocial factors) to investigate how the health patterns after work-disengagement differ across job-type (Santi *et al.*, 2013). Such index is based on the four digits of the ISCO occupational classification, and allows for a better understanding of the impact of job-exposure than the standard blue- vs white collar distinctions. Furthermore, the theoretical and empirical literature has shown that several health behaviors may be affected by retirement, which are well known to affect health and the onset of diseases. Examples are participation in physical activity, smoking and drinking, which we are able to analyze in our study (Zhu, 2016).
Our preliminary results indicate that retirement has a consistent beneficiary effect on mental health, which is driven primarily by individuals who disengaged from a demanding occupation.

2. Methods

Empirical strategy: a panel fixed effects with instrumental variable

We define a respondent as retired if she is disengaged from the labour market (out of the labour force) based on self-reported economic activity status (see, e.g., Zhu (2016), Mazzonna and Peracchi (2017), Belloni *et al.* (2016)).
To estimate the effect of retirement on health, we start with a basic model to be estimated through OLS:

$$H_{it} = \beta_0 + \beta_1 DRET_{it} + X'_{it}\beta_2 + U_{it} \quad (1)$$

where y_{it} is the health measure for individual i at time t, $DRET_{it}$ is a dummy variable for being out of the labour force at time t, X_{it} includes controls (described below). However, OLS estimates may be biased due to potential reverse causality (people with poor health may decide to retire earlier) or correlation between the retirement choice and unobservable factors included in the regression error (Mazzonna and Peracchi, 2017). We solve this important empirical issue in two steps. First, we add individual fixed-effects d_i, to control for unobservable time-invariant characteristics affecting both the retirement decision and health. Second, we use instrumental variables technique to control for time-varying unobservables and reverse causality. In particular, we instrument retirement decision with eligibility to State Pension, determined by a combination of respondents’ year-and-month of birth, and date-of-interview (both available in the data). To be valid, an instrument has to be sufficiently correlated with the retirement variable, and it has to be orthogonal to the idiosyncratic error term ν_{it}. Both requirements have already been shown to hold for this instrument: there is a strong existing evidence that pension eligibility is related to the retirement decision (retirement incentives increase significantly when reaching the SPA, leading to a discontinuous jump in the probability of being retired), and correlated to health only through the effect of retirement (Coe and Zamarro, 2011). Moreover,

we can exploit a recent reform which, since 2010, gradually increased female SPA from 60 years old to 65 (in 2018), depending on year-and-month of birth, for women born after March 1950.[1] The fact that the SPA is not set at 60 for all women strengthens the identification of our analysis, as we can effectively disentangle the incentive to retire due to crossing the SPA, from the incentive possibly coming from turning a specific age (e.g., 60), after controlling for age (see below).
We estimate the following model, using the two-stage least squares FE estimator

$$H_{it} = \beta_1 \text{DRET}_{it} + X'_{it}\beta_2 + d_i + \nu_{it} \quad (2)$$

Where DRET is instrumented with a dummy for being below Statutory Pension Age.

Controls

The vector X includes standard demographic controls as a second-order polynomial for age, fixed effects for marital status (married/living in couple, single, divorced/widowed), and for having zero, one/two, or three or more children. We also control for socio-demographic characteristics such as education (A-level or higher, GCSE level or no education), the hierarchical Socioeconomic Classification (NS-SEC) of the current/last job (routine, intermediate, managerial), and an indicator for home-ownership. We further include fixed effects for interview year, and country of residence (within the UK). Adding such set of controls strengthens both the conditional independence and the exclusion restriction hypothesis on which the IV-estimation is based, as well as the efficiency of the estimates (Angrist and Pischke, 2009).

Retirement duration

In a subsequent specification, we investigate both the short- and the long-run effects of retirement by including in (2) a control for the years spent in retirement. Retirement duration is defined as the distance between interview date and the date when the last job ended for those who are retired, and is set at zero for people still in the labor market. Being this duration variable endogenous as well, we instrument it with the duration of being eligible for the state pension, at the time of interview, which takes value zero for those not retired (Zhu, 2016).

Data and main health measures

We use data from the *Understanding Society* survey, a longitudinal dataset collected between 2009 and 2016 (7 waves) on a sample of household members aged 16+ in the UK (see Lynn (2009) for technical details) It includes rich information about mental and physical health, employment status, financial situation, activities and birthdate of the respondents. We restrict our sample to respondents aged 50 to 70 years old. As we use an instrumental variables technique based on statutory pensionable age (SPA), we restrict the sample to women only, to exploit a recent reform which substantially altered the female SPA. We drop respondents who never worked, proxy interviews, as well as those who reported returning to work during the considered time interval (3,061 obs.), as we cannot assume that stopping or re-starting paid-work leads to symmetric effects on health (Bonsang *et al.*, 2012). Our main sample consists of 34,722 observations (7,684 women).

We evaluate the causal effect of retirement on the General Health Questionnaire (GHQ) and the Short Form-12 scores, which are widely adopted measures of mental and physical health (e.g., Mitra and Jones (2017), Dustmann and Fasani (2016)). The GHQ measures psychological distress through 12 items, each scored with a four-steps Likert scale, which are summed in a final index (between 0 and 36), with higher values signalling worse health (Goldberg *et al.*, 1997). We also exploit the disaggregation in three clinically meaningful factors (anxiety and depression, social dysfunction and loss of confidence) proposed by Graetz (1991). Following Dustmann and Fasani (2016), we normalise both the GHQ and the factors scores between zero and 100 (higher scores, worse health). We also create dummy variables based on validated cutoffs for

[1] The rate of increase is nonlinear in birth-date: for women born between 6th April 1950 and 6th April 1953, or between December 1953 and October 1954, SPA increases by one month for those born after the sixth day of each month. For women born between 6th April and 5th December 1953, SPA rises by three months for those born after the sixth day of each month.

the presence of depressive disorders, corresponding to a 12+ score in the 0-36 GHQ scale and to a 3+ score in the 0-12 GHQ scale (Goldberg *et al.* (1997); Makowska *et al.* (2002)).
The Short Form-12 (SF-12, version 2) is a generic health-related instrument which comprises 12 items (Ware, 2002). Each item is evaluated with either 1-3 or 1-5 Likert scales, then aggregated in factors, which are in turn aggregated in a summary physical (PCS) and a mental (MCS) score, each ranging from 0 to 100 (with a mean of 50). Higher values signal better health (see Ware (2002) for further details).

We also exploit information on whether respondents have been diagnosed (by a doctor) with a chronic condition since the previous interview, or before joining the survey. Health conditions include arthritis; cardiovascular disease; endocrine disease, liver disease; high blood pressure; clinical depression (see Davillas and Pudney (2017) for details). We thus evaluate whether retirement (and its duration) impact the likelihood of new diagnoses, for those respondents who were free from the condition at baseline.

We also test the impact of retirement on hours and quality of sleep (available for waves 1,4,7).

To complement the analysis based on previous measures, we run a separate analysis on biomarkers data based on non-fasted blood sample, which have been collected within Understanding Society by trained nurses, during waves 2 and 3 (Benzeval *et al.*, 2014), for 13,107 respondents. Biomarkers include indicators for inflammation (C-reactive protein fibrinogen) fatty substances in the blood (the ratio of total cholesterol to high-density lipoprotein cholesterol; and triglycerides) and level of sugar in the blood (Glycated haemoglobin). When applying the selection as described above, our biomarker sample includes 2,224 women (cross-section).

Analysis of mechanisms and heterogeneity
We investigate whether the health effect of retirement differs across job type, using a novel measure of job-strain described in Santi *et al.* (2013) which we match to 99% of our sample through the ISCO employment code. The exposure measure is a combination of 5 dimensions of occupational burdens: Ergonomic Stress, Environmental Pollution, Mental Stress, Social Stress and Temporal Loads. The index ranges from 1 (low exposure) to 10 (high exposure), has been externally validated and more informative than the widely used distinctions based on the assumed skill level from the first digit of the ISCO-code (Santi *et al.*, 2013). Following Santi *et al.* (2013) and Mazzonna and Peracchi (2017), we identify as "high exposure" those occupations with a score of 6 or higher (40% of our sample). We thus add in interact the retirement dummy in model (2) with the exposure dummy (this interaction is instrumented with another interaction term between the pension eligibility dummy and the exposure dummy).
We further explore the causal effect of retirement on health-related behaviors, namely the participation and intensity of smoking, and alcohol drinking, as well as of exercising (for waves 2, 5 and 7), which are potential mechanisms for the observed retirement-health effects.

3. Preliminary Results

Preliminary results for GHQ and SF-12 outcomes are shown in Table 1, where each row refers to a separate regression. Column (i) shows the sample mean for each dependent variable (health outcome). Column (ii) reports the coefficients for the impact of being retired when neglecting endogeneity (OLS estimation, model (1)), while column (ii) refers to the fixed effects 2SLS model in (2). Column (iii) reports the standard Durbin Wu Hausman test for exogeneity of retirement in the 2SLS model. At the bottom, we report the F-stat for the instrument's strength test.
OLS coefficients highlight that there is a negative statistical association between retirement and health, across all the considered measures of mental and physical health. However, when causality is addressed (the fixed-effects IV model) results point to a much different effect: retirement reduces the GHQ depression scores, improves mental health as measured by the MCS and several SF12 factors, while no significant result is found for physical health. All the diagnostic tests suggest that the 2SLS model is the best suited to capture a casual effect. The first stage regression shows that pension eligibility is a strong predictor of labor market

disengagement (First stage F statistics = 122, full results in Table 2), and the exogeneity of the retirement variable is almost always significantly rejected.
Albeit further analysis will be provided for sleep, chronic conditions and biomarkers variables, we already have a first evidence on the mechanism driving these results. Table 3 reports the heterogeneous health-effect of retirement between women in high- and low-strain jobs, which are much stronger for the former group: indeed, only women in demanding jobs have a statistically significant reduction in GHQ depression scores, and even in the risk of depression (the GHQ cutoff).

Table 1, Impact of being out of the labor force on health, OLS and IV-Fixed-effects models

		Indep. variable: being retired		
	(i) Population mean	(ii) OLS	(iii) FE-IV 2SLS	(iv) Exogeneity DWH test: p-value
GHQ (higher score, worse health)				
GHQ overall score	31.6	2.893***	-6.323**	0.017
		(3.482 - 2.305)	(-10.558 - -2.088)	
GHQ cutoff 3+ (0-12 scale)	0.24	0.083***	-0.114	0.15
		(0.094 - 0.062)	(-0.233 - -0.006)	
GHQ – anxiety	29.6	2.270***	-10.880**	0.019
		(3.041 - 1.500)	(-17.836 - -3.924)	
GHQ – social dysfunction	37	3.118***	-2.523	0.298
		(3.617 - 2.619)	(-6.366 - 1.320)	
GHQ – Loss of confidence	19.3	4.622***	-8.609**	0.003
		(5.457 - 3.787)	(-14.472 - -2.747)	
SF-12 (higher score, better health)				
MCS mental component score	48	-1.482***	5.124***	0.002
		(-1.136 - -1.828)	(2.487 - 7.760)	
PCS physical component score	50	-3.173***	1.7	0.065
		(-2.810 - -3.537)	(-1.615 - 5.016)	
Physical functioning	48.8	-2.804***	2.526	0.019
		(-2.429 - -3.180)	(-0.344 - 5.395)	
Role physical	49.3	-3.562***	4.015*	0.003
		(-3.188 - -3.935)	(0.527 - 7.504)	
Bodily pain	47.9	-3.646***	1.132	0.228
		(-3.247 - -4.045)	(-2.475 - 4.738)	
General health	46.7	-3.066***	2.355	0.003
		(-2.665 - -3.466)	(-0.972 - 5.682)	
Vitality	49.9	-1.659***	4.566***	0.003
		(-1.328 - -1.990)	(2.006 - 7.125)	
Social functioning	49.9	-3.441***	4.504*	0.003
		(-3.047 - -3.836)	(0.916 - 8.092)	
Role emotional	49.5	-3.068***	4.052**	0.004
		(-2.661 - -3.475)	(1.215 - 6.889)	
Mental health	49.3	-1.192***	4.739**	0.010
		(-0.844 - -1.541)	(1.860 - 7.619)	
N	34,722	34,722	34,722	
First stage F statistics (instrument: being below SPA)	-	-	F(1,307)=122.04***	-

*Note: we report OLS coefficients in column (ii), Panel Fixed-Effects Instrumental Variables 2SLS coefficient in column (iii), and the p-value for the Durbin–Wu–Hausman (DWH) exogeneity test for the retirement variable in column (iv). Each row corresponds to one regression model. P-values correspondence: *<0.05,**<0.01,***<0.001. We report 95% confidence intervals (in brackets), clustered by year-and-month-of-birth (308 clusters).*
Sample selection: women aged 50-69 between 2009 and 2016, having been engaged in paid work in their life, excluding proxy respondents and individuals who reported returning to work during the sampling period. Additional controls include a 2nd order polynomial for age, fixed effects for interview year, living arrangements (married,

widowed/divorced/separated, single), country, number of children (none, one-two, three or more), education (low, mid or high degree), Socio Economic classification (routine, intermediate, management), and home-ownership.

Table 2, first-stage for the probability of being retired

Dep variable: being retired	Coefficient	Confidence interval
being below SPA	-0.162***	(-0.187 - -0.137)
routine SES (ref.)		
Intermediate SES	0.016	(-0.008 - 0.040)
Manager SES	0.01	(-0.014 - 0.033)
Marital status (ref. “married/couple”)		
Widowed/divorced	-0.013	(-0.042 - 0.016)
Single	-0.053*	(-0.101 - -0.006)
no children (ref.)		
1-2 children	-0.076**	(-0.132 - -0.020)
3+ children	-0.156**	(-0.272 - -0.041)
No education (ref.)		
low education	0.012	(-0.053 - 0.076)
mid or high education	-0.068*	(-0.133 - -0.003)
Owning dwelling	0.018	(-0.031 - 0.066)
Age (years)	-0.100***	(-0.130 - -0.071)
Age squared	0.001***	(0.001 - 0.001)
constant	2.763***	(1.467 - 4.060)
N	34,722	

*Note: Linear Probability Model coefficients for the probability of being retired. The status of being above/below State Pension Age (SPA) is defined by comparing the individual SPA (based on month-year of birth) and the date of interview. Additional controls include country dummies (for the UK), and dummies for year of interview. P-values correspondence: *<0.05,**<0.01,***<0.001. Sample selection as in Table 1.*

Table 3, health effect of retirement, heterogeneity by job-type

		Indep. variable: being out of labour force	
Panel IV Fixed Effects	(i) Population mean	(ii) Low exposure job	(iii) High exposure job
GHQ overall score	31.6	-4.491 *(-9.1 – 0.122)*	-10.02*** *(-15.14 - -4.9)*
GHQ cutoff 3+ (0-12 scale)	0.24	*-0.095* *(-0.228 – 0.038)*	*-0.153*** *(-0.293 - -0.012)*
MCS mental component score	48	*4.766**** *(2.195 – 7.336)*	*5.433**** *(1.913 – 8.953)*
PCS physical component score	50	*1.083* *(-2.464 – 4.631)*	*2.301* *(-1.32 – 5.923)*
N = 34,411			
First stage F statistics F(2,307)=169.77***			

Note: we report Panel Fixed-Effects Instrumental Variables 2SLS coefficients. Each row corresponds to one regression model. In each regression, the labour force status is interacted with a dummy for high-strain job. Job exposure is defined as having a score of 6+ in the exposure index associated to the respondent’s current (or last) job ISCO code

*(Santi et al., 2013). P-values correspondence: *<0.05, **<0.01, ***<0.001. We report 95% confidence intervals, clustered by year-and-month-of-birth (308 clusters). Sample selection as in Table 1.*

References

Angrist, J. D. and J.-S. Pischke (2009). 'Mostly harmless econometrics', *An Empiricist's Companion. Princeton*.

Avendano, M. and L. F. Berkman (2014). 'Labor markets, employment policies, and health', *Social epidemiology*, pp. 182-233.

Behncke, S. (2012). 'Does retirement trigger ill health?', *Health economics*, vol. **21**(**3**), pp. 282-300.

Belloni, M., E. Meschi and G. Pasini (2016). 'The effect on mental health of retiring during the economic crisis', *Health economics*, vol. **25**(**S2**), pp. 126-140.

Benzeval, M., A. Davillas, M. Kumari and P. Lynn (2014). 'Understanding society: the UK household longitudinal study biomarker user guide and glossary', *Institute for Social and Economic Research, University of Essex*.

Bertoni, M., S. Maggi and G. Weber (2017). 'Work, retirement, and muscle strength loss in old age', *Health economics*, pp. n/a-n/a.

Bonsang, E., S. Adam and S. Perelman (2012). 'Does retirement affect cognitive functioning?', *Journal of health economics*, vol. **31**(**3**), pp. 490-501.

Coe, N. B., H.-M. von Gaudecker, M. Lindeboom and J. Maurer (2012). 'The effect of retirement on cognitive functioning', *Health economics*, vol. **21**(**8**), pp. 913-927.

Coe, N. B. and G. Zamarro (2011). 'Retirement effects on health in Europe', *Journal of health economics*, vol. **30**(**1**), pp. 77-86.

Davillas, A. and S. Pudney (2017). 'Concordance of health states in couples: Analysis of self-reported, nurse administered and blood-based biomarker data in the UK Understanding Society panel', *Journal of health economics*, vol. **56**, pp. 87-102.

Dustmann, C. and F. Fasani (2016). 'The Effect of Local Area Crime on Mental Health', *The Economic Journal*, vol. **126**(**593**), pp. 978-1017.

Goldberg, D. P., R. Gater, N. Sartorius, T. B. Ustun, M. Piccinelli, O. Gureje and C. Rutter (1997). 'The validity of two versions of the GHQ in the WHO study of mental illness in general health care', *Psychological medicine*, vol. **27**(**1**), pp. 191-197.

Graetz, B. (1991). 'Multidimensional properties of the general health questionnaire', *Social psychiatry and psychiatric epidemiology*, vol. **26**(**3**), pp. 132-138.

Hessel, P. (2016). 'Does retirement (really) lead to worse health among European men and women across all educational levels?', *Social Science & Medicine*, vol. **151**, pp. 19-26.

Lynn, P. (2009). 'Sample design for understanding society', *Underst. Soc. Work. Pap. Ser*, vol. **2009**.

Makowska, Z., D. Merecz and W. Kolasa (2002). 'The validity of general health questionnaires, GHQ-12 and GHQ-28, in mental health studies of working people', *International journal of occupational medicine and environmental health*, vol. **15**(**4**), pp. 353-362.

Mazzonna, F. and F. Peracchi (2017). 'Unhealthy retirement?', *Journal of Human Resources*, vol. **52**(**1**), pp. 128-151.

Mitra, S. and K. Jones (2017). 'The impact of recent mental health changes on employment: new evidence from longitudinal data', *Applied Economics*, vol. **49**(**1**), pp. 96-109.

Santi, I., L. E. Kroll, A. Dietz, H. Becher and H. Ramroth (2013). 'Occupation and educational inequalities in laryngeal cancer: the use of a job index', *BMC public health*, vol. **13**(**1**), pp. 1080.

Ware, J. E. (2002). *The SF-12v2TM how to score version 2 of the SF-12® health survey:(with a supplement documenting version 1)*: Quality Metric.

Westerlund, H., M. Kivimäki, A. Singh-Manoux, M. Melchior, J. E. Ferrie, J. Pentti, M. Jokela, C. Leineweber, M. Goldberg and M. Zins (2009). 'Self-rated health before and after retirement in France (GAZEL): a cohort study', *The Lancet*, vol. **374**(**9705**), pp. 1889-1896.

Zhu, R. (2016). 'Retirement and its consequences for women's health in Australia', *Social Science & Medicine*, vol. **163**, pp. 117-125.