The end of the grind or the beginning of the end?

Inequalities in the short- and medium-term impact of retirement on women's mental health in the UK

Laura Sochas ; Ludovico Carrino ; Mauricio Avendano

INTRODUCTION

In the face of population ageing, high-income countries are concerned about increased dependency ratios, declining tax resources, and their capacity to sustain healthcare and public pension schemes. In response, many OECD countries have implemented reforms aiming to incentivise people to work longer, often by increasing the age at which citizens become eligible for state pension benefits.

While such reforms have been shown to delay retirement (OECD 2017), the potential health impacts of retirement are still misunderstood. Theoretically, the impact of retirement on health, both mental and physical, could be positive or negative (Hessel, 2016; Mazzonna and Peracchi, 2015). On the one hand, retirement may improve health by lowering psychosocial and physical stress caused by work and freeing up time for healthy activities such as cooking or exercise, or meaningful activities such as volunteering, learning a new skill, or spending time with family. Conversely, retirement could have negative effects on health by separating retirees from an important social network, reducing market incentives to invest in health-enhancing activities and intellectually engaging tasks, lowering their sense of fulfilment, and decreasing their levels of exercise. Retirement could also increase the leisure time that people can choose to invest in harmful activities, such as drinking or smoking. Crucially, theory predicts that these effects of retirement may depend on retirement duration (Bonsang et al., 2012; Coe et al., 2012; Mazzonna and Peracchi, 2015, 2012). Work-disengagement may produce an immediate positive health effect, e.g., a reduction in depression, due to the changed environment at the time of retirement (the *honeymoon phase* theorised by Atchley (1982)). Subsequently, the reduction in health-investment would progressively and cumulatively translate to an increased rate of health deterioration post-retirement (the use-it-or-lose-it hypothesis, see Rohwedder and Willis (2010)).

Many of these pathways are likely to be sensitive to socio-economic status. For example, the benefits of retirement may depend on the extent to which a person's work was taxing for health, either physically or psycho-socially. People who had jobs that were not socially supportive, fulfilling, or cognitively challenging wouldn't have work-related health benefits to lose from exiting the labour force. One's pension wealth could also affect the extent to which the retirement period could be used for fulfilling activities, or whether poverty in retirement and its psycho-social and material effects would be a concern.

From a policy perspective, it is particularly interesting to investigate the impact of retirement within the female population, since women are those most impacted by pension reforms, with employment growth in the 55-65 female age group exceeding 40% since 2000 (OECD, 2019). In addition, older women are more likely to conduct unpaid caring activities for spouses and grandchildren, the burden of which could also affect their health post-retirement.

The existing literature on this topic reports mixed results for both men and women. Being retired improves a wide range of health measures according to some studies (Atalay and Barrett, 2014; Belloni et al., 2016; Bloemen et al., 2017; Coe and Zamarro, 2011; Hessel, 2016) but has negative effects in others (Behncke, 2012; Mazzonna and Peracchi, 2012). This contradiction is very likely explained by the fact that these papers do not explicitly model both the effect of the short-term transition to retirement *and* the cumulative effect of years retired. Those that do typically find a short-term improvement in physical or cognitive health, followed by an acceleration of the negative age-health trend compared to those who are not retired (Bertoni et al., 2018; Celidoni et al., 2017; Mazzonna and Peracchi, 2015; Westerlund et al., 2009).

There are currently no studies jointly modelling the short and medium-term impact of retirement for mental health outcomes. Papers that investigate the impact of retirement on mental health typically look at the average effect of being retired, and find a positive effect (Atalay and Barrett, 2014; Belloni et al., 2016; Mein et al., 2003; Zhu, 2016), or no effect (Coe and Zamarro, 2011). Johnston and Lee (2009) find that the effect of the initial transition to retirement on mental health is positive, while Zhu (Zhu, 2016) finds that retirement duration confers additional mental health benefits. If the effect of retirement on mental health follows the same pattern as for physical health and cognition, the average positive effect of being retired on mental health (relative to those who are not retired) might mask an initial positive improvement, followed by a reversal over time.

In terms of heterogeneity in the effect of retirement on mental health, Atalay and Barrett (Atalay and Barrett, 2014) find stronger positive effects for the lower half of the income distribution in women, as do Belloni et al (Belloni et al., 2016) for male blue-collar workers exposed to the 2008 economic crisis. In terms of physical health and cognition, the greater number of papers investigating heterogeneity reach a range of conclusions, with some finding more positive effects of retirement for people of high socio-economic status (Bertoni et al., 2018; Hessel, 2016; Mazzonna and Peracchi, 2012) while others find the reverse (Hernaes et al., 2013; Mazzonna and Peracchi, 2015; Westerlund et al., 2009). Atalay and Barrett (2014) find no evidence of heterogeneity for most outcomes in women and all outcomes in men, and neither do Xue et al (2018). Most papers investigate heterogeneity by repeating the analysis in stratified samples, but do not directly test the significance of differences in sub-groups using interaction terms or post-estimation tests, except for Bertoni et al (2018) and Belloni et al (2016).

This study proposes to extend the existing literature by analysing, for the first time, the short and medium-term impact of retirement on validated mental health outcomes, and by testing for heterogeneous effects by socio-economic status. Moreover, we are the first to investigate these effects in the UK, after the implementation of the 2010 reform that increased the female State Pension Age (SPA). We do so by exploiting recent longitudinal data from the Understanding Society survey (2009-2017). Previous evidence analysing both the short and medium-term health effects of retirement is based on European (SHARE), American (HRS), and French (GAZEL) longitudinal surveys.

METHODS

Data and sample selection

This study employs data from seven waves of Understanding Society, a large, nationally representative, longitudinal survey of UK individuals (Lynn, 2009), covering a time-period from 2009 to 2017. The sample selection includes female individuals, aged 50-70 at the time of the interview, first observed in the labour force.[1] As our focus is on the transition from work to retirement, we exclude respondents who report being inactive (home-caring or sick) or who go back from

[1] As the male SPA was fixed at 65 in the survey period, we exclude men from the analysis due to absence of credible exogenous variation in retirement incentives.

[2] The asinh is comparable to a log-transformation, with the advantage of being defined at 0 such that there is no need to add +1 year to those who spent zero years in retirement.

retirement to paid-work during the survey period (Belloni et al., 2016; Bonsang et al., 2012; Celidoni et al., 2017). After dropping proxy- or partial interviews and any observations with missing data on outcome or control variables, our sample comprises 27,498 person-years and 6,391 individuals, of which 1,158 retire over the period of observation. On average, individuals in our sample who retire over the period of observation spend 2.9 years in retirement. 25% of those who retire are observed for 4.5 years or more, while the maximum length of time spent in retirement is 7 years.

Variables

Two indices are used to measure mental health outcomes. Depression is measured using the GHQ-12 index, scored from 0 to 36 using 12 items, each measured 1-3 using a Likert scale. The GHQ items measure the appearance of new and distressing psychological phenomena, as well as an inability to carry out normal functions (D.P. et al., 1997; Goldberg and Williams, 1988). We treat the index as a continuous variable, and normalise it from 0 to 100, with higher values signalling worse health (greater depression).

We also use a generic measure of mental health, the SF-12 Mental Health Composite Score (MCS). The Short Form-12 (SF-12, version 2) is a generic health-related quality of life instrument which comprises 12 items from eight health concepts (Ware, 2002). Four of these concepts relate to mental health: role limitations due to emotional problems (2 items), psychological distress, psychological wellbeing, and social functioning. The items are evaluated with either 1-3 or 1-5 Likert scales, and then aggregated into a mental (MCS) component score, ranging from 0 to 100, with higher values signalling better health.

As a comparison, we also estimate the effect of retirement on physical health, which is measured using the SF-12 Physical Health Composite Score (PCS), from 0 (bad) to 100 (good health). This score captures physical functioning, role limitations due to physical health issues, bodily pain, and vitality (Ware et al., 1994, 1996).

We define retirement status as an indicator variable taking value 1 if the respondent self-reports being retired. We generate a continuous variable for years in retirement, defined as the number of years since the respondent first reported being retired (the variable equals 0 prior to retirement). In order to allow for a non-linear relationship between retirement duration and health outcomes, we adopt the inverse hyperbolic sine (asinh) transformation of years retired in all specifications, and also include the asinh[2] of age at the year-month level as a control (Bertoni et al., 2018; Mazzonna and Peracchi, 2015). Examination of the Akaike Information Criterion and the Bayesian Information Criterion confirms that the non-linear specification of age has a superior fit to the linear one, which we test as a robustness check (results available upon request).

Conditional on covariates (discussed below), and given our sample selection, which is comprised solely of individuals observed in paid-work when the survey started, the coefficient on age measures the health-age trend independently from retirement, the coefficient on the dummy for being retired measures the step-change in health at the point of the transition into retirement, and the coefficient on years retired quantifies the change in the health-age trend after retirement (Lagarde, 2012; Mazzonna and Peracchi, 2015).

---

[2] The asinh is comparable to a log-transformation, with the advantage of being defined at 0 such that there is no need to add +1 year to those who spent zero years in retirement.

The following controls are also included: dummies for interview year, hierarchical job-category (routine (ref), intermediate or managerial)[3], highest educational attainment (no qualifications (ref), GCSE or below, A levels and above), number of children (no children (ref), 1-2 children, 3+ children), a dummy for home ownership, and a dummies for country (England (ref), Wales, Scotland, Northern Ireland). Baseline income is not controlled for, because it is likely affected by retirement. However, the results are robust to including income as an additional control (CPI adjusted; logged or quantile; household or individual) (results available upon request).

Identification strategy

Identifying the causal impact of retirement is complicated by several factors. First, there could be reverse causation, as people may retire because of health problems. Second, there may also be a selection bias whereby people who work in jobs that are harmful to their health may not be able to afford an early retirement. Finally, retirement and health problems both become more likely as people age (Coe and Zamarro, 2011).

In order to resolve these identification issues, we use an instrumental variable approach in order to identify the effect of retirement on health based on an exogenous source of variation in the treatment (retirement). Specifically, we employ the individual's pension eligibility status as the instrument for being retired (i.e.: whether or not someone has reached State Pension Age (SPA)), and the number of years an individual has exceeded SPA (which is 0 if the individual is below SPA) as an instrument for years in retirement (Battistin et al., 2009; Mazzonna and Peracchi, 2015). This approach identifies the Local Average Treatment Effect, which is the effect of retirement on health for those who retired when they reached State Pension Age. In order to consistently estimate the Local Average Treatment Effect, the instrument must (1) be associated with retirement, (2) it cannot be the case that people return to work from retirement upon reaching the State Pension Age, (3) the instrument must be as good as randomly assigned, and (4) the instrument must only be associated with health through its association with retirement. Proof of the first condition is given by the first-stage diagnostic tests, which show that the instrument is highly relevant (Cf: the Cragg-Donald statistic and the Kleibergen-Paap Wald F-test in Table 2). Moreover, a large stream of literature has shown that pension eligibility constitutes an important incentive to retire in most Western economies (Rohwedder and Willis, 2010).The second condition is met due to the sample selection having eliminated any individuals going from retirement to work. Although conditions (3) and (4) cannot be proven, it is very likely that they hold in this case because reaching the State Pension Age provokes a sharp discontinuity in the probability of being retired conditional on age (see Appendix Figure 1), and because people cannot strategically position themselves across this cut-off. While age is also related to health, it is unlikely that such a sharp discontinuity is present in the age-health function (Coe and Zamarro, 2011).

Three additional precautions are taken to strengthen this assumption. First, we control for the asinh function of age, measured at the year-month level, to ensure that the instrument captures the effect of the discontinuity only, and not the underlying age trend. Second, we include individual fixed effects, which control for all time invariant individual characteristics that may result, for example, in a link between being of state pension age and experiencing a sharp discontinuity in health (e.g. being of a generation exposed to catastrophic historical events such as the Second World War (Kesternich et al., 2014)), or in non-random attrition rates. Third, we exploit a reform in the UK that

[3] Job-category corresponds to the hierarchical Standard Occupational Classification (SOC 2000), and is based on current/last occupational class, firm-size, and employment status employer/self-employed/employee).

progressively increased the SPA from age 60 for women in 2010 to age 65 in 2018 (and eventually to age 66 by 2020, see Thurley (2017)). For the individuals in our dataset who have reached State Pension Age, there is a range of SPAs from 60 to 63, often differing by a matter of months. For example, a woman born in March 1950 (or earlier) could retire at age 60, while a woman born in January 1952 could only retire at age 62. This exogenous change in SPA ensures that the discontinuity in health at the time of reaching the State Pension Age is not caused by the relationship between reaching a specific age and health, since people of same age face different incentives to retire, depending on how they are affected by the reform.

Specification

(1)	$y_{it} = \alpha + \beta_1 * \widehat{retired}_{it} + \beta_2 * \text{asinh}\,(\widehat{yearsret}_{it}) + \beta_3 * \text{asinh}\,(age_{it}) + \delta_i + \theta_t + \gamma\,\Phi_{it} + \varepsilon_{it}$

Our analysis uses the above baseline specification, where *y* is a continuous health outcome: GHQ (depression) or MCS (mental health); *retired* and *yearsret* are the independent variables of interest, instrumented with being below the State Pension Age and years since reaching the State Pension Age; *age* is measured at the year-month level, δ is an individual fixed effect, θ are dummies for interview year, and ϕ is a vector of other time-varying controls, including type of job, number of children, educational attainment, marital status, home ownership, and country of residence. We estimate robust standard errors clustered at the month-year of birth, which is the source of variation in the treatment variable caused by pension eligibility rules.

RESULTS

Overall sample

Table 1 presents descriptive statistics of the outcome and control variables, according to whether the observations are for a retired respondent or not. 18.7% of sampled individuals retire over the study period. Those who are retired have better mental health but worse physical health than those who are not retired. Many control variables, such as marital status, number of children, and educational attainment are also significantly different between retired and non-retired observations, likely due to cohort and/or age effects.

Table 1: Mean and frequencies of outcome and control variables in retired vs not retired

	retired = 0 N=23,779	**retired = 1** N=3,719	**p-value**
Age in years	56 (5)	64 (4)	<0.001
Marital status			
Married	15,040 (63.2%)	2,452 (65.9%)	0.002
Single	2,253 (9.5%)	233 (6.3%)	<0.001
Divorced or separated	6,486 (27.3%)	1,034 (27.8%)	0.5
Number of children	1.61 (1.33)	1.56 (1.27)	0.036
Home ownership	20,038 (84.3%)	3,363 (90.4%)	<0.001

Max educational attainment			
No qualifications	1,819 (7.6%)	436 (11.7%)	<0.001
GCSE or other lower qual	7,848 (33.0%)	1,248 (33.6%)	0.5
Above GCSE	14,112 (59.3%)	2,035 (54.7%)	<0.001
Category of last job			
Routine	7,504 (31.6%)	1,167 (31.4%)	0.83
Intermediate	9,875 (41.5%)	1,570 (42.2%)	0.43
Management	6,400 (26.9%)	982 (26.4%)	0.51
Household income at baseline[4]			
Low income (£395-£2,412)	5,348 (23.1%)	963 (26.5%)	<0.001
Medium income (£2,413-£3,897)	7,535 (32.5%)	1,241 (34.1%)	0.053
High income (£3,901-£11,300)	10,288 (44.4%)	1,431 (39.4%)	<0.001
Health outcomes			
Depression: GHQ score 0-100	30.75 (14.07)	28.19 (13.57)	<0.001
Mental health: MCS score 0-100	50.21 (9.01)	52.32 (8.61)	<0.001
Physical health: PCS score 0-100	51.22 (9.09)	48.23 (10.58)	<0.001

Note: % mean or standard deviation in parentheses

Table 2 displays results for model (1) for all health outcomes. We report the coefficients of interest for both the fixed effects regressions and the fixed effects-instrumental variable regressions. According to the results from the fixed effects regression, the transition to retirement has a short-term positive effect on depression and mental health, followed by a progressive worsening according to retirement duration. The effect of the short-term transition to retirement on physical health, in contrast, appears to be negative, with no significant effects for retirement duration.

In order to better identify the causal relationship between retirement and health, we turn to the results from the fixed effects-instrumental variable regressions (FE-IV). Our two instruments are informative and strong: the Cragg-Donald statistic and the Kleibergen-Paap Wald F-test allow us to strongly reject the hypothesis of weak-instruments, and largely exceeds the standard rule-of-thumb value of 10 (see Baum et al. (2007, p. 490)). Results from these models suggest that the effect of retirement on mental health is much stronger and more precisely estimated, especially for retirement duration, than in the non-instrumented model. The FE-IV results imply that the initial transition to retirement causes a 10 point decrease in the depression score and an 8 point improvement in the mental health score, which correspond to a 34% decrease in depression a 15% improvement in mental health relative to the mean. After an initial improvement, progressive worsening with each year retired would mean that within 3.9 years (depression) or 5.9 years (mental health), the retiree would return to their pre-retirement level. Conversely, no causal effect on physical health is found, suggesting that the mental health effects of retirement are not linked to physical health effects.

Table 2: Fixed Effects, Fixed Effects Instrumental Variable results (full sample)

	Depression (GHQ)		Mental health (MCS)		Physical health (PCS)	
	FE	FE-IV	FE	FE-IV	FE	FE-IV
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)

[4] Terciles of household income per month, adjusted for CPI in the first wave the individual is observed, trimmed at the 99% percentile

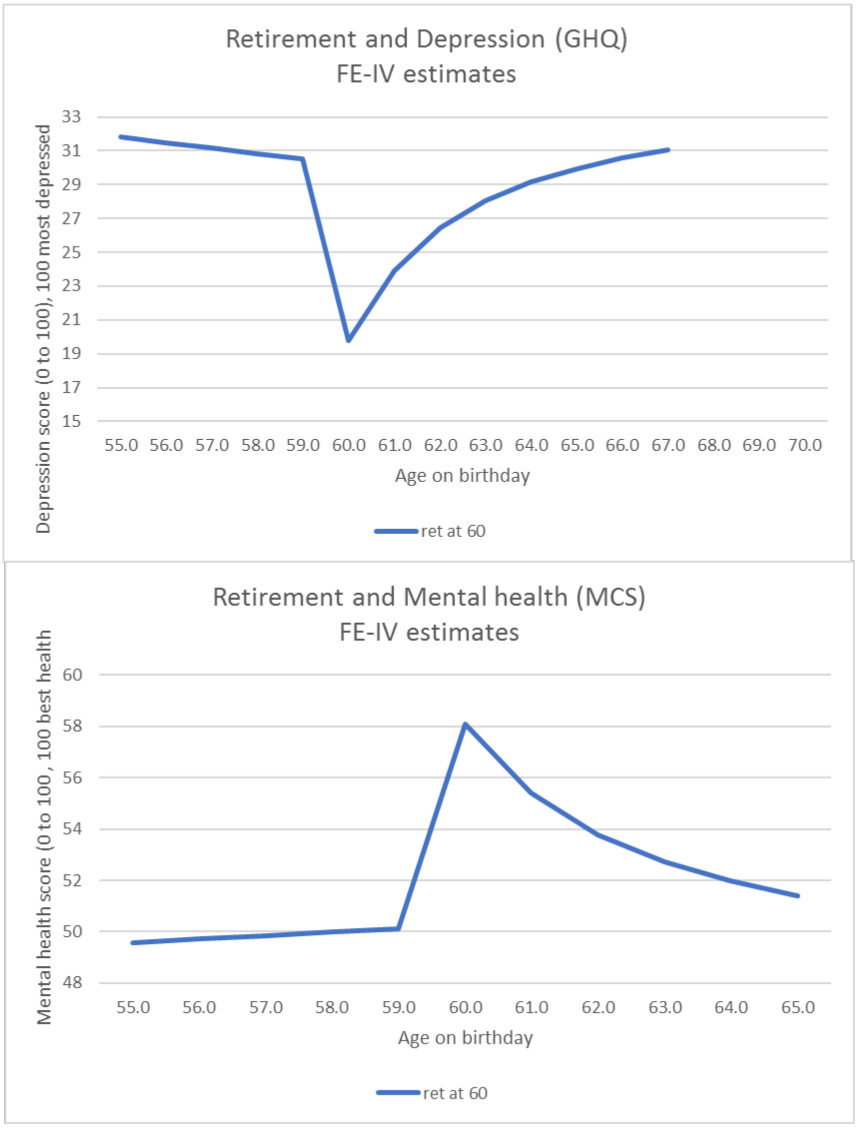
retired	-2.022***	-10.42***	1.303***	7.805***	-0.773***	-1.455
	(0.500)	(3.973)	(0.280)	(2.176)	(0.282)	(2.294)
logyret	0.622**	5.048**	-0.201	-3.161***	-0.0977	0.434
	(0.292)	(1.963)	(0.180)	(1.063)	(0.206)	(1.125)
Constant	135.3		38.23		132.1***	
	(87.18)		(62.08)		(49.54)	
Observations	27,497	26,266	27,497	26,266	27,497	26,266
R-squared	0.007	-0.012	0.005	-0.023	0.015	0.015
Nb of pidp	6,390	5,159	6,390	5,159	6,390	5,159
Individual FE	YES	YES	YES	YES	YES	YES
CDF weak IV		158.6		158.6		158.6
RKF weak IV		56.69		56.69		56.69

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 1 graphically displays the FE-IV results in columns (2) and (4) of Table 1, for a hypothetical, average individual who retires at 60, when reaching her SPA.[5] Because we observe individuals for a maximum of seven years after the reform (2010-2017), we limit the extent of the post-retirement extrapolation to this length. This graph shows that the post-retirement age-health trend seems to level out around the same level of depression and mental health as the pre-retirement level after about 6 or 7 years. Future waves of the Understanding Society survey could confirm this pattern further. In other words, it appears after a large short-term improvement in mental health, individuals experience a return to their pre-retirement mental health in the medium-term. This would imply that mental-health benefits from retirement are not permanent, similar to what recent studies have found for cognitive and physical health (Bertoni et al., 2018; Celidoni et al., 2017; Mazzonna and Peracchi, 2015)

[5] Due to our model's design, the size of the effect of retirement and retirement duration would be exactly the same for individuals retiring when reaching the SPA at later age (e.g., 61 or 62).

Figure 1: Effect of retirement on depression and mental health for a hypothetical individual



Sensitivity analysis

We conduct several sensitivity tests to test the robustness of these results (Appendix Table 1a and 1b). First, we include age dummies instead of a logged age trend, in order to allow the age effect on mental health to vary non-parametrically. This is important to ensure that the discontinuity in health at the State Pension Age is not manufactured by a mis-specified trend. For both MCS and GHQ, the direction of the effects is preserved, and the coefficient measuring the short-term effect of retirement retains 10% significance. However the coefficient on logyret loses significance and falls below 1. Second in order to account for the possibility of heteroscedasticity and serial correlation, we replace standard errors clustered at the cohort (month of birth) level with standard errors clustered at the individual level. For both mental health outcomes, the level of significance is preserved (and improves for logyret with the GHQ outcome).

Third, we test whether our results are robust to the inclusion of clinically relevant cut-offs for both mental health outcomes, using a linear probability model. In the case of GHQ, we use a validated, clinically relevant cut-off of 12 points out of 36 (D.P. et al., 1997; Kelly et al., 2008). For MCS, we use a cut off of 45 out of 100[6]. For both outcomes, we find that the direction and significance of the effects is preserved.

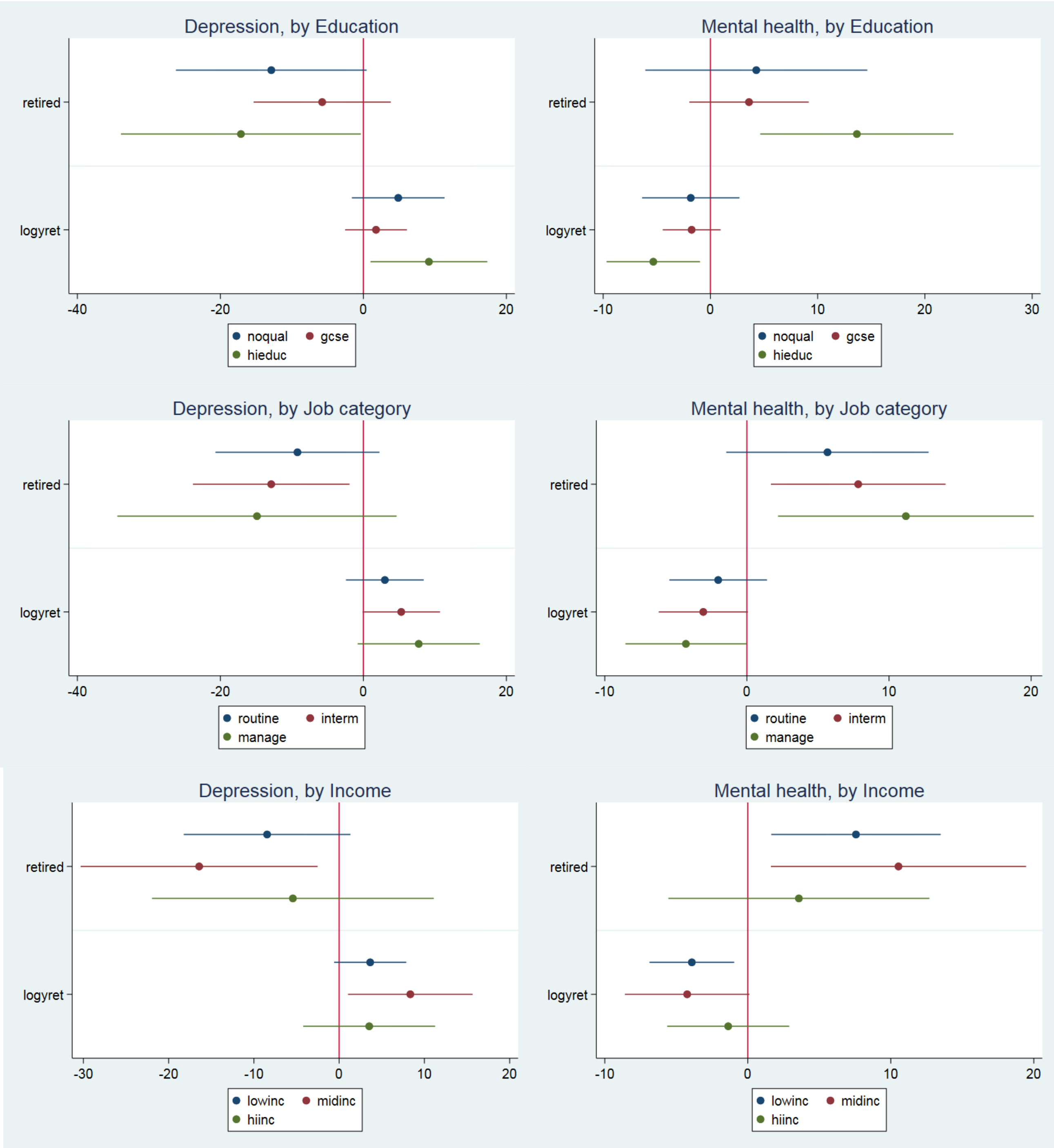
Finally, we test a set of alternative specifications where the short-term effect of the transition to retirement (measured by the coefficient on "retired") is allowed to last longer than one year, as the duration of the short-term period is assumed rather than tested by our base specification. We do this by redefining the "years retired" variable to start 1, 2, and 3 years after retirement respectively. We find that as the length of the "medium-term" period is artificially reduced relative to the "short-term" period, the coefficients retain significance at the 1% or 5% level, the short-term coefficient reduces, and the medium-term coefficient increases in magnitude. Although the Akaike and Bayesian Information Criterions, which are used to compare non-nested models, suggest that the original model is the best fit for the data (Appendix Tables 2a and 2b), we cannot be sure that the initial improvement in mental health is restricted to the first year.

Heterogeneity analysis

Figure 2 shows the point estimates and 95% confidence intervals for the coefficients on retired and asinh(years retired) from models run in separate samples according to the respondent's socio-economic status, defined according to educational attainment, job category, or baseline income. The confidence intervals demonstrate considerable overlap between different socio-economic categories, indicating that the estimates are not significantly different from each other at the 5% level. This is confirmed in FE-IV models with interaction terms run in the full sample, where the socio-economic categories are interacted with the variables of interest and instrumented accordingly. The interaction terms in these models are not significant (see Appendix Tables 3a and 3b), except for the interaction between "retired" and "high education" for the mental health outcome (MCS) in Appendix Table 3b, where the high SES category is compared to a joint category of low or medium.

[6] Although there is no universally accepted cut-off level for SF-12 (Córdoba-Doña et al., 2016), an optimal screening cut-off to evaluate 30-day depressive disorders (MCS) was set at 45.6 for the European population by Vilagut et al. (2013).

Figure 2: FE-IV results in different samples, by educational attainment, job category, and baseline income



Given that the social gradient patterns is consistent across mental health outcomes (left to right), and that we employ a tight specification combined with a modest sample size, the lack of significant heterogeneity may be due to a lack of power. An OLS-IV specification (with standard errors clustered at the individual level) uncovers further evidence of heterogeneity for the MCS outcome by education and job category (results available upon request).

The nature of the gradient appears to differ according to the definition of socio-economic status. When split according to education and job category, the highest SES category appears to derive the strongest ST benefits of retirement, combined with a stronger negative "catch-up" effect after the initial transition. When the sample is split by baseline income, however, wo men with baseline incomes in the top tercile experience the weakest effects of retirement, particularly when mental health is measured using the MCS index.

DISCUSSION

This study is the first to provide evidence of the differential short- and medium-term effects of retirement on mental health, and focuses on UK women who retired between 2010 and 2017. While previous studies have indicated that the average effect of being retired on mental health is positive[7], this study suggests that while the initial transition to retirement has positive mental health effects for women, mental health returns to pre-retirement levels within six years. This is very similar to estimated patterns for physical health and cognition (Bertoni et al., 2018; Celidoni et al., 2017; Mazzonna and Peracchi, 2015). There is a lack of strong heterogeneity in the effect of retirement by socio-economic status in our sample, although this could be explained by power limitations. The limited evidence on heterogeneity suggests that, in line with Bertoni et al (Bertoni et al., 2018), higher SES women (defined by their education or the category of their last job) may experience stronger beneficial effects of the initial transition to retirement on mental health, followed by a more negative change in trend.

Limitations and strengths

Our findings are derived from highly rigorous causal inference methods that exploit a reform in the State Pension Age to effectively control for the endogeneity of retirement decision. The corollary of using these methods is that we estimate the Local Average Treatment Effect of retirement on health, i.e. for those women who retired upon reaching the State Pension Age. These findings may therefore not apply to early or late retirees (Celidoni et al., 2017). Moreover, due to the limited time interval covered by the survey, we can be most confident of our findings for women who have only spent a limited amount of time in retirement (by construction, up to seven years). Additional waves of longitudinal datasets are required in order to study effects over the long-term, and for younger cohorts who will be subject to substantially higher SPA and will become eligible for State Pension benefits from 2017 onwards.

The pattern of short-term and medium-term effects identified in this article cannot be pinned down in terms of number of years, as shown in the sensitivity tests (Appendix Tables 2a and 2b). This is because identifying retirement effects non-parametrically using dummies (which must each be instrumented separately) is too demanding on the data. What we can say is that an initial period of improvement of a certain number of years or months is followed by a period of decline.

Possible explanations for findings

There are a number of explanations for our results. The initial improvement in mental health following retirement may be linked to the relief of stress experienced at the workplace, combined with greater time to engage in rewarding or relaxing activities. The subsequent worsening of mental health trends, reverting to pre-retirement levels over the period of observation, may be explained by the end of a "honeymoon" period at the start of retirement characterised by lower levels of daily social interactions or struggling to find meaning in life if work was instrumental to one's identity. A social gradient of retirement effects on mental health, with lower SES women experiencing weaker effects, is consistent with these explanations. Women of lower SES may not experience such strong

[7] When we run our model by including only a "retirement-status" dummy as the main regressor (instrumented with "being of State Pension Age", we observe that being retired is significantly associated to lower depression (GHQ) and better mental health (MCS). Results available upon request.

reductions in stress levels given that they may remain exposed to other stressors such as financial or security concerns post-retirement. They might also be less likely to define themselves through work if they did not have the opportunity to choose a career that is meaningful to them, and may therefore not lose out on the positive mental health effects of work in the same way.

Public health implications

These findings imply that the positive effects of retirement on mental health are likely to fade over the medium-term, while there is no evidence of retirement impacting the physical health of women in the UK. Mental health interventions for retirees after their first year in retirement could be trialled in order to investigate whether the short-term positive mental health effects of retirement could be extended for longer. Public health policies could also leverage the short-term positive effects of retirement on mental health to encourage recent retirees to undertake lasting behaviour change such as smoking cessation or increasing levels of exercise.

References

Atalay, K., Barrett, G.F., 2014. The causal effect of retirement on health: New evidence from Australian pension reform. Econ. Lett. 125, 392–395. https://doi.org/10.1016/j.econlet.2014.10.028

Atchley, R.C., 1982. Retirement as a Social Institution. Annu. Rev. Sociol. 8, 263–287. https://doi.org/10.1146/annurev.so.08.080182.001403

Battistin, E., Brugiavini, A., Rettore, E., Weber, G., 2009. The retirement consumption puzzle: evidence from a regression discontinuity approach. Am. Econ. Rev. 99, 2209–2226. https://doi.org/doi: 10.1257/aer.99.5.2209

Baum, C.F., Schaffer, M.E., Stillman, S., 2007. Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing. Stata J. Promot. Commun. Stat. Stata 7, 465–506. https://doi.org/10.1177/1536867X0700700402

Behncke, S., 2012. Does retirement trigger ill health? Health Econ. 21. https://doi.org/10.1002/hec.1712

Belloni, M., Meschi, E., Pasini, G., 2016. The effect on mental health of retiring during the economic crisis. Health Econ. 25, 126–140. https://doi.org/10.1002/hec.3377

Bertoni, M., Maggi, S., Weber, G., 2018. Work, retirement, and muscle strength loss in old age. Heal. Econ. (United Kingdom) 27, 115–128. https://doi.org/10.1002/hec.3517

Bloemen, H., Hochguertel, S., Zweerink, J., 2017. The causal effect of retirement on mortality: Evidence from targeted incentives to retire early. Heal. Econ. (United Kingdom) 26, e204–e218. https://doi.org/10.1002/hec.3493

Bonsang, E., Adam, S., Perelman, S., 2012. Does retirement affect cognitive functioning? J. Health Econ. 31, 490–501. https://doi.org/10.1016/j.jhealeco.2012.03.005

Celidoni, M., Dal Bianco, C., Weber, G., 2017. Retirement and cognitive decline. A longitudinal analysis using SHARE data. J. Health Econ. 56, 113–125. https://doi.org/10.1016/J.JHEALECO.2017.09.003

Coe, N.B., von Gaudecker, H.-M., Lindeboom, M., Maurer, J., 2012. The effect of retirement on cognitive functioning. Health Econ. 21, 913–927. https://doi.org/10.1002/hec.1771

Coe, N.B., Zamarro, G., 2011. Retirement effects on health in Europe. J. Health Econ. 30, 77–86. https://doi.org/10.1016/j.jhealeco.2010.11.002

D.P., G., R., G., N., S., T.B., U., M., P., O., G., 1997. The validity of two versions of the GHQ in the WHO study of mental illness in general health care. Psychol. Med. 27, 191–197.

Goldberg, D., Williams, P., 1988. A user's guide to the General Health Questionnaire. Windsor, Berks. : NFER-Nelson.

Hernaes, E., Markussen, S., Piggott, J., Vestad, O.L., 2013. Does retirement age impact mortality? J. Health Econ. 32, 586–598. https://doi.org/10.1016/j.jhealeco.2013.03.001

Hessel, P., 2016. Does retirement (really) lead to worse health among European men and women across all educational levels? Soc. Sci. Med. 151, 19–26. https://doi.org/10.1016/j.socscimed.2015.12.018

Johnston, D.W., Lee, W., 2009. Retiring to the good life ? The short-term effects of retirement on health. Econ. Lett. 103, 8–11. https://doi.org/10.1016/j.econlet.2009.01.015

Kelly, M.J., Dunstan, F.D., Lloyd, K., Fone, D.L., 2008. Evaluating cutpoints for the MHI-5 and MCS using the GHQ-12: a comparison of five different methods. BMC Psychiatry 8, 10. https://doi.org/10.1186/1471-244X-8-10

Kesternich, I., Siflinger, B., Smith, J.P., Winter, J.K., 2014. The Effects of World War II on Economic and Health Outcomes across Europe. Rev Econ Stat 96. https://doi.org/10.1162/REST

Lagarde, M., 2012. How to do (or not to do) ... Assessing the impact of a policy change with routine longitudinal data. Health Policy Plan. 27, 76–83. https://doi.org/10.1093/heapol/czr004

Lynn, P., 2009. Sample design for understanding society. Underst. Soc. Work. Pap. Ser 2009.

Mazzonna, F., Peracchi, F., 2015. Unhealthy Retirement? J. Hum. Resour. 52, 128–151. https://doi.org/10.3368/jhr.52.1.0914-6627R1

Mazzonna, F., Peracchi, F., 2012. Ageing, cognitive abilities and retirement. Eur. Econ. Rev. 56, 691–710. https://doi.org/10.1016/j.euroecorev.2012.03.004

Mein, G., Martikainen, P., Hemingway, H., Stansfeld, S., Marmot, M., 2003. Is retirement good or bad for mental and physical health functioning? J. Epidemiol. Community Health 57, 46–49.

OECD, 2019. Labour force participation rate (indicator) [WWW Document]. https://doi.org/OECD (2019), Labour force participation rate (indicator). doi: 10.1787/8a801325-en (Accessed on 18 February 2019)

Pensions at a Glance 2017 OECD and G20 Indicators, 2017. https://doi.org/https://doi.org/10.1787/pension_glance-2017-en

Rohwedder, S., Willis, R.J., 2010. Mental retirement. J. Econ. Perspect. 24, 119–138.

Thurley, D., Keen, R., Library, H. of C., 2017. State Pension age increases for women born in the 1950s, Briefing Paper. House of Commons Library, London.

Ware, J., M, K., Keller, S., 1994. SF-36 Physical and Mental Health Summary Scales: A User's Manual. Boston, MA: The Health Institute, New England Medical Center.

Ware, J.E., Kosinski, M., Keller, S.D., 1996. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity. Med. Care 34. https://doi.org/10.2307/3766749

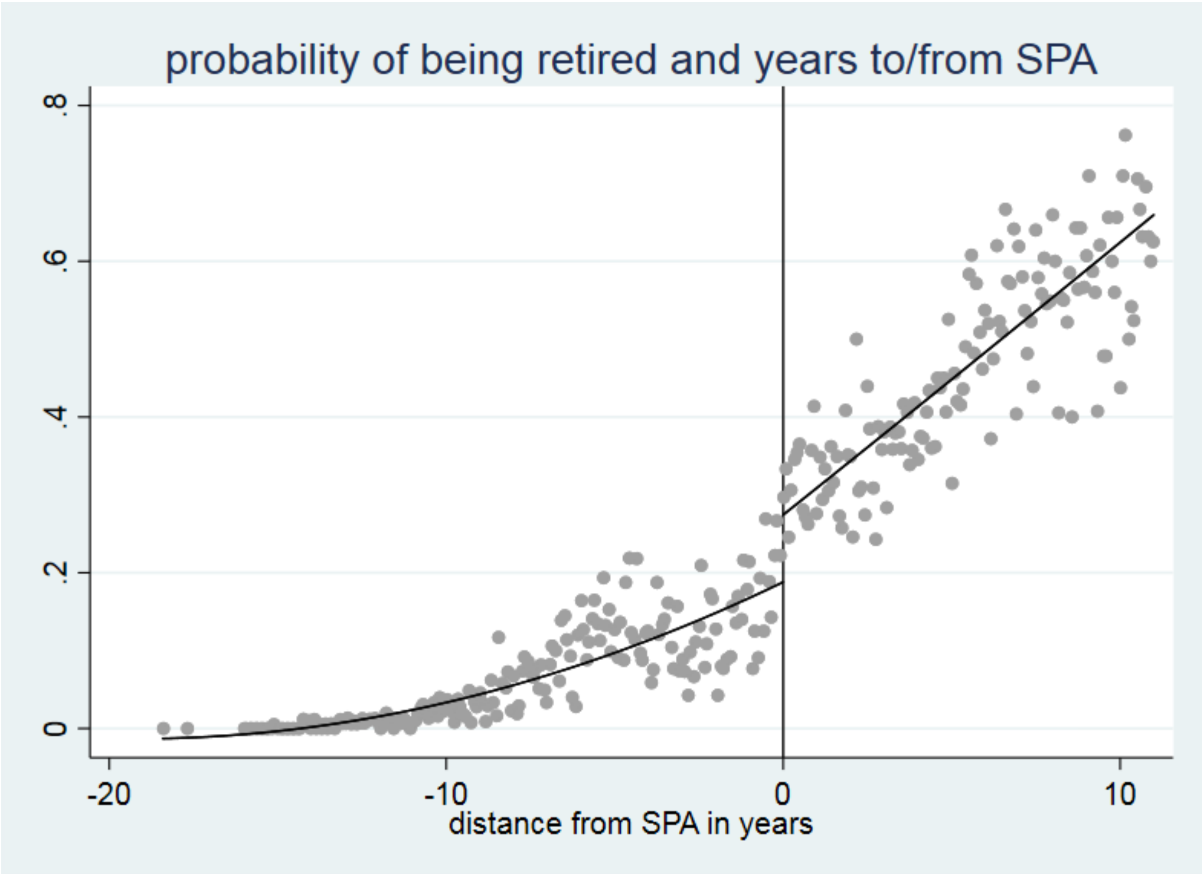
Westerlund, H., Kivimäki, M., Singh-Manoux, A., Melchior, M., Ferrie, J.E., Pentti, J., Jokela, M., Leineweber, C., Goldberg, M., Zins, M., Vahtera, J., 2009. Self-rated health before and after retirement in France (GAZEL): a cohort study. Lancet 374, 1889–1896. https://doi.org/10.1016/S0140-6736(09)61570-1

Xue, B., Cadar, D., Fleischmann, M., Stansfeld, S., Carr, E., Kivimaki, M., NcMunn, A., Head, J., 2018. Effect of retirement on cognitive function : the Whitehall II cohort study. Eur. J. Epidemiol. 33, 989–1001. https://doi.org/10.1007/s10654-017-0347-7

Zhu, R., 2016. Retirement and its consequences for women's health in Australia. Soc. Sci. Med. 163, 117–125. https://doi.org/10.1016/j.socscimed.2016.04.003

APPENDIX

Appendix Figure 1: Visual depiction of the first stage



Appendix Table 1a: Sensitivity tests for Depression outcome (GHQ)

VARIABLES	(1) Base specification	(2) Age dummies	(3) SEs clustered at individual-level	(4) Binary outcome
retired	-10.42***	-8.457*	-10.42***	-0.351**
	(3.973)	(4.971)	(3.783)	(0.154)
logyret	5.048**	-0.730	5.048***	0.187**
	(1.963)	(3.243)	(1.777)	(0.0744)
Observations	26,266	26,266	26,266	26,266
R-squared	-0.012	-0.014	-0.012	-0.016
Number of persons	5,159	5,159	5,159	5,159
Individual FE	YES	YES	YES	YES
Control for age	TREND	DUMMIES	TREND	TREND
Clustered SEs	COHORT	COHORT	INDIV	COHORT

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 1b: Sensitivity tests for Mental health outcome (MCS)

VARIABLES	(1) Base specification	(2) Age dummies	(3) SEs clustered at individual-level	(4) Binary outcome
retired	7.805***	5.148*	7.805***	-0.411***
	(2.176)	(2.654)	(2.409)	(0.124)
logyret	-3.161***	-0.816	-3.161***	0.172***
	(1.063)	(2.149)	(1.150)	(0.0586)
Observations	26,266	26,266	26,266	26,266
R-squared	-0.023	-0.005	-0.023	-0.028
Number of persons	5,159	5,159	5,159	5,159
Individual FE	YES	YES	YES	YES
Control for age	LOG TREND	DUMMIES	LOG TREND	LOG TREND
Clustered SEs	COHORT	COHORT	INDIV	COHORT

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2a: Test different lengths for "short-term" retirement effect – Depression (GHQ)

VARIABLES	(1) Base spec	(2) +1 year	(5) +2 years	(6) +3 years
retired	-10.42***	-7.778***	-6.356**	-5.442**
	(3.973)	(2.959)	(2.475)	(2.172)
logyret	5.048**			
	(1.963)			
logyret_1		4.841***		
		(1.754)		
logyret_2			5.634***	
			(1.926)	
logyret_3				7.368***
				(2.484)
Observations	26,266	26,266	26,266	26,266
R-squared	-0.012	-0.015	-0.017	-0.018
Number of persons	5,159	5,159	5,159	5,159
Individual FE	YES	YES	YES	YES
AIC	192115	192186	192238	192255
BIC	192303	192374	192426	192443

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 2b: Test different lengths for "short-term" retirement effect – Mental health (MCS)

VARIABLES	(1) Base spec	(2) +1 year	(3) +2 years	(4) +3 years
retired	7.805***	6.115***	5.268***	4.724***
	(2.176)	(1.656)	(1.422)	(1.271)
logyret	-3.161***			
	(1.063)			
logyret_1		-2.981***		
		(0.931)		
logyret_2			-3.560***	
			(1.043)	
logyret_3				-4.928***
				(1.403)
Observations	26,266	26,266	26,266	26,266
R-squared	-0.023	-0.024	-0.026	-0.029
Number of persons	5,159	5,159	5,159	5,159
Individual FE	YES	YES	YES	YES
AIC	168307	168339	168389	168469
BIC	168495	168527	168577	168657

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 3a: Heterogeneity by SES (3 sub-categories: low, medium, high)

	By education		By job category		By baseline income	
	GHQ	MCS	GHQ	MCS	GHQ	MCS
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
retired	-13.27*	3.432	-7.623	5.269	-8.338*	7.806***
	(6.920)	(5.511)	(6.224)	(3.578)	(4.820)	(2.949)
logyret	5.171	-1.652	3.082	-2.014	4.328**	-4.015***
	(3.340)	(2.497)	(2.937)	(1.737)	(2.178)	(1.485)
retgcse	6.526	0.0164				
	(8.811)	(6.308)				
Ret*hieduc	-2.991	10.07				
	(10.71)	(7.275)				
Logyret*gcse	-2.984	0.252				
	(4.204)	(2.878)				
Logyret*hieduc	3.500	-3.722				
	(5.377)	(3.518)				
Ret*interm			-1.340	0.142		
			(8.199)	(4.460)		
Ret*manage			-9.364	7.394		
			(12.35)	(6.359)		
Logyret*interm			1.537	-0.581		
			(3.932)	(2.225)		
Logyret*manage			5.092	-2.649		
			(5.602)	(3.080)		
Ret*midinc					-7.379	2.806
					(8.618)	(5.307)
Ret*hiinc					2.432	-4.212
					(8.925)	(5.142)
Logyret*midinc					3.623	-0.0858
					(4.158)	(2.654)
Logyret*hiinc					-0.947	2.566
					(4.272)	(2.481)
Observations	26,266	26,266	26,266	26,266	25,607	25,607
R-squared	-0.029	-0.050	-0.024	-0.036	-0.015	-0.024
Number of persons	5,159	5,159	5,159	5,159	5,024	5,024
Individual FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 3b: Heterogeneity by SES (2 sub-categories: high vs. medium or low)

	By education		By job category		By baseline income	
	GHQ	MCS	GHQ	MCS	GHQ	MCS
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
retired	-8.592**	3.425	-8.332**	5.329**	-11.58***	9.143***
	(3.800)	(2.496)	(3.824)	(2.442)	(3.831)	(2.429)
logyret	3.154*	-1.434	3.800**	-2.261*	5.797***	-3.924***
	(1.762)	(1.202)	(1.895)	(1.192)	(1.938)	(1.198)
ret*hieduc	-7.663	10.07*				
	(8.938)	(5.341)				
logyret*hieduc	5.522	-3.944				
	(4.361)	(2.639)				
ret*manage			-8.687	7.417		
			(11.28)	(5.844)		
logyret*manage			4.386	-2.430		
			(4.975)	(2.748)		
ret*hiinc					5.692	-5.537
					(8.524)	(5.425)
logyret*hiinc					-2.436	2.460
					(4.127)	(2.563)
Observations	26,266	26,266	26,266	26,266	25,607	25,607
R-squared	-0.029	-0.050	-0.024	-0.036	-0.011	-0.022
Number of persons	5,159	5,159	5,159	5,159	5,024	5,024
Individual FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1