

Panel Data Research Center, Keio University

PDRC Discussion Paper Series

Use it Too Much and Lose Everything? The Effects of Hours of Work on Health

Shinya Kajitani, Colin McKenzie, Sakata Kei

3 February, 2020

DP2016-009

<https://www.pdrc.keio.ac.jp/en/publications/dp/2306/>



Panel Data Research Center, Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
info@pdrc.keio.ac.jp
3 February, 2020

Use it Too Much and Lose Everything? The Effects of Hours of Work on Health

Shinya Kajitani, Colin McKenzie, Sakata Kei

PDRC Keio DP2016-009

3 February, 2020

JEL Classification: I10, J2

Keywords: health;working hours;endogeneity;pensions;retirement

Abstract

The “use it or lose it” hypothesis examines whether active involvement in work can prevent cognitive decline for elderly workers, but how much work is good for health is as yet unknown. We examine the causal impact of working hours on various health outcomes of Australian men aged 40 years and over using panel data from the HILDA survey over the period 2001 to 2012. To capture the potential non-linear dependence of health status on working hours, the models for health outcomes include working hours and its square as explanatory variables. We deal with the potential endogeneity of working hours by using the instrumental variable estimation technique with instruments based on the age for pension eligibility. There is time series variability in the pension eligibility ages for some men because in 2009, significant changes in pension eligibility ages were announced for some, but not all males. A non-linear causal effect of working hours on health is confirmed. For males working relatively moderate hours (up to around 24–27 hours for a week), an increase in working hours has a positive impact on health, but thereafter an increase in working hours has a negative impact on health.

Shinya Kajitani

Faculty of Economics, Kyoto Sangyo University

Motoyama, Kamigamo, Kita-ku, Kyoto City 603-8555, JAPAN

kajitani@cc.kyoto-su.ac.jp

Colin McKenzie

Faculty of Economics, Keio University

2-15-45 Mita, Minato-ku, Tokyo 108-8345, JAPAN

mckenzie@z8.keio.jp

Sakata Kei

Australian Institute of Family Studies

Level 4, 40 City Road, Southbank, Victoria 3006, AUSTRALIA

kei.sakata@aifs.gov.au

Acknowledgement: The authors would also like to thank Daiji Kawaguchi and participants at the following workshops and conferences, the Ageing and Longevity in the UK and Japan: Perspectives

from the Health and Social Sciences Workshop (King's College London), the Globalization and the Building of a High Quality Economic System Conference (Keio University), the 10th International Conference of the Thailand Econometric Society 2017 (Chiang Mai University), the 22th Euroasia and Business and Economics (EBES) Society Conference (Sapienza University of Rome), the 2019 Gerontological Society of America Annual Scientific Meeting (Austin, Texas), and the Spring Conference of the Japanese Economic Association (Ritsumeikan University) for their helpful and constructive comments on earlier versions of this paper. This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute: Applied Economic and Social Research, the University of Melbourne (the Melbourne Institute). The findings and views reported in this paper, however, are those of the authors and should not be attributed to DSS, the Melbourne Institute, or the Australian Institute of Family Studies.

This work was supported by Japan Society for the Promotion of Science (JSPS) Grant in Aid for Scientific Research (B) No. JP24330093 for a project on "Retirement Behaviour of the Aged and their Cognitive Ability and Health" 24330093 2012-2016.

Use it Too Much and Lose Everything? The Effects of Hours of Work on Health

Shinya Kajitani^a, Faculty of Economics, Kyoto Sangyo University

Colin McKenzie^b, Faculty of Economics, Keio University

Kei Sakata^c, Australian Institute of Family Studies

Acknowledgements

The authors would also like to thank Daiji Kawaguchi and participants at the following workshops and conferences, the Ageing and Longevity in the UK and Japan: Perspectives from the Health and Social Sciences Workshop (King's College London), the Globalization and the Building of a High Quality Economic System Conference (Keio University), the 10th International Conference of the Thailand Econometric Society 2017 (Chiang Mai University), the 22th Euroasia and Business and Economics (EBES) Society Conference (Sapienza University of Rome), the 2019 Gerontological Society of America Annual Scientific Meeting (Austin, Texas), and the Spring Conference of the Japanese Economic Association (Ritsumeikan University) for their helpful and constructive comments on earlier versions of this paper. This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute: Applied Economic and Social Research, the University of Melbourne (the Melbourne Institute). The findings and views reported in this paper, however, are those of the authors and should not be attributed to DSS, the Melbourne Institute, or the Australian Institute of Family Studies.

Declarations

The authors declare they do not have any actual or potential conflict of interest in relation to this research. No ethic approval was required for this study.

Funding

This research was supported by the Japan Society for the Promotion of Science (JSPS) Grant

^a Shinya Kajitani, Faculty of Economics, Kyoto Sangyo University, Motoyama, Kamigamo, Kita-ku, Kyoto City 603-8555, JAPAN. Email: kajitani@cc.kyoto-su.ac.jp

^b Corresponding author: Colin McKenzie, Faculty of Economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, JAPAN. Email: mckenzie@keio.jp

^c Kei Sakata, Australian Institute of Family Studies, Level 4, 40 City Road, Southbank, Victoria 3006, AUSTRALIA. Email: kei.sakata@aifs.gov.au

in Aid for Scientific Research (B) No. JP24330093 for a project on “Retirement Behaviour of the Aged and their Cognitive Ability and Health” (P.I.: Kei Sakata), a Japan Society for the Promotion of Science (JSPS) Grant in Aid for Scientific Research (B) No. 16H03607 for a project "An Empirical Analysis of Parental Employment Status, Time Allocation and Way of Thinking, and their Children's Human Capital Formation" (Project Leader: Midori Wakabayashi), and Kyoto Sangyo University Research Grant No. E1901 (P.I.: Shinya Kajitani).

ABSTRACT

The “use it or lose it” hypothesis examines whether active involvement in work can prevent cognitive decline for elderly workers, but how much work is good for health is as yet unknown. We examine the causal impact of working hours on various health outcomes of Australian men aged 40 years and over using panel data from the HILDA survey over the period 2001 to 2012. To capture the potential non-linear dependence of health status on working hours, the models for health outcomes include working hours and its square as explanatory variables. We deal with the potential endogeneity of working hours by using the instrumental variable estimation technique with instruments based on the age for pension eligibility. There is time series variability in the pension eligibility ages for some men because in 2009, significant changes in pension eligibility ages were announced for some, but not all males. A non-linear causal effect of working hours on health is confirmed. For males working relatively moderate hours (up to around 24–27 hours for a week), an increase in working hours has a positive impact on health, but thereafter an increase in working hours has a negative impact on health.

Keywords: health, working hours, endogeneity, pensions, retirement.

JEL Classification Nos: I10, J2

Highlights

- For Australian men aged 40 years and over, a non-linear causal relationship between working hours and health is observed.
- Moderate hours of work provide health benefits, and these benefits peak around 24–27 hours of work for a week. The positive health effects decline after the peak of 24–27 hours of work, and around 50 hours of work all the positive effects of work disappear.
- Working slightly more than 50 hours leads to health outcomes that are worse than those observed for men who are not working at all.

1. Introduction

How does work affect health? Is work bad for health? Does work have any health benefits? There is an extremely large literature in epidemiology, occupational psychology, and health economics that examines these issues (see, for example, Bannai and Tamakoshi, 2014, Bassanini and Caroli, 2015). Some papers examine the *extensive* margin of work (working or not working), for example, by examining the impact of unemployment and job loss on health outcomes. Using fixed-effect models for Australian, Canadian and UK panel data, Llena-Nozal (2009) shows that the shift from being employed to being unemployed has adverse effects on mental health. On the other hand, using German data, Schmitz (2011) finds no significant effect of plant closures on various health outcomes. Heller-Sahlgren (2017) finds no immediate effect, but rather a delayed impact of retirement on mental health. Another stream of research on the extensive margin examines whether retirement has any impact on cognitive functioning and health. Such studies test the so-called “use it or lose it” hypothesis (Rohwedder and Willis, 2010, Coe and Zamarro 2011, Bonsang *et al.* 2012, De Grip *et al.* 2012, Mazzonna and Peracchi 2012, Blake and Garrouste 2017, Kajitani *et al.* 2017a, Mazzonna and Peracchi 2017, Atalay *et al.* 2019). Overall, these studies tend to suggest that retirement has a negative impact on cognitive functioning, but positive impacts on health outcomes.

Other papers examine the *intensive* margin of work, that is, the number of hours worked. The main focus of these analyses is on the effects of working long hours on various health outcomes. Such studies reveal that working long hours has adverse effects on health (Spurgeon *et al.* 1997, Sparks *et al.* 1997, Frijters *et al.* 2009, Bannai *et al.* 2015, Nie *et al.* 2015). In a paper that emphasises the importance of employment policy in Europe in the context of the relationship between work and health, Barnay (2016) argues that while there is a large literature in epidemiology on the intensive margin of work on health outcomes, there is next to no literature in economics (see also Bannai and Tamakoshi 2014). Lee and Lee (2016) examine the impacts of working hours on injury rates at the worker’s workplace which is regarded as one of the risk factors for their health, exploiting a quasi-natural experiment in South Korea. Assuming that the injury risk is a quadratic function of working hours, they find that the function is convex, which indicates that shortening very long working hours could be effective in reducing the injury rate. However, previous studies do not examine the effects of moderate working hours on health or the optimal number of hours worked. Robone *et al.* (2011) indicate that having a part-time job, as compared to a full-time job, has a positive impact on the health of people who are satisfied with their working hours. For males, Llena-Nozal (2009) finds that compared to working full-time, working overtime leads to a worsening of mental health outcomes¹. This highlights the fact that the relationship

¹ In contrast, for women, Llena-Nozal *et al.* (2004) find that compared to working full-time,

between work and health may not be linear. Work can be a double-edged sword in that it can have both positive and negative effects. Interactions with people at work may help maintain an individual's cognitive functions and his/her mental health. Moreover, working individuals have more incentive to invest in health repair activities in order to be 'fit' in the labour market. On the other hand, long working hours can cause fatigue and stress on both physical and mental levels which potentially damage an individual's overall health and also reduce the amount of time that can be investment in health repair activities. Most of the previous studies treat long working hours as a 0–1 dummy variable which defines long working hours as working more than 50 or 60 hours per week. This means that they implicitly assume that long working hours have a constant shift effect on health status. They do not deal with the potential non-linear effects of working hours on health.

For health outcomes, the literature on the extensive margin suggests that working may be better than not working, while the literature on the intensive margin suggests that working extremely long hours is worse than working a normal working week. Combining these two observations suggests there is a non-linear relationship between workhours and health outcomes which we seek to capture using a quadratic form in workhours. Although not directly focusing on health outcomes, one explanation for the results generated in Pencavel's (2014, 2016) analysis of the nonlinear relationship between working hours and productivity in munitions factories in Britain during the First World War is stress and fatigue generated by long working hours that then affects productivity (see also Collewet and Sauermann, 2017).

We contribute to the existing literature in two ways. First, we focus on not only labor market participation (the extensive margin), but also working hours (the intensive margin). Secondly, the literature examining the impact of retirement on cognitive function examines the 'use it or lose it' hypothesis, namely that not working (not using your brain) leads to losses of cognitive functioning. Here, we also examine the relevance of this hypothesis for a broader set of health outcomes. In addition, we focus on the 'use it too much and lose everything' hypothesis which refers to the situation where working too much can lead to not just a loss of cognitive functioning (see Kajitani *et al.* 2017b), but also declines in health status across the board.

We examine the causal impact of working hours on the health outcomes of middle-aged and older male adults (men aged 40 years and over) in Australia using Wave 1 to Wave 12 of the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The health outcomes are measured using five self-assessed health score components computed from the SF-36 (the 36-Item Short Form Health Survey) which is one of the most widely used self-assessed measures of health status. These score components cover both an individual's

moving to part-time work or overtime lead to improvements in mental health outcomes.

physical and mental health (see Ware, 2000). One of the issues in estimating the causal relationship between hours worked and health is what is called the ‘healthy worker effect’ (Bassanini and Caroli, 2015), that is, healthy workers are more likely to be employed and work longer. Thus, the presence of the healthy worker effect implies the existence of reverse causality. We deal with the potential endogeneity of decisions relating to working hours by using the instrumental variable estimation technique. One advantage of using a sample of middle-aged and older adults is that it enables us to use information related to the eligibility age for pension benefits as instruments for variables related to working hours.

Our empirical evidence shows that there is non-linearity in the effects of working hours on self-assessed health status. To be more specific, there is an inverted U-shaped relationship. When working hours are less than around 24–27 hours a week, increases in working hours have a positive impact on health. However, when working hours are greater than this threshold, increases in working hours have negative impacts on health. Compared to males who do not work, working hours slightly over 50 hours will lead to worse health outcomes depending on the measure of health status. These results suggest that men in old age could maintain or improve their health status ability compared to not working by working in a part-time job that requires them to work around 24–27 hours per week. The results are consistent with the analysis on cognitive functioning of Kajitani *et al.* (2017b).

The rest of this paper is organized as follows: Section 2 provides a brief description of the Australian pension system. Section 3 presents the empirical framework used in this paper. Section 4 describes the data, and Section 5 reports the results of estimation and discusses their implications. The last section concludes this paper.

2. Australian pension system

In Australia, the timing of retirement is closely related to the pension system. Retirement income consists of three sources: a means tested public pension, a mandatory employer-contributed private retirement savings account, and voluntary private retirement savings. Since there is no mandatory retirement age in Australia, an elderly Australian can also continue to work to supplement his/her pension.

The maximum benefit payment from the public age pension is set at 25% of male total average earnings. Since the introduction of the good and service tax in 2000, a supplement for compensation has been added to the payment. The maximum basic rates of the public pension were A\$10,262.20 per annum for the standard rate and A\$8,569.60 per annum for the partnered rate in September 2001. In September 2012, at the time of HILDA Wave 12, the maximum standard rate had been increased to A\$ 18,512.00 and the partnered rate to A\$ 13,954.20. It is important to note that the maximum benefit is subject to both an income test and an asset test (for details see Atalay and Barrett, 2015).

Age and residency conditions are applied for the eligibility of age pension. To satisfy the

residency condition, individuals are required to have been resident in Australia for at least ten years. The aged pension eligibility age for males was set at 65 in 1909. In June 2008, a little over 80% males and females aged 65 or over were receiving aged pensions (Hammer (2008, Chart 20)). In 2009–10, ABS (2011, Table 3) estimates that for 59.2% of those people receiving the aged pension the government pensions and allowances contributed to more than 90% of their gross household income. That is, for a large number of Australians the aged pension provides an important source of income.

On 12 May 2009, Wayne Swan, the Treasurer, and Jenny Macklin, the Minister for Families, Housing, Community Services and Indigenous Affairs, jointly announced that starting in 2017 the qualifying age for the aged pension for males would be gradually increased from 65 to 67 by 2023². As can be seen from Table I, this policy change raised the pension eligibility ages for males in the “younger” generations from 65 to 65.5, 66, 66.5 or 67 depending on their birthdates. For males born before or on 30 June 1952, there was no change in their pension eligibility age. We will use variations in the pension eligibility age to identify the labour supply behavior of the middle age and elderly male workers in Australia.

[Table I around here]

3. Estimation model and identification strategy

Our identification strategy exploits the variation in working hours, while controlling for time-invariant individual characteristics. In order to capture the possible non-linear effects of working hours on health status, we consider the following model for health outcomes³:

$$y_{it} = \alpha_1 + \alpha_2 WH_{it}^2 + X1_{it}\beta + u_{it}, \quad i = 1, \dots, N, t = 1, \dots, T_i \quad (1a)$$

$$u_{it} = \mu_i + \epsilon_{it} \quad (1b)$$

where y_{it} denotes various health outcomes (physical functioning, bodily pain, general health, vitality, and mental health) for individual i at the time of the survey t , and WH_{it} is working hours. In estimating equation (1a), we include those individuals whose working hours are zero. $X1_{it}$ denotes a vector of time variant control variables: a spouse 0–1 dummy

² The joint press release is available from the following URL: <http://ministers.treasury.gov.au/DisplayDocs.aspx?doc=pressreleases/2009/056.htm&pageID=003&min=wms&Year=2009&DocType=0> (Accessed 30 June 2019)

³ An alternative to the parametric model in equation (1a) to account for the non-linear effect of working hours on health status would be to estimate a semi-parametric or non-parametric model. However, such an approach makes it rather difficult to deal with the potential endogeneity between working hours and health status. Here, we put priority on dealing with the potential endogeneity of working hours.

variable, *Married*, which takes the value one if the respondent has a spouse and zero otherwise; the number of dependent children, *Number of dependent children*; the respondent's age, *Age*, which controls for age-related effects; and a house ownership 0–1 dummy variable, *Ownhouse*, which indicates whether the respondent owns or is in the process of owning his house as a proxy for assets. The variables related to the respondent's marital status and the number of dependent children are included because communication and interaction with other family members may prevent declines in health, particularly in mental health. In addition, the number of dependent children is included since it can be argued that people with dependent children may be likely to invest more in their health capital. The house ownership dummy is included to control for the effects of assets holdings on health. In addition, $X1_{it}$ includes four 0–1 regional dummies. N is the number of individuals and T_i is the number of observations available for individual i indicating that we have an unbalanced panel. As equation (1b) indicates, u_{it} is an error term which consists of a time invariant individual fixed effect, μ_i , and an idiosyncratic error, ϵ_{it} , so that we allow for some degree of individual heterogeneity. The coefficients α_1 and α_2 in equation (1a) capture the non-linear effect of working hours on a health outcome. Given the discussion in section 1 that some work is better than no work, and that too much work may be worse than some work, it is expected that $\alpha_2 < 0$ and $\alpha_1 > 0$. Holding everything else constant, it is easy to see that the value of a health score is maximized when $WH_i = -\alpha_1/(2\alpha_2)$, and that for $WH_i = -\alpha_1/\alpha_2$ the level of health is the same as it would be if the respondent is not working.

The possibility of the endogeneity of the respondents' working hours in equation (1a) is a major obstacle to estimate the causal impact of working hours on health. As discussed in section 1, this particular identification problem is called the 'healthy worker effect', that is healthy individuals are more likely to be employed and to work longer whereas unhealthy workers may decide to leave the workforce or work short hours. Individuals, who are healthier and, therefore, tend to earn a relatively higher wage, could decide to reduce their hours of work. The same logic can be applied to health outcomes.

In order to account for the potential endogeneity of working hours and its square in equation (1a), we use an instrumental variable estimator using two instruments, *Age difference 1* and *Age difference 2*, that are based on pension eligibility ages. These instruments measure the distance from an individual's retirement age. More specifically, *Age difference 1* is the difference between a respondent's age and the respondent's pension eligibility age provided the respondent has reached the eligibility age and zero otherwise, while *Age difference 2* is the difference between a respondent's age and the respondent's pension eligibility age provided the respondent has *not* reached the eligibility age and zero otherwise. That is, we assume with the following models:

$$WH_{it} = \mu_1 Age\ difference\ 1_{it} + \mu_2 Age\ difference\ 2_{it} + X1_{it}\mu + w_{it}, \quad (2)$$

$$WH_{it}^2 = \mu\mu_1 Age\ difference\ 1_{it} + \mu\mu_2 Age\ difference\ 2_{it} + X1_{it}\mu\mu + ww_{it}, \quad (3)$$

where the error terms, w_{it} and ww_{it} , include individual fixed effects.

As can be seen from Table I, in our sample, a reasonable amount of variation in the eligibility age is observed. Although the pension eligibility age for men remained at the age of 65 for individuals born before or on 30 June 1952, the proportion of this group in Wave 9 (just after the 2009 pension reform) is just 47.9%. The proportion of men whose pension eligibility age is 67 in Wave 9 consists of 42.0%.

[Table I around here]

The two instruments in equations (2) and (3) are closely related to one of the standard instruments used for the analysis of the causal relationship between retirement and cognitive functioning (Atalay *et al.* 2019, Bonsang *et al.* 2012, Coe and Zamarro 2011, Mazzonna and Peracchi 2012, Mazzonna and Peracchi 2017, Rohwedder and Willis 2010). It is worth pointing out that, in the literature that tries to determine the causal impacts of retirement on health outcomes, these pension related instruments are used to explain the retirement decision, that is, the extensive margin, not the number of hours worked.

Here we rely on Neumark and Powers (2000, 2003/4, 2005, 2006), Kudrna and Woodland (2011a,b) and Vere (2011) to justify using these pension eligibility age related variables as instruments for also explaining the intensive margin of working hours as well. Neumark and Powers (2000, 2003/4, 2005, 2006) provide a theoretical explanation of the connection between social security in the United States and hours worked prior to retirement, while Vere (2011) provides some empirical evidence for the US on the connection between pension payments and hours worked. In the context of a dynamic general equilibrium model, Kudrna and Woodland (2011a,b) provide evidence for Australia on the effect of the major reform of the Australian pension system that was announced by the Australian Government in the 2009/2010 national budget that we are considering on working hours. Furthermore, in the data set we are using for individuals who report their working hours are positive, Figure I graphs the relationship between the average of working hours and the difference between a respondent's age and pension eligibility age. For values of this difference less than 10, there would appear to be a negative relationship between this variable and the average of working hours⁴. One reason for limiting our sample to males aged 40 and over is to ensure that there

⁴ For values of *Age minus Aged pension eligibility age* greater than or equal to 15, the sample size to compute each average is rather small (less than 10), so that the estimation of the average is not so precise.

is a sufficient connection between the changes in pension eligibility ages and working hours.

[Figure I around here]

In both equations (2) and (3), the dependent variable is non-negative, but Wooldridge (2010, p. 90) makes the important point that the application of instrumental variable (IV) estimator is not limited to the case where the dependent variable(s) in the first stage is (are) continuous, so even though working hours are non-negative this is not an impediment to the application of the IV technique.

4. Data: Overview of the HILDA Survey

Our data are drawn from the first 12 waves of the “Household, Income and Labour Dynamics in Australia (HILDA) Survey,” from Wave 1 conducted in 2001 to Wave 12 conducted in 2012. The HILDA Survey which is conducted by the Melbourne Institute: Applied Economics and Social Research, the University of Melbourne is a broad social and economic longitudinal survey. Since 2001, the HILDA Survey has asked Australian respondents about their economic and subjective well-being, family structures, and labor market dynamics. Household included in the survey were selected using a three-stage approach. First, a sample of 488 Census Collection Districts (CDs) were randomly selected from across Australia. Second, within each of these CDs, a sample of dwellings was selected based on expected response rates and occupancy rates. Finally, within each dwelling, up to three households were selected to be part of the sample. In addition, the sample was replenished in Wave 11. One aim of this replenishment was to provide better coverage of migrants for inclusion in the HILDA Survey⁵.

The HILDA survey contains the SF-36 (the 36-Item Short Form Health Survey) which is one of the most widely used self-assessed measures of health status⁶. It consists of eight scaled self-assessed health scores: physical functioning; role physical; bodily pain; general health; vitality; social functioning; role emotional and mental health. The eight categories are scaled by the weighted sums of their questions, and are converted to a 0–100 scale. 0 is equivalent to the highest disability, and 100 is equivalent to the lowest disability⁷. Of the

⁵ Detailed information on the sample design of the HILDA Survey is available in Wooden *et al.* (2002) and Watson and Wooden (2013).

⁶ Cai and Kalb (2006) in an analysis of the SF-36 data in HILDA report some discrepancies in health status when it is self-reported compared to when the data is obtained through interviews.

⁷ Although each self assessed health score is bounded from below by zero and above by 100, and there are individuals who score one of these two boundary values (see Table II), we do not employ a double sided Tobit type estimator when estimating equation (1a).

eight health scores, we use physical functioning, bodily pain, general health, vitality, and mental health. The other three scores, role physical; social functioning; and role emotional, are eliminated from the analysis because they display little variation.

The general release HILDA data sets that we are using do not contain information on the day or month of birth of the respondent. In each wave, for each individual two ages are reported: the age as at 30 June of the year of the wave, and the age at the time of the survey. Using information on the age of respondents as at 30 June in the first wave it is possible to determine the pension eligibility ages of all individuals except those whose age is reported to be 44 and 47 in this wave. Using both ages and the information on when the respondent was interviewed, it is possible to determine the pension eligibility age for 25% and 32% of those respondents whose ages are reported to be 44 and 47 at 30 June 2001 in wave 1, respectively. Those respondents whose pension eligibility age could not be determined were excluded from the analysis. Table O1 in the online supplementary material indicates the impact of this exclusion on the sample size available in each wave.

As stated in section 2, the 2009 pension reform was announced on 12 May 2009. As the interview periods for HILDA's waves 8 and 9 are from 20 August 2008 to 27 February 2009 and from 20 August 2009 to 11 March 2010, respectively, the 2009 pension reform falls right between waves 8 and 9.

The exact definitions of all the variables used in the analysis in this paper are summarized in Appendix I. The sample is restricted to individuals who meet the following five criteria: (i) males aged 40 and over in Wave 1 of the survey or males who turn 40 after Wave 1 but are only included for the Waves where they are aged 40 and over; (ii) all five scores relating to health status are available; (iii) age and working hours are less than age and working hours in the top 1% percentile, respectively; (iv) respondents who report they are unemployed are excluded; and (v) information on all the relevant variables is available. We target this age group as people start experiencing some health declines. For example, the Australian Heart Foundation recommends that people over 45 (over 35 for Aboriginal and Torres Strait Islanders) have a heart health check⁸. The National Bowel Cancer Screening Program is offered to all Australians aged 50 – 74 in 2020⁹. In addition, this sample selection is applied for our identification strategy as younger age groups are thought to be indifferent to changes in their retirement age. Even if we did not have attrition, the second part of criterion (i) means we will not have a balanced panel data set. After imposing these restrictions, a sample of 35,196 observations on 6,022 individuals remains. Table O1 in the online supplementary

⁸ See, for example, <https://www.heartfoundation.org.au/your-heart/know-your-risks/heart-health-check> (Accessed 29 January 2020).

⁹ See, for example, <http://www.cancerscreening.gov.au/internet/screening/publishing.nsf/Content/nbcsp-fact-sheet> (Accessed 29 January 2020).

material indicates the effect of these criterion on the sample size available in each wave.

Table II displays descriptive statistics on all the variables used in the analysis. In this table, *Working hours* is the respondent's usual hours of working per week. As a result, the mean value of *Working hours* for males is 27.52 hours.

[Table II around here]

5. Estimation results

All regression results reported in this section are estimated using STATA version 15 (StataCorp, 2017).

5.1 Estimation using a fixed effect instrumental variable (FEIV) estimator

Table III presents the estimates for the equations in both the first and second stages of the fixed effect instrumental variable (FEIV) estimator. In stage 2 (columns (3.B1)–(3.B5)), where the fixed effect instrumental variable estimates of equation (1) are presented, it is observed that for each health variable both working hours and its square are individually statistically significant with the estimated coefficients have signs that are consistent with the a priori expectations suggested in section 3, so that there is an inverted U shape relationship between health and working hours. An examination of the endogeneity tests in Table III which test the joint significance of two residuals added to the model estimated by FE (see Wooldridge 2010, pp. 352–354), the results clearly reject the null hypothesis that *Working hours* and *Working hours squared* are exogenously determined. The Cragg-Donald (1993) test indicates we do not have a problem of “weak” instruments. Finally, the Hausman (1978) tests for choosing between the pooled IV model and FEIV indicate clearly that the FEIV estimator is to be preferred. In stage 1 (columns (3.A1)–(3.A2)) estimates of the “reduced form” equations for working hours (equation (2)) and working hours squared (equation (3)) are displayed. The key variables, *Age difference 1* and *Age difference 2*, are both individually and jointly significant in explaining the variation in working hours and working hours squared.

[Table III around here]

The results in Table III indicate that, for all health measures for males as working hours increase from zero the magnitude of the positive impact of working hours on their health status is decreasing until working hours reaches a threshold. Above the threshold, further increases in working hours have a negative impact on their self-assessed health status.

Where does the threshold occur? In other words, when does the impact of working hours on health status change from being positive to negative? The peaks occur around 26 hours for physical functioning, 27 hours for role physical, 26 hours for general health, 26 hours for

vitality, and 24 hours for mental health¹⁰. These are all slightly higher than the estimated working hours (22–30 hours) for when cognitive abilities peak that are reported in Kajitani *et al.* (2017b).

In Figure II, we graph the relationship between impacts of working hours and self-assessed health status after controlling for other variables, using the estimated coefficients presented in Table III. Moreover, Figure II also shows that the health status of those working extremely long hours can be lower than those who are not working at all. This suggests that long working hours can lead to a deterioration of health status across the board for men.

[Figure II around here]

The results presented in Table III and graphed in Figure II show that there is non-linearity in the causal effects of working hours on self-assessed health status for middle aged and older males living in Australia. Our findings are consistent with our hypothesis, that is, work can help in maintaining health status for middle age and elderly male workers, but long working hours have a negative effect. Our results indicate that part-time work is an effective way to maintain to health in retirement.

5.2 Robustness Checks

Four robustness checks of the results in section 5.1 are presented here. They involve using a different estimator, using an additional instrumental variable, excluding full-time workers, and restricting the age of respondents. All these checks provide results that are consistent with the results in Table III.

The first robustness check involves an alternative way to generate some instruments by assuming the following equation to explain an individual's hours worked:

$$WH_{it}^* = \gamma_1 Age\ difference\ 1_{it} + \gamma_2 Age\ difference\ 2_{it} + X2_{it}\delta + e_{it} \quad (4a)$$

$$\begin{aligned} WH_{it} &= 0 \quad \text{if } WH_{it}^* \leq 0 \\ &= WH_{it}^* \quad \text{if } 0 < WH_{it}^*, \end{aligned} \quad (4b)$$

where WH_{it}^* denotes an unobserved latent variable, which is connected to the observed working hours WH_{it} through equation (4b). $X2_{it}$ consists of the same vector of control variables as used in equation (1a) and two control variables related to the level of schooling, and e_{it} is a disturbance which is assumed to be normally, independently and identically

¹⁰ The 95% confidence intervals for these thresholds computed using bootstrapped estimates of the standard errors based on 500 repetitions are: (24.8, 28.2) (Physical functioning), (25.2, 29.1) (Bodily pain), (24.1, 28.4) (General health), (24.1, 27.7) (Vitality) and (21.3, 26.5) (Mental health).

distributed with a zero mean and variance σ^2 .

As is clear from equations (4a) and (4b), we have another issue in examining the effects of working hours on health, that is, working hours are censored (for example, retirees report zero working hours), so we estimate equations (4a) and (4b) using a Tobit estimator¹¹. From equations (4a) and (4b), the conditional expectation of WH_{it} can be computed as

$$E(WH_{it} | Z_{it}) = \Phi\left(\frac{Z_{it}\varepsilon}{\sigma}\right) Z_{it}\varepsilon + \sigma\phi\left(\frac{Z_{it}\varepsilon}{\sigma}\right) \quad (5),$$

where Z_{it} and ε are the vectors of regressors and parameters in equation (2), respectively, and $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative distribution function and probability distribution function of the standard normal distribution, respectively (see Greene 2008, p. 871). Using estimates of the parameters of equation (4a), this conditional expectation can be estimated. This estimate is denoted by \widehat{WH}_{it} . In estimating equation (1a) using a 2SLS procedure, \widehat{WH}_{it} and \widehat{WH}_{it}^2 are then used as *instruments* for WH_{it} and WH_{it}^2 , respectively (see Wooldridge 2010, p. 268). This estimator is labelled as Tobit+FEIV estimator. Zhang (2013) and Chi and Drewianka (2014) provide examples of this type of estimator in different contexts as a way to avoid the “forbidden regression” problem.

The results for this Tobit+FEIV estimator are reported in Table IV. The second stage estimates of equation (1a) for the five health variables are reported in columns (4.C1)–(4.C5). The results are similar to the results reported in Table III. The Tobit estimates of equations (4a) and (4b) are reported in column (4.A1) where we see that both the key variables, *Age difference 1* and *Age difference 2*, are both individually and jointly significant in explaining the variation in working hours. The first stage of the FEIV estimates for *Working hours* and *Working hours squared* are reported in columns (4.B1) and (4.B2), respectively, and it can be seen that \widehat{WH}_{it} and \widehat{WH}_{it}^2 are individually significant. It should be noted that because \widehat{WH}_{it} and \widehat{WH}_{it}^2 are non-linear functions of Z_{it} the conditions that Z_{it} is required to satisfy for this estimator to be consistent are stronger than an instrumental estimator using just Z_{it} , so it is possible to compare these two estimators using a Hausman (1978) style testing procedure. This Hausman test is reported as Hausman test of three-stage estimator vs two-stage estimator, and indicates for three of the five health scores that the three-stage estimator is problematic compared to the two stage estimator reported in section 5.1. For this reason, the remaining results in this paper adopt the estimator use to generate the results in

¹¹ Since we are estimating this model on panel data, it could be argued that we should employ a Tobit estimator with fixed effects. Given the incidental parameter problem, this estimator would not provide a consistent estimator of any of the parameters of the model (see Greene, 2004), so we employ a pooled Tobit estimator.

Table III.

The results of the second robust check are reported in Table V reports where in addition to *Age difference 1* and *Age difference 2* the vacancy rate is also used as an instrument when estimating equation (1a) by FEIV. The vacancy rate shows both geographical (differing across each state) and temporal variation, and since it provides a measure of macroeconomic conditions at the state level is likely to affect the hours worked. On the other hand, it is unlikely to directly affect any of the health scores of individuals and so it is an appropriate instrument. As can be seen in columns (5.B1)–(5.B5), *Working hours* and *Working hours squared* are individually highly significant in all the equations and have estimated coefficients with the expected signs. All five health score equations pass the overidentifying test. In the first stage (columns (5.A1)–(5.A2)), the vacancy rate is highly significant in both equations.

The third robustness check restricts the analysis to exclude full-time workers. One potential criticism of our approach is that not all workers are free to choose their hours of work, so that it is inappropriate to treat hours worked as being endogenous for all workers. For example, in the absence of a collective agreement or award for *full-time* workers the ordinary hours of hours per week that they work is set at 38 hours¹². In addition, full-time employees have the right to refuse to work unreasonable additional hours above this minimum number of hours worked. To take account of this possibility Table VI replicates Table III but excludes full-time workers. As a result of excluding full-time workers, the sample size is reduced from 35,196 to 16,137, and the number of individuals is reduced from 6,022 to 3,301. The results in Table VI are entirely consistent with the results obtained in Table III.

The final robustness test displayed in Table VII restricts the ages of respondents to the range 45–75. This reduces the sample from 35,196 to 25,865, and the number of individuals is reduced from 6,022 to 4,471. Excluding men aged over 75 years is likely to exclude people who have already stopped working due to retirement and whose average health status is lower than the rest of the sample. Individuals aged 40–44 in 2001 are more than 20 years away from their pension eligibility age, so that changes in their pension eligibility age may have little impact on their working hours. The results in Table VII are again consistent with the results in Table III, so that making these exclusions do not undermine our results.

5.3 Testing the Identifying Assumptions

¹² Section 20 of the federal *Fair Work Act* of 2009. Available from URL: <https://www.legislation.gov.au/Details/C2018C00512> (Accessed 21 December 2019). Its predecessor, the federal *Workplace Relations Act* of 1996, contained a similar provision.

The final robustness check seeks to check some of the identifying assumptions underlying the analysis¹³. Given the non-linearities in (1a) and, we cannot write down an explicit reduced form for y_{it} . If we assume $\alpha_2 = 0$ and make the structural equation in (1a) linear in working hours, then the reduced form is¹⁴

$$y_{it} = \lambda_1 \text{Age difference } 1_{it} + \lambda_2 \text{Age difference } 2_{it} + X1_{it}\lambda + v_{it}, \quad (6)$$

Table VIII presents estimates of equation (6) for the five health measures estimated using all the data and a fixed effects (FE) estimator. In four of the five equations, *Age difference 1* is statistically significant and in the other equation for mental health *Age difference 2* is significant.

[Table VIII around here]

As can be seen from Table I, the pension eligibility age for individuals born on or before 30 June 1952 (hereafter called the 1951 “cohort”) does not change with the 2009 reform. If we imagine the reduced form age-health profiles prior to the 2009 pension reform, then in the analysis in section 5.1 for the 1951 cohort whose pension eligibility age does not change before and after the 2009 pension reform it is implicitly assumed that when they are under the age of 65 that their age-health profiles do not change around the time of the pension reform. Based on equation (6), we test this assumption using the following model estimated using data from Waves 1–12 for those individuals in the 1951 cohort who are not yet 65¹⁵:

¹³ We are greatly indebted to Daiji Kawaguchi for suggesting robustness tests along these lines.

¹⁴ For women in the United Kingdom, Carrino *et al.* (2018) present estimates of the impact of pension reform on health outcomes using a similar type of reduced form model.

¹⁵ Although equation (6) contains three age related variables, *Age difference 1_{it}*, *Age difference 2_{it}*, and *Age_{it}*, so for a general test of structural change we would have the variables *Age difference 1_{it}*, *Age difference 2_{it}*, *Age_{it}*, *Age difference 1_{it}AfterReform_t*, *Age difference 2_{it}AfterReform_t*, *Age_{it}AfterReform_t* and *AfterReform_t*. When the sample is limited to the 1951 cohort which has a constant pension eligibility age of 65, *Age difference 1_{it}* + *Age difference 2_{it}* = *Age_{it}* – 65 for all observations, so perfect multicollinearity exists among these three variables. This justifies dropping *Age difference 1_{it}*, and *Age difference 1_{it}AfterReform_t*. Similarly when the sample is limited to those members of the 1951 Cohort aged under 65 then *Age difference 2_{it}* = *Age_{it}* – 65, and *Age difference 1_{it}* = 0 for all observations, so we can only include one of these three variables; *AfterReform_t*, *Age difference 2_{it}AfterReform_t* and *Age_{it}*.

$$y_{it} = \eta_1 AfterReform_t + \eta_2 Age\ difference\ 2_{it} AfterReform_t + X1_{it}\lambda + v_{it}, \quad (7)$$

where $AfterReform_t$ is a 0–1 dummy variable taking the value 1 if the observation is after the pension reform of 2009, and 0 if the observation is before the pension reform of 2009. Equations (9.1)–(9.5) in Table IX present estimates of equation (7) estimated using a fixed effect (FE) estimator. It is easy to observe that the coefficients η_1 and η_2 are individually and jointly insignificant. That is, the pension reform has not affected the age-health profiles for individuals in the 1951 cohort who have not reached the age of 65.

[Table IX around here]

As can be seen from Table I, prior to the pension reform of 2009 the pension eligibility age for individuals in the 1951 cohort and for individuals born on or after 1 July 1952 (hereafter called the 1952 “cohort”) are the same. If we imagine the reduced form age-health profiles for these two groups of individuals aged under 65 prior to the 2009 pension reform, we are implicitly assuming them to be the same. Based on equation (6), we test this assumption using the following model estimated using data from Waves 1–8 for those individuals in the 1951 and the 1952 cohorts who have not reached the age of 65¹⁶:

$$y_{it} = \delta_1 Agedifference\ 2_{it} Cohort1951_i + X1_{it}\lambda + v_{it}, \quad (8)$$

where $Cohort1951_i$ is a 0–1 dummy variable taking the value 1 if the individual is in the 1951 cohort, and taking the value zero if the individual is in the 1952 cohort. Equations (9.6)–(9.10) in Table IX presents FE estimates of (8) using individuals in the 1951 cohort who are under 65, and for individuals in the 1952 cohort using only data from before the 2009 reform. In no case is the parameter δ_1 significantly different from zero supporting the assumption that before the 2009 reform the 1951 and 1952 cohorts are similar.

[Table IX around here]

¹⁶ Although equation (6) contains three age related variables, $Age\ difference\ 1_{it}$, $Age\ difference\ 2_{it}$, and Age_{it} , so for a general test of structural change we would have the variables $Age\ difference\ 1_{it}$, $Age\ difference\ 2_{it}$, Age_{it} , $Age\ difference\ 1_{it}Cohort1951_i$, $Age\ difference\ 2_{it}Cohort1951_i$, $Age_{it}Cohort1951_i$ and $Cohort1951_i$. When the sample is limited to the pre-reform sample where all individuals have a constant pension eligibility age of 65 and individuals are assumed to be younger than 65, $Age\ difference\ 1_{it} + Age\ difference\ 2_{it} = Age_{it} - 65$, and $Age\ difference\ 1_{it} = 0$ for all observations. This justifies dropping $Age\ difference\ 1_{it}$, $Age\ difference\ 1_{it}Cohort1951_i$, $Age_{it}Cohort1951_i$ and $Age\ difference\ 2_{it}$. In addition, since $Cohort1951_i$ is a time-independent variable and we are using a fixed effect estimator, its coefficient is not identified.

6. Concluding remarks

We examine the causal impact of working hours on the self-assessed health status of middle-aged and elderly males living in Australia using longitudinal data from the Household Income and Labour Dynamics in Australia (HILDA) Survey. The literature in this area is very limited in that it does not consider a non-linearity in the effect of working hours on health. Many previous studies examine the ‘use it or lose it’ hypothesis which tests whether or not retirement (not using your brain) leads to losses of cognitive functioning. On the other hand, we also examine the relevance of this hypothesis for a broader set of health outcomes. In addition, we focus on the ‘use it too much and lose everything’ hypothesis which refers to the situation where working too much can lead to not just a loss of cognitive functioning (see Kajitani *et al.* 2017b), but also declines in health status across the board. This study is unique in that we focus on not only labor market participation (the extensive margin), but also the intensive margin of work (working hours) and that we determine the optimal working hours for middle aged and elderly male workers in terms of maximizing their health status.

Using five measures of self-assessed health status in the SF-36, it is found that for working hours up to 24–27 hours per week increases in working hours have a positive impact on cognition for males depending on the health measure. After that, working hours have a negative impact on health status. Compared to males who do not work, working hours over 48–54 hours will lead to worse health outcomes depending on the measure. This indicates that the differences in working hours is an important factor in explaining differences in the health outcomes of middle aged and elderly male adults.

Thus, in middle and old age, adopting part-time work as a pattern of work could be effective in maintaining/improving the health status of individuals compared to when they do not work. Previous studies on retirement and cognitive functioning indicate that increasing the qualifying age for a pension can not only reduce the government social security expenditures but can potentially reduce the risk of cognitive deterioration. However, our study highlights that raising the qualifying age for a pension can reduce the risk of health deterioration, but that too much work can have quite adverse effects on health status.

Two important areas for future research are whether these results reported here for Australian males can be found for females and workers in other countries.

[Appendix I around here]

Online Supplementary Material

Table O1: Effect of Selection Criterion on Sample Sizes

References

- Atalay K, Barrett GF. 2015. The impact of age pension eligibility age on retirement and program dependence: Evidence from an Australian experiment. *Review of Economics and Statistics* **97**(1), 71–87. DOI: 10.1162/REST_a_00443
- Atalay K, Barrett, GF, Staneva A. 2019. The effect of retirement on elderly cognitive functioning. *Journal of Health Economics*, **66**, 37–53. DOI: 10.1016/j.jhealeco.2019.04.006
- Australian Bureau of Statistics (ABS) (2011). &530.0 - Household Expenditure Survey, Australia: Summary of Results, 2009-10.9 September 2011. Accessed 26 September 2019. Available from URL: <https://www.abs.gov.au/ausstats/abs@.nsf/featurearticlesbyCatalogue/82CB91B523027894CA25790200154C25?OpenDocument>
- Bannai A, Tamakoshi A. 2014. The association between long working hours and health: A systematic review of epidemiological evidence. *Scandinavian Journal of Work, Environment and Health* **40**(1), 5–18. DOI:10.5271/sjweh.3388
- Bannai A, Ukawa S, Tamakoshi A. 2015. Long working hours and psychological distress among school teachers in Japan. *Journal of Occupational Health* **57**(1), 20–27. DOI: 10.1539/joh.14-0127-OA
- Barnay T. 2016. Health, work and working conditions: A review of the European economic literature. *European Journal of Health Economics* **17**, 693–709. DOI: 10.1007/s10198-015-0715-8
- Bassanini A, Caroli E. 2015. Is work bad for health? The role of constraint vs. choice. *Annals of Economics and Statistics* **119–120**, 13–37. DOI: 10.15609/annaeconstat2009.119-120.13
- Blake H, Garrouste C. 2017. Collateral effects of a pension reform in France. Accessed 10 December 2019. Available from URL: <https://hal.archives-ouvertes.fr/hal-01500683/>
- Bonsang E, Adam S, Perelman S. 2012. Does retirement affect cognitive functioning? *Journal of Health Economics* **31**(3), 490–501. DOI: 10.1016/j.jhealeco.2012.03.005
- Cai L, Kalb G. 2006. Health status and labour force participation: Evidence from Australia. *Health Economics* **15**, 241–261. DOI: 10.1002/hec.1053
- Carrino L., Glaser K. & Avendano M. (2018). Later pension, poorer health? Evidence from the new state pension in the UK. Harvard Center for Population and Development Studies Working Paper. DOI: 10.2139/ssrn.3195760
- Chi M, Drewianka S. (2014). How much is a green card worth? Evidence from Mexican men who marry women born in the U.S. *Labour Economics*, **31**, 103–116. DOI:10.1016/j.labeco.2014.10.004
- Coe NB, Zamorro G. 2011. Retirement effects on health in Europe. *Journal of Health Economics* **30**(1), 77–86. DOI: 10.1016/j.jhealeco.2010.11.002
- Collewet M, Sauermann J. 2017 Working hours and productivity. *Labour Economics* **47**(1), 96–106. DOI: 10.1016/j.labeco.2017.03.006
- Commonwealth of Australia 2009. *Securable and Sustainable Pensions*. Accessed 4

- December 2016. Available from URL: http://www.budget.gov.au/2009-10/content/glossy/pension/download/pensions_overview.pdf#search='Securable+and+sustainable+pensions'
- Cragg JG, Donald SG. 1993. Testing identifiability and specification in instrumental variable models. *Econometric Theory* **9**(2), 222–240. DOI: 10.1017/S0266466600007519
- De Grip A, Lindeboom M, Montizaan R. 2012. Shattered dreams: The effects of changing the pension system late in the game. *Economic Journal* **122**(559), 1–25. DOI: 10.1111/j.1468-0297.2011.02486.x
- Frijters P, Johnston D, Meng X. 2009. The mental health cost of long working hours: The case of rural Chinese migrants. mimeo.
- Greene W. 2004. The behavior of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal* **7**(1), 98–119. DOI: 10.1111/j.1368-423X.2004.00123.x
- Greene WH. 2008. *Econometric Analysis*, 6th edition. Pearson: New Jersey.
- Hammer J. (2008). Pension Review. Background Paper. Accessed 25 September 2019. Available from URL: https://www.dss.gov.au/sites/default/files/documents/06_2012/pension_review_paper.pdf
- Hausman JA. (1978). Specification tests in econometrics. *Econometrica* **46**(6), 1251–1271. DOI: 10.2307/1913827
- Heller-Sahlgren G. 2017. Retirement blues. *Journal of Health Economics* **54**(1), 66–78. DOI: 10.1016/j.jhealeco.2017.03.007
- Kajitani S, Sakata K, McKenzie C. 2017a. Occupation, retirement and cognitive functioning. *Ageing & Society* **37**(8), 1568–1596. DOI: 10.1017/S0144686X16000465
- Kajitani S, McKenzie C, Sakata K. 2017b. Use it too much and lose it? The effect of working hours on cognitive ability, Panel Data Research Center at Keio University Discussion Paper no 2016-008.
- Kudrna G, Woodland AD. 2011a. Implications of the 2009 age pension reform in Australia: A dynamic general equilibrium analysis. *Economic Record* **87**, 183–201. DOI: 10.1111/j.1475-4932.2010.00703.x
- Kudrna G, Woodland A. 2011b. An inter-temporal general equilibrium analysis of the Australian age pension means test. *Journal of Macroeconomics* **33**(1), 61–79. DOI: 10.1016/j.jmacro.2010.09.006
- Lee J, Lee YK. 2016. Can working hour reduction save workers? *Labour Economics* **40**, 25–36. DOI: 10.1016/j.labeco.2016.02.004
- Llena-Nozal A. 2009. The effect of work status and working conditions on mental health in four OECD countries. *National Institute Economic Review* **209**(1), 72–87. DOI: 10.1177/0027950109345234

- Llena-Nozal A, Lindeboom M, Portrait F. 2004. The effect of work on mental health: Does occupation matter? *Health Economics* **13**, 1045–1062. DOI: 10.1002/hec.929
- Mazzonna F, Peracchi F. 2012. Ageing, cognitive abilities and retirement. *European Economic Review* **56**(4), 691–710. DOI: 10.1016/j.euroecorev.2012.03.004
- Mazzonna F, Peracchi F. 2017. Unhealthy retirement? *Journal of Human Resources* **52**(1), 128–151. DOI:10.3368/jhr.52.1.0914-6627R1
- Neumark D, Powers E. 2000. Welfare for the elderly: The effects of SSI on pre-retirement labor supply. *Journal of Public Economics* **78**(1-2), 51–80. DOI: 10.1016/S0047-2727(99)00111-5
- Neumark D, Powers E. 2003/2004. The effect of the SSI program on labor supply: Improved evidence from social security administrative files. *Social Security Bulletin* **65**(3), 45–60.
- Neumark D, Powers E. 2005. The effects of changes in state SSI supplements on preretirement labor supply. *Public Finance Review* **33**(1), 3–35. DOI: 10.1177/1091142104270655
- Neumark D, Powers E. 2006. Supplemental security income, labor supply, and migration. *Journal of Population Economics* **19**(3), 447–479. DOI: 10.1007/s00148-006-0074-y
- Nie P, Otterbach S, Sousa-Poza A. 2015. Long hours and health in China. *China Economics Review* **33**, 212–229. DOI: 10.1016/j.chieco.2015.02.004
- Pencavel J. 2014. The productivity of working hours. *Economic Journal* **125**, 2052–2076. DOI: 10.1111/ecoj.12166
- Pencavel J. 2016. Recovery from work and the productivity of working hours. *Economica* **83**, 545–563. DOI: 10.1111/ecca.12206
- Robone S, Jones A, Rice N. 2011. Contractual conditions, working conditions and their impact on health and well-being. *European Journal of Health Economics* **12**(5), 429–444. DOI: 10.1007/s10198-010-0256-0
- Rohwedder S, Willis RJ. 2010. Mental retirement. *Journal of Economic Perspectives* **24**(1), 119–138. DOI: 10.1257/089533010797456247
- Schmitz H. 2011. Why are the unemployed in worse health? The causal effect of unemployment on health. *Labour Economics* **18**(1), 71–79. DOI: 10.1016/j.labeco.2010.08.005
- Sparks K, Cooper C, Fried Y, Shimon A. 1997. The effects of hours of work on health: A meta-analytic review. *Journal of Occupational and Organizational Psychology* **70**(4), 391–408.
- Spurgeon A, Harrington JM, Cooper C. 1997. Health and safety problems associated with long working hours: A review of the current position. *Occupational and Environmental Medicine* **54**(6), 367–375. DOI: 10.1136/oem.54.6.367
- StataCorp. 2017. *Stata: Release 15*. College Station, TX: StataCorp LLC.

- Vere JP. 2011. Social security and elderly labor supply: Evidence from the Health and Retirement Study. *Labour Economics* **18**, 676–686. DOI: 10.1016/j.labeco.2011.02.001
- Ware JD. 2000. SF-36 Health Survey Update, *Spine* **25**(24), 3130–3139.
- Watson N, Wooden M. 2013. Adding a top-up sample to the Household, Income and Labour Dynamics in Australia Survey. *Australian Economic Review* **46**(4), 489–498.
DOI: 10.1111/1467-8462.12027
- Wooden M, Freidin S, Watson N. 2002. The Household, Income and Labour Dynamics in Australia (HILDA) Survey: Wave 1. *Australian Economic Review* **35**(3), 339–348.
DOI: 10.1111/1467-8462.00252
- Wooldridge J. 2010. *Econometric Analysis of Cross Section and Panel Data*, 2nd edition, The MIT Press: Cambridge.
- Zhang Y. (2013). Does private tutoring improve students' National College Entrance Exam performance? A case study from Jinan, China. *Economics of Education Review* **32**, 1–28. DOI: 10.1016/j.econedurev.2012.09.008

Appendix I: Definitions of Variables

Name	Definition
Physical functioning	The SF-36 physical functioning score (0–100)
Bodily pain	The SF-36 bodily score (0–100)
General health	The SF-36 general health score (0–100)
Vitality	The SF-36 vitality score (0–100)
Mental health	The SF-36 mental health score (0–100)
Working hours	The number of usual or average working hours per week the respondent works.
Working hours-squared	(Working hours) ²
Age	Respondent's age in years at the time of the survey
School years 7–10 (benchmark: the respondent's highest years of school completed are under 7)	0–1 dummy variable taking the value of unity if the respondent's highest years of school completed are between 7 and 10, and 0 otherwise.
School years 11 and over (benchmark: the respondent's highest years of school completed are under 7)	0–1 dummy variable taking the value of unity if the respondent's highest years of school completed are 11 and over, and 0 otherwise.
Married	0–1 dummy variable taking the value of unity if the respondent is currently married, and 0 otherwise.
Number of dependent children	The number of the respondents' children who reside with him/her and are either are aged under 15 years or are aged 16–24 years and are enrolled in full-time education.
Ownhouse	0–1 dummy variable taking the value of unity if the respondent owns his/her own house or is currently paying off a mortgage, and 0 otherwise.
Inner regional	0–1 dummy variable taking the value unity if the respondent lives in inner regional Australia, and 0 otherwise.
Outer regional	0–1 dummy variable taking the value unity if the respondent lives in outer regional Australia, and 0 otherwise.
Remote	0–1 dummy variable taking the value unity if the respondent lives in remote Australia, and 0 otherwise.
Very remote	0–1 dummy variable taking the value unity if the respondent lives in very remote Australia, and 0 otherwise.
Age difference 1	=(Respondent's age in years at the time of the survey)-(Aged pension eligibility age) if the respondent <i>has reached</i> Aged pension eligibility age at the time of the survey, =0 otherwise
Age difference 2	=(Respondent's age in years at the time of the survey)-(Aged pension eligibility age) if the respondent <i>has not yet reached</i> Aged pension eligibility age at the time of the survey, =0 otherwise
Vacancy rate	(Job vacancy/Employed)*100, where Job vacancy denotes the number of job vacancies in the state where the respondent lives in each waves which are reported by the Australian Bureau of Statistics (ABS), and Employed denotes the number of total employed persons in the relevant state in November at each waves which are reported by the ABS. Noted that we use the values for May 2008 because of a lack of values for November 2008.
AfterReform	0-1 dummy variable taking the value unity if the observation is after May 2009, and 0 otherwise.
Cohort1951	0-1 dummy variable taking the value unity if individual was born before or on 30 June 1952, and 0 otherwise.

Table I: Australian Age Pension Eligibility Ages and the Sample Distribution

Birthday (Age)	Pension eligibility age before April 2009	Sample (%)		Pension eligibility age after May 2009	Sample (%)
		Wave 1	Wave 8		Wave 9
Up to 30 June 1952 (Aged 49 and older in Wave 1)	65	68.6%	50.8%	65	47.9%
1 July 1952 to 31 December 1953 (Aged 47–48 in Wave 1)	65	5.0%	3.8%	65.5	3.7%
1 January 1954 to 30 June 1955 (Aged 46 in Wave 1)	65	3.8%	2.5%	66	2.8%
1 July 1955 to 31 December 1956 (Aged 44–45 in Wave 1)	65	4.9%	3.8%	66.5	3.6%
From 1 January 1957 (Aged 43 and younger in Wave 1)	65	17.7%	39.1%	67	42.0%
Total		100.0%	100.0%		100.0%
Number of Observations		2,893	2,713		2,725

Source: For pension eligibility ages: Commonwealth of Australia (2009), p. 9. Sample proportions in each group are authors' calculations using data from the HILDA Survey.

Table II: Descriptive Statistics

Variable	Mean	Std. Dev.	Min	Max
Physical functioning	79.61	23.81	0	100
Bodily pain	70.06	24.63	0	100
General health	65.17	21.67	0	100
Vitality	61.59	19.70	0	100
Mental health	76.36	16.59	0	100
Working hours	27.52	23.42	0	84
Working hours-squared	1305.50	1363.31	0	7056
Age	57.04	11.94	40	86
School years 7–10	0.48	0.50	0	1
School years 11 and over	0.48	0.50	0	1
Married	0.79	0.41	0	1
Number of dependent children	0.63	1.07	0	9
Ownhouse	0.82	0.38	0	1
Inner regional	0.27	0.44	0	1
Outer regional	0.13	0.33	0	1
Remote	0.02	0.12	0	1
Very remote	0.00	0.06	0	1
Vacancy rate	1.42	0.42	0.56	3.26
Age difference 1	2.17	4.58	0	21
Age difference 2	-10.50	9.15	-27	0

Source: Authors' calculations using data from Waves 1–12 of the HILDA Survey.

Note: 35,196 observations from 6,022 individuals are used to compute the descriptive statistics in each case.

Table III: The impacts of working hours on health

VARIABLES	1st Stage (FE)		2nd stage (FEIV)				
	(3.A1) Working hours	(3.A2) Working hours- squared	(3.B1) Physical functioning	(3.B2) Bodily pain	(3.B3) General health	(3.B4) Vitality	(3.B5) Mental health
Age difference 1	-2.137*** [0.182]	-101.079*** [12.824]					
Age difference 2	-2.644*** [0.206]	-137.930*** [14.178]					
Working hours			5.516*** [1.079]	3.204*** [0.736]	2.446*** [0.608]	4.021*** [0.826]	1.608*** [0.469]
Working hours-squared			-0.104*** [0.019]	-0.059*** [0.013]	-0.047*** [0.011]	-0.078*** [0.015]	-0.033*** [0.008]
Age	1.312*** [0.170]	63.360*** [12.317]	-0.734*** [0.237]	-0.615*** [0.152]	-0.687*** [0.127]	-0.424** [0.183]	-0.184* [0.101]
Married	0.975* [0.577]	71.488* [38.663]	1.645 [1.751]	-0.123 [1.155]	-0.105 [0.964]	1.359 [1.337]	3.070*** [0.785]
Number of dependent children	-0.578** [0.247]	-14.403 [17.335]	1.943** [0.925]	1.117* [0.589]	0.802 [0.496]	0.975 [0.713]	0.258 [0.391]
Ownhouse	-0.441 [0.511]	-37.369 [34.174]	-1.096 [1.483]	-0.697 [0.946]	-0.954 [0.793]	-1.115 [1.115]	-0.232 [0.626]
Inner regional	-4.010*** [1.148]	-202.885*** [68.519]	-0.531 [2.624]	0.465 [1.659]	0.008 [1.383]	0.140 [2.010]	0.403 [1.073]
Outer regional	-2.542 [1.625]	-92.198 [98.354]	3.015 [3.714]	2.469 [2.346]	0.414 [2.009]	2.835 [2.716]	2.125 [1.564]
Remote	2.960 [3.391]	258.600 [208.603]	10.528 [7.779]	7.415 [4.792]	3.637 [4.284]	6.513 [5.990]	4.069 [3.369]
Very remote	11.635* [6.661]	770.223** [347.581]	12.233 [18.438]	7.899 [10.995]	5.215 [8.348]	9.563 [15.006]	6.552 [6.146]
Constant	-69.216*** [11.450]	-3495.364*** [822.698]	103.698*** [20.886]	93.183*** [13.161]	98.546*** [11.059]	75.208*** [16.175]	83.389*** [8.843]
Sample size	35,196	35,196	35,196	35,196	35,196	35,196	35,196
Number of individuals	6,022	6,022	6,022	6,022	6,022	6,022	6,022
Test of the exclusion restrictions	82.17***	53.33***					
Test of zero restrictions	69.13***	56.04***	97.61***	132.93***	173.79***	58.55***	35.97***
Hausman test for fixed effects			17.64*	49.98***	62.64***	15.80	21.41**
Endogeneity test (F-statistic)			74.86***	20.51***	21.34***	66.01***	23.61***
Cragg-Donald Wald F statistic for weak instruments					30.36		

Notes:

1) The first stage equations have been estimated using an individual fixed effects (FE) estimator and the second stage equations have been estimated using an individual fixed effect instrumental variable (FEIV) estimator.

2) *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

3) Figures reported in square brackets are robust standard errors adjusted for clustering.

4) Test of the exclusion restrictions reports an F-test for the null hypothesis that the coefficients on Age difference 1 and Age difference 2 are jointly zero.

5) Test of zero restrictions reports a test of the null hypothesis that all the coefficients except the constant are jointly zero. This test is computed as an F-test in Stage 1 and a Wald test in Stage 2.

6) The Hausman test for fixed effects reports a Hausman test for the null pooled model estimated by instrumental variables against the fixed effect model estimated by instrumental variables. The test is based on 500 bootstrap repetitions using the "rhausman" command in STATA 15.

7) The Endogeneity test tests the null hypothesis that Working hours and Working hours-squared in the second stage can be treated as being exogenous. Following Wooldridge (2010, pages 352–354), this test is implemented as an F-test that tests the joint significance of the residuals from the two first stage equations when added to each of the second stage models and these second stage models are estimated by a fixed effects estimator.

8) The Cragg-Donald Wald F statistic reported is computed using the "xtivreg2" command in STATA 15.

Table IV: The impacts of working hours on health (Tobit + FEIV)

VARIABLES	0th stage (Tobit)	1st Stage (FE)		2nd stage (FEIV)				
	(4.A1)	(4.B1)	(4.B2)	(4.C1)	(4.C2)	(4.C3)	(4.C4)	(4.C5)
	Working hours	Working hours	Working hours- squared	Physical functioning	Bodily pain	General health	Vitality	Mental health
Age difference 1	-3.304*** [0.304]							
Age difference 2	-0.884*** [0.203]							
<i>Fitted values</i> of Working hours		1.299*** [0.089]	62.802*** [5.276]					
Squared of (<i>Fitted values</i> of Working hours)		-0.011*** [0.001]	-0.451*** [0.063]					
Working hours				6.289*** [1.850]	3.219*** [1.075]	2.469*** [0.846]	4.274*** [1.254]	1.641** [0.642]
Working hours-squared				-0.136*** [0.037]	-0.069*** [0.022]	-0.053*** [0.017]	-0.092*** [0.025]	-0.037*** [0.013]
Age	-0.636*** [0.201]	-0.087 [0.066]	-5.528 [3.893]	-1.642*** [0.271]	-1.152*** [0.180]	-0.962*** [0.143]	-0.929*** [0.186]	-0.347*** [0.106]
School years 7–10	7.204** [3.490]							
School years 11 and over	11.764*** [3.463]							
Married	7.857*** [1.010]	-2.657*** [0.623]	-135.385*** [39.877]	2.770 [2.427]	0.457 [1.422]	0.195 [1.107]	1.951 [1.653]	3.252*** [0.866]
Number of dependent children	-0.065 [0.315]	-0.203 [0.250]	4.452 [17.644]	2.073 [1.409]	0.995 [0.790]	0.746 [0.614]	0.968 [0.963]	0.234 [0.470]
Ownhouse	6.243*** [0.936]	-3.300*** [0.579]	-199.276*** [37.219]	-1.949 [2.232]	-1.065 [1.245]	-1.147 [0.961]	-1.535 [1.488]	-0.353 [0.740]
Inner regional	-1.246 [0.821]	-3.511*** [1.143]	-177.535*** [68.383]	-4.043 [3.567]	-1.705 [1.999]	-1.099 [1.568]	-1.850 [2.422]	-0.251 [1.178]
Outer regional	1.836 [1.246]	-4.000** [1.617]	-172.963* [97.889]	1.658 [5.146]	1.308 [2.864]	-0.167 [2.292]	1.937 [3.410]	1.796 [1.742]
Remote	11.605*** [2.672]	-2.189 [3.226]	-41.102 [199.097]	16.424 [11.457]	10.124 [6.315]	5.049 [5.100]	9.481 [7.807]	4.943 [3.876]
Very remote	11.622** [4.515]	4.492 [6.505]	358.033 [347.541]	26.337 [24.036]	15.200 [13.824]	8.986 [9.743]	16.990 [17.645]	8.840 [7.084]
Constant	31.488** [13.652]	15.094*** [4.938]	738.855** [294.420]	176.251*** [20.564]	137.757*** [13.244]	121.296*** [10.452]	116.217*** [14.141]	96.838*** [7.718]
Sample size	35,196	35,196	35,196	35,196	35,196	35,196	35,196	35,196
Number of individuals	6,022	6,022	6,022	6,022	6,022	6,022	6,022	6,022
Left-censored observations	12,658							
Test of the exclusion restrictions	60.22***	106.22***	75.44***					
Test of zero restrictions	257.93***	70.31***	55.77***	56.39***	90.34***	120.91***	35.48***	27.28***
Hausman test for fixed effects				23.00**	38.11***	47.88***	30.09***	32.37***
Endogeneity test (F-statistic)				63.70***	14.63***	11.94***	42.77***	12.89***
Hausman test of three-stage estimator vs two-stage estimator				8.45**	8.60**	3.44	5.12*	2.11
Cragg-Donald Wald F statistic for weak instruments						12.71		

Notes

1) In the 0th stage, the Working hours is estimated using a Tobit estimator. In the first stage, the equations have been estimated using an individual fixed effects (FE) estimator with the fitted value of working hours and its square included as explanatory variables. In the second stage equations have been estimated using an individual fixed effect instrumental variable (FEIV) estimator.

2) The Hausman test of the three stage estimator vs two-stage estimator tests the null hypothesis of the three stage estimator (Table IV) against alternative hypothesis of the two stage estimator (Table III) using a Hausman test. The estimated significance of the Hausman chi-squared statistic is based on 500 bootstrap repetitions using the "rhausman" command in STATA 15.

3) See Notes 2–8 for Table III.

Table V: Robustness check for FEIV (Using the vacancy rate as an additional instrument)

VARIABLES	1st stage (FE)		2nd stage (FEIV)				
	(5.A1)	(5.A2)	(5.B1)	(5.B2)	(5.B3)	(5.B4)	(5.B5)
	Working hours	Working hours-squared	Physical functioning	Bodily pain	General health	Vitality	Mental health
Age difference 1	-2.284*** [0.188]	-108.059*** [13.089]					
Age difference 2	-2.811*** [0.213]	-145.873*** [14.511]					
Vacancy rate	1.121*** [0.285]	53.390*** [17.734]					
Working hours			5.458*** [1.043]	3.155*** [0.712]	2.481*** [0.597]	4.011*** [0.801]	1.710*** [0.463]
Working hours-squared			-0.103*** [0.019]	-0.058*** [0.013]	-0.047*** [0.011]	-0.078*** [0.014]	-0.035*** [0.008]
Age	1.398*** [0.173]	67.453*** [12.417]	-0.748*** [0.228]	-0.627*** [0.147]	-0.679*** [0.123]	-0.426** [0.176]	-0.159 [0.099]
Married	0.974* [0.576]	71.431* [38.625]	1.649 [1.743]	-0.120 [1.149]	-0.108 [0.968]	1.360 [1.336]	3.063*** [0.796]
Number of dependent children	-0.568*** [0.247]	-13.928 [17.328]	1.913** [0.910]	1.092* [0.579]	0.820* [0.493]	0.970 [0.704]	0.310 [0.393]
Ownhouse	-0.439 [0.509]	-37.269 [34.110]	-1.088 [1.476]	-0.690 [0.940]	-0.959 [0.796]	-1.114 [1.114]	-0.247 [0.635]
Inner regional	-3.985*** [1.145]	-201.727*** [68.350]	-0.601 [2.589]	0.406 [1.633]	0.050 [1.381]	0.129 [1.992]	0.526 [1.083]
Outer regional	-2.502 [1.626]	-90.276 [98.299]	2.937 [3.685]	2.403 [2.325]	0.461 [2.018]	2.822 [2.711]	2.263 [1.583]
Remote	2.859 [3.386]	253.831 [207.956]	10.498 [7.735]	7.389 [4.754]	3.655 [4.311]	6.508 [5.987]	4.123 [3.435]
Very remote	11.630* [6.657]	769.988** [346.594]	12.289 [18.273]	7.947 [10.889]	5.181 [8.428]	9.572 [14.972]	6.453 [6.370]
Constant	-77.154*** [11.723]	-3873.493*** [834.960]	105.101*** [19.887]	94.373*** [12.545]	97.702*** [10.632]	75.435*** [15.401]	80.928*** [8.565]
Sample size	35,196	35,196	35,196	35,196	35,196	35,196	35,196
Number of individuals	6,022	6,022	6,022	6,022	6,022	6,022	6,022
Test of the exclusion restrictions	58.01***	37.83***					
Test of zero restrictions	63.63***	51.40***	99.29***	134.83***	172.86***	59.12***	35.17***
Hansen J test			0.09	0.16	0.13	0.00	1.67
Cragg-Donald Wald F statistic for weak instruments					20.50		

Notes

1) The first stage equations have been estimated using an individual fixed effects (FE) estimator and the second stage equations have been estimated using an individual fixed effect instrumental variable (FEIV) estimator. This Table differs from Table III as it uses an additional instrument, the vacancy rate.

2) The Hansen J test is an overidentification test, and in this case has one degree of freedom.

3) Notes 2–8 for Table III.

Table VI: The impacts of working hours on health (FEIV) (Excluding full-time workers)

VARIABLES	1st stage		2nd stage (FEIV)				
	(6.A1) Working hours	(6.A2) Working hours- squared	(6.B1) Physical functioning	(6.B2) Bodily pain	(6.B3) General health	(6.B4) Vitality	(6.B5) Mental health
Age difference 1	-1.008*** [0.265]	-24.094*** [7.803]					
Age difference 2	-1.152*** [0.284]	-29.090*** [8.297]					
Working hours			24.238** [10.493]	17.716** [7.799]	13.533** [6.264]	16.968** [7.651]	9.881** [4.737]
Working hours-squared			-0.948** [0.377]	-0.684** [0.282]	-0.528** [0.224]	-0.675** [0.275]	-0.398** [0.170]
Age	0.758*** [0.263]	18.363** [7.777]	-0.764 [0.688]	-0.543 [0.505]	-0.593 [0.409]	-0.479 [0.499]	-0.091 [0.311]
Married	-0.513 [0.387]	-13.019 [11.402]	-1.406 [3.892]	-1.806 [3.182]	-1.822 [2.402]	-0.490 [2.977]	1.919 [1.933]
Number of dependent children	0.350 [0.370]	2.764 [11.168]	-4.477 [3.742]	-2.902 [2.690]	-2.921 [2.170]	-3.905 [2.701]	-1.606 [1.714]
Ownhouse	-0.325 [0.347]	-7.261 [9.942]	2.919 [3.320]	0.648 [2.503]	0.535 [1.893]	1.038 [2.423]	0.842 [1.570]
Inner regional	-0.484 [0.582]	-13.481 [15.634]	-3.640 [4.481]	-0.358 [3.424]	-0.407 [2.598]	-0.915 [3.246]	0.960 [2.195]
Outer regional	-0.028 [0.864]	-9.218 [25.072]	-8.899 [7.674]	-4.113 [5.902]	-4.993 [4.415]	-4.813 [5.592]	-1.378 [3.658]
Remote	-3.126 [3.436]	-123.176 [107.826]	-39.756 [31.823]	-27.512 [23.769]	-23.395 [18.458]	-31.041 [23.761]	-17.839 [14.730]
Very remote	-5.878** [2.711]	-182.667** [80.481]	-54.107** [27.335]	-40.029** [18.512]	-34.429** [14.719]	-27.862 [18.224]	-21.327* [11.876]
Constant	-44.863*** [17.217]	-1091.247** [509.258]	115.199** [57.290]	94.791** [42.091]	96.114*** [33.980]	88.228** [41.596]	78.049*** [25.941]
Sample size	16137	16137	16137	16137	16137	16137	16137
Number of individuals	3301	3301	3301	3301	3301	3301	3301
Test of the exclusion restrictions	8.41***	7.41***					
Test of the zero restrictions	14.71***	12.70***	58.28***	84.03***	80.90***	41.79***	18.14*
Hausman test for fixed effects			0.55	3.05	5.63	2.40	5.94
Endogeneity test (F-statistic)			46.94***	22.15***	18.48***	41.82***	20.25***
Cragg-Donald Wald F statistic for weak instruments					9.51		

Notes

1) The first stage equations have been estimated using an individual fixed effects (FE) estimator and the second stage equations have been estimated using an individual fixed effect instrumental variable (FEIV) estimator. This Table differs from Table III because it excludes all respondents who reported that they were full-time workers.

2) Notes 2–8 for Table III.

3) When calculating the Hausman's statistic for fixed effects, Stata reports that the estimated covariance matrix was not of full rank in 4 cases for column (6.B2) and in 1 case for column (6.B3) out of the 500 bootstrap replications, respectively.

Table VII: The impacts of working hours on health (FEIV) (Respondents aged 45–75)

VARIABLES	2nd stage (FEIV)						
	(7.A1) Working hours	(7.A2) Working hours- squared	(7.B1) Physical functioning	(7.B2) Bodily pain	(7.B3) General health	(7.B4) Vitality	(7.B5) Mental health
Age difference 1	-2.282*** [0.236]	-108.227*** [16.904]					
Age difference 2	-2.565*** [0.248]	-136.330*** [17.701]					
Working hours			2.924*** [0.716]	1.686*** [0.587]	1.552*** [0.502]	2.390*** [0.603]	1.006** [0.421]
Working hours-squared			-0.053*** [0.013]	-0.028*** [0.011]	-0.029*** [0.009]	-0.045*** [0.011]	-0.022*** [0.008]
Age	1.028*** [0.216]	50.253*** [16.083]	-0.376 [0.232]	-0.304* [0.173]	-0.524*** [0.156]	-0.163 [0.202]	-0.162 [0.141]
Married	1.620** [0.815]	91.003* [53.582]	-1.714 [1.280]	-2.792*** [1.020]	-1.858** [0.937]	-0.577 [1.143]	2.508*** [0.871]
Number of dependent children	-0.847*** [0.319]	-26.238 [22.000]	1.122* [0.662]	0.780 [0.485]	0.571 [0.467]	0.787 [0.582]	0.172 [0.396]
Ownhouse	-0.811 [0.684]	-56.466 [44.994]	-0.320 [1.062]	0.393 [0.815]	-0.130 [0.757]	-0.187 [0.906]	0.185 [0.656]
Inner regional	-4.652*** [1.361]	-215.190*** [79.441]	0.741 [1.729]	0.628 [1.378]	1.219 [1.150]	1.704 [1.424]	0.779 [1.006]
Outer regional	-2.919 [1.992]	-73.062 [117.803]	3.934 [2.551]	4.678** [2.030]	1.994 [1.731]	3.605* [2.067]	2.052 [1.681]
Remote	1.674 [3.894]	176.520 [244.688]	4.508 [5.029]	4.851 [3.308]	2.085 [3.442]	2.537 [4.411]	2.230 [3.144]
Very remote	10.876 [8.975]	742.896* [415.892]	4.364 [12.881]	0.859 [7.155]	1.618 [7.195]	5.928 [12.839]	4.957 [5.668]
Constant	-50.416*** [14.351]	-2625.706** [1064.690]	89.717*** [18.390]	77.888*** [13.596]	90.272*** [12.339]	63.513*** [16.034]	84.167*** [11.214]
Sample size	25,865	25,865	25,865	25,865	25,865	25,865	25,865
Number of individuals	4,741	4,741	4,741	4,741	4,741	4,741	4,741
Test of the exclusion restrictions	54.14***	31.27***					
Test of the zero restrictions	71.74***	57.99***	70.83***	116.29***	138.48***	24.40***	26.26***
Hausman test for fixed effects			3.85	5.56	20.70**	15.11	4.34
Endogeneity test (F-statistic)			19.03***	4.80***	7.12***	20.38***	10.28***
Cragg-Donald Wald F statistic for weak instruments					25.93		

Notes

1) The first stage equations have been estimated using an individual fixed effects (FE) estimator and the second stage equations have been estimated using an individual fixed effect instrumental variable (FEIV) estimator. This Table differs from Table III because it restricts the sample to those respondents aged 45 years or older and 75 years and younger.

2) Notes 2–8 for Table III.

Table VIII: Robustness check (Reduced form)

VARIABLES	(8.1) Physical functioning	(8.2) Bodily pain	(8.3) General health	(8.4) Vitality	(8.5) Mental health
Age difference 1	-1.254*** [0.183]	-0.900*** [0.209]	-0.487*** [0.169]	-0.743*** [0.165]	-0.057 [0.146]
Age difference 2	-0.208 [0.192]	-0.355 [0.216]	0.002 [0.174]	0.081 [0.173]	0.360** [0.153]
Age	-0.098 [0.167]	-0.138 [0.194]	-0.449*** [0.157]	-0.067 [0.155]	-0.192 [0.138]
Married	-0.427 [0.610]	-1.206* [0.726]	-1.073* [0.560]	-0.271 [0.544]	2.248*** [0.538]
Number of dependent children	0.257 [0.231]	0.113 [0.264]	0.064 [0.207]	-0.229 [0.212]	-0.189 [0.184]
Ownhouse	0.365 [0.537]	0.089 [0.561]	-0.280 [0.451]	0.013 [0.461]	0.308 [0.414]
Inner regional	-1.500* [0.818]	-0.442 [0.882]	-0.283 [0.696]	-0.224 [0.676]	0.740 [0.634]
Outer regional	-1.398 [1.208]	-0.251 [1.328]	-1.480 [1.081]	-0.227 [0.937]	1.120 [0.998]
Remote	-0.101 [2.032]	1.679 [2.085]	-1.254 [1.562]	-1.672 [1.657]	0.181 [1.551]
Very remote	-3.873 [4.647]	-0.150 [4.555]	-2.454 [2.370]	-3.477 [2.711]	-0.494 [2.705]
Constant	86.237*** [11.161]	77.110*** [12.854]	93.191*** [10.395]	68.380*** [10.288]	88.967*** [9.178]
Sample size	35,196	35,196	35,196	35,196	35,196
Number of individuals	6,022	6,022	6,022	6,022	6,022
Test of exclusion restrictions	75.88***	21.53***	20.37***	67.80***	22.22***
Test of zero restrictions	38.17***	29.49***	38.90***	19.86***	7.48***

1) All equations are estimated using the fixed effects (FE) estimator.

2) Test of the exclusion restrictions reports an F-test for the null hypothesis that the coefficients on Age difference 1 and Age difference 2 are jointly zero.

3) Test of zero restrictions reports an F-test of the null hypothesis that all the coefficients except the constant are jointly zero.

4) Notes 2 and 3 for Table III.

Table IX : Robustness check of the reduced form

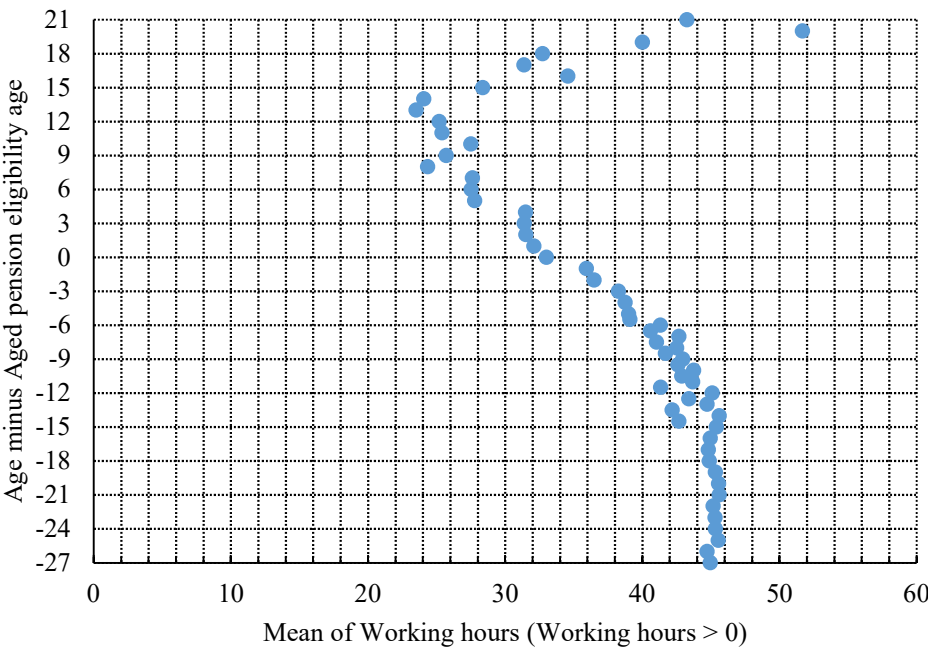
VARIABLES	(9.1)	(9.2)	(9.3)	(9.4)	(9.5)	(9.6)	(9.7)	(9.8)	(9.9)	(9.10)
	Physical functioning	Bodily pain	General health	Vitality	Mental health	Physical functioning	Bodily pain	General health	Vitality	Mental health
AfterReform	-1.514	0.388	-0.290	-0.409	0.038					
	[0.965]	[1.136]	[0.849]	[0.862]	[0.759]					
Age difference 2*AfterReform	-0.283	-0.020	-0.184	-0.141	-0.102					
	[0.185]	[0.219]	[0.180]	[0.179]	[0.157]					
Age difference 2*Cohort1951						-0.104	-0.115	0.049	0.155	0.008
						[0.129]	[0.143]	[0.111]	[0.114]	[0.100]
Age	-0.383***	-0.613***	-0.435***	0.007	0.158**	-0.211***	-0.463***	-0.448***	-0.110	0.129*
	[0.099]	[0.103]	[0.085]	[0.081]	[0.072]	[0.077]	[0.096]	[0.073]	[0.077]	[0.071]
Married	-1.409	-5.113***	-3.574***	-0.807	1.917	-0.172	-1.635*	-0.043	0.065	3.327***
	[1.361]	[1.537]	[1.169]	[1.191]	[1.246]	[0.752]	[0.943]	[0.708]	[0.780]	[0.816]
Number of dependent children	0.079	-0.284	0.235	-0.136	0.359	0.118	0.056	0.344	0.007	-0.116
	[0.623]	[0.649]	[0.536]	[0.499]	[0.413]	[0.311]	[0.373]	[0.270]	[0.296]	[0.271]
Ownhouse	-0.877	1.482	-0.919	1.601	1.510	-0.518	0.794	-0.940	-0.022	0.593
	[1.334]	[1.140]	[0.969]	[1.022]	[0.935]	[0.671]	[0.792]	[0.645]	[0.667]	[0.622]
Inner regional	-0.811	-1.514	-0.399	0.969	0.256	-0.887	-0.959	-0.187	0.241	1.528
	[1.383]	[1.503]	[1.192]	[1.095]	[1.004]	[1.005]	[1.344]	[1.009]	[1.001]	[0.946]
Outer regional	1.984	3.621	-0.825	1.378	1.178	-1.324	-0.247	-2.374	-0.695	0.576
	[2.091]	[3.082]	[1.925]	[1.823]	[2.161]	[1.543]	[2.195]	[1.584]	[1.462]	[1.681]
Remote	2.718	-1.118	-4.191*	-4.508	2.432	4.601	2.460	-4.559*	-1.045	2.492
	[5.048]	[3.409]	[2.191]	[3.734]	[2.053]	[3.701]	[3.343]	[2.759]	[2.199]	[2.486]
Very remote	-8.952	0.746	-7.873	-10.086***	-5.027*	-7.595	0.342	-7.309*	-7.470*	-2.509
	[12.342]	[8.621]	[6.841]	[2.601]	[2.886]	[8.194]	[6.804]	[4.127]	[4.164]	[3.707]
Constant	102.829***	107.615***	93.977***	61.032***	64.480***	94.996***	96.234***	90.825***	68.136***	65.527***
	[5.912]	[6.131]	[5.118]	[4.853]	[4.219]	[4.227]	[5.294]	[3.975]	[4.259]	[3.873]
Sample size	9,065	9,065	9,065	9,065	9,065	16,035	16,035	16,035	16,035	16,035
Number of individuals	1,714	1,714	1,714	1,714	1,714	3,541	3,541	3,541	3,541	3,541
Test of zero restrictions	4.83***	8.08***	5.86***	3.11***	2.98***	2.70***	6.55***	8.63***	0.69	3.51***
F-test H ₀ : all the coef. η are jointly zero	1.34	0.24	0.90	0.35	0.66					

1) All equations are estimated using the fixed effects (FE) estimator.

2) Figures reported in square brackets are robust standard errors adjusted for clustering.

3) Equations (9.1)–(9.5) are estimated respondents in the 1951 Cohort provided they are aged under 65. Equations (9.6)–(9.10) are estimated for respondents in the 1951 and 1952 Cohorts who are under the age of 65 for the pre-reform sample.

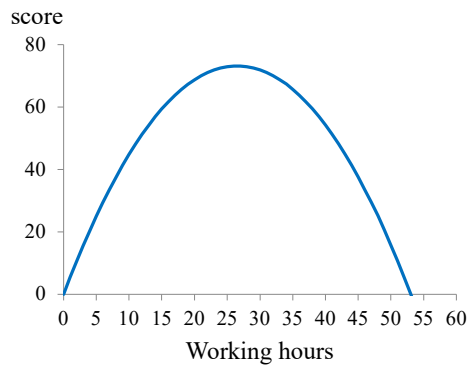
Figure I: The relationship between Average working hours and Distance from pensionable age



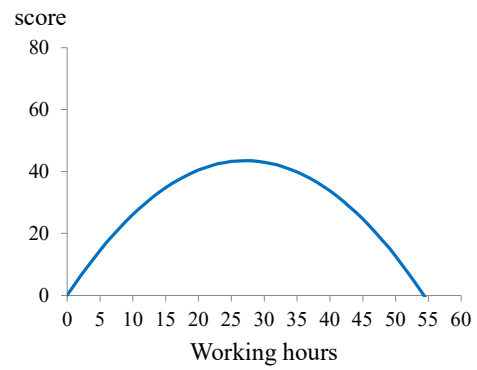
Note:
(1) This figure is constructed by computing the mean of working hours across all waves for all individuals in the relevant group provided the individuals were working.

Figure II : Estimated impacts of working hours on health outcomes

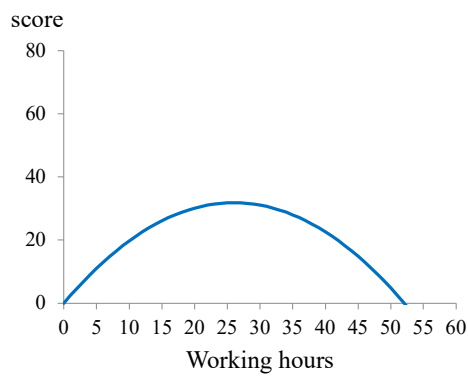
Panel A: Physical functioning



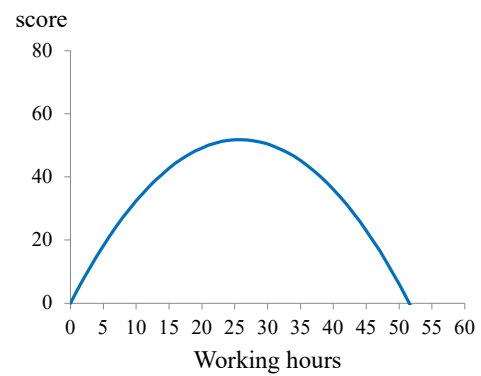
Panel B: Bodily pain



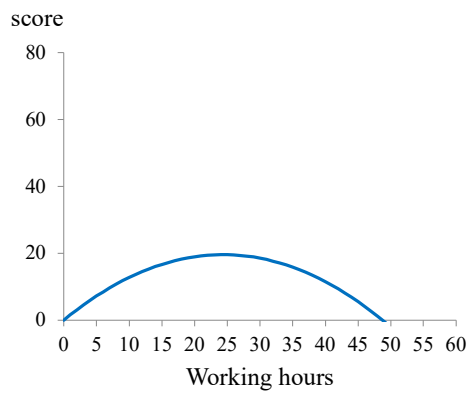
Panel C: General health



Panel D: Vitality



Panel E: Mental health



Note:

(1) In each of the panels, the fitted values of that part of the score that depends on working hours, namely $\alpha_1 WH_{it} + \alpha_2 WH_{it}^2$, computed using the estimated coefficients of *Working hours* and *Working hours-squared* reported in Columns (3.B1)–(3.B5) in Table III are displayed.

Online Supplementary Table O1: Effect of Selection Criterion on Sample Sizes

	Original Number of Observations (Males)	Eliminating observations where the respondent is under the age of 40 at the time of the Wave	Eliminating observations whose pension eligibility age was not identified	Eliminating observations who are unemployed	Dropping observations where the relevant information was not available	Eliminating observations who are working on working hours in the top one percentile	Eliminating observations on age in the top first percentile
Wave 1	6,634	3,642	3,449	3,346	2,933	2,911	2,893
Wave 2	6,222	3,480	3,293	3,218	2,800	2,778	2,764
Wave 3	6,034	3,386	3,210	3,144	2,763	2,748	2,734
Wave 4	5,872	3,295	3,131	3,083	2,754	2,739	2,723
Wave 5	6,029	3,374	3,212	3,156	2,769	2,758	2,737
Wave 6	6,079	3,404	3,241	3,196	2,814	2,803	2,788
Wave 7	6,003	3,347	3,198	3,150	2,721	2,712	2,694
Wave 8	6,021	3,376	3,227	3,177	2,748	2,741	2,713
Wave 9	6,318	3,483	3,330	3,270	2,761	2,757	2,725
Wave 10	6,414	3,536	3,382	3,309	2,956	2,948	2,919
Wave 11	8,357	4,642	4,441	4,346	3,795	3,781	3,746
Wave 12	8,266	4,589	4,398	4,310	3,808	3,798	3,760
Total	78,249	43,554	41,512	40,705	35,622	35,474	35,196