Spillover Effects of Retirement: does health vulnerability matter?

Dominic Byrne[*], Do Won Kwak[†], Kam Ki Tang[‡] and Myra Yazbeck[§]

May 6, 2020

Abstract

The current literature investigating the impact of retirement and the associated spousal spillover effects overlooks the unintended effects of retirement on spouses in vulnerable health, namely spouses with long-term health conditions. In this paper, we fill this gap in the literature and investigate the impact of an individual's retirement on their partner's health outcomes when their partners have long-term health conditions. Given the inherent identification challenges associated with entry into retirement, we exploit an exogenous variation to pension-qualifying age in Australia. Using a Fuzzy Regression Discontinuity Design and data from the Household Income and Labour Dynamics in Australia survey, we find that the retirement of the husband has a positive impact on the wife's Quality Adjusted Life Years (QALY) and other physical and mental health outcomes. Drawing upon the literature on QALY and cost-effectiveness thresholds, we estimate that the dollar value of the husband-to-wife spillovers could be worth somewhere between AUD8,354 and AUD25,062.

JEL Code: I-31; I-19

Keywords: Spillovers; Retirement; Long-term Health Conditions; Regression Discontinuity Design

[*] School of Economics, University of Queensland, Australia; e-mail: d.byrne@uq.edu.au

[†] Graduate School of International Studies, Korea University, Seoul, Korea; e-mail: dwkwak@korea.ac.kr (corresponding author).

[‡] School of Economics, University of Queensland, Australia; e-mail: kk.tang@uq.edu.au

[§] Department of Economics, University of Ottawa, Canada; School of Economics, University of Queensland; e-mail: myazbeck@uottawa.ca

1 Introduction

In the past decades, aging has become one of the most significant factors in setting social policies in developed countries. The rising life expectancy and the declining birthrate are putting pressure on social security funds and thus imperilling their viability. Given that the increasing life expectancy is accompanied by improvement in average population health, raising the qualifying age for pension (and hence retirement) may be a promising avenue to explore. This being said, the suitability of such policy reforms hinges on three assumptions. Specifically, it posits that the reform is primarily based on the population's mean lifespan regardless of its dispersion. It also implicitly assumes that retirement has no negative impact on one's health and that no spillovers are occurring within the family. However, these assumptions are not always satisfied. Some people may not be in a position health-wise to work longer. Besides, retirement can have effects that go beyond the individual sphere of those who retired. Indeed spillovers are ubiquitous within the family; thus, the impact of a reform that increases the qualifying age for pension may not only affect the eligible person but also all their family members, particularly those who are in vulnerable health.

There is a large body of literature investigating the health impacts of retirement; however, the conclusions remain mixed. On the one hand, some of the evidence indicates that retirement has positive effects on a person's subjective health (Johnston and Lee, 2009), life satisfaction (Gorry et al. 2018), psychological well-being (Latif, 2011), physical activity (Kämpfen and Maurer, 2016) and health in general (Coe and Zamarro, 2011). On the other hand, there is evidence that retirement increases the risk of cardiovascular disease, cancer, and metabolic syndrome (Behncke, 2012), as well as obesity (Godard, 2016) and cognitive impairment (Bonsang, Adam and Perelman, 2012). Regardless of whether retirement has positive or negative effects on health, empirical evidence suggests that the impact of retirement on health outcomes cannot be ignored and thus, should be carefully considered in pension reforms. In fact, the health impact of pension reforms could be larger if there are spillover effects from prospective retirees to other family members, such as their spouses.

The interdependence of health outcomes between couples is well documented in the literature[1] however, the studies assessing the impact of an individual's retirement decision on their spouse's health outcomes remain limited. In a recent paper, Muller and Shaikh (2018) analyse the impact of retirement decisions on the spouse's health behaviours such as alcohol consumption and physical activity. The authors find that a person's retirement has a statistically significant negative impact on their spouse's health-related behaviours. However, the authors do not consider the impacts on a population segment that has great importance for

[1] Clark and Etile, 2006; Clark and Etile, 2011; Cutler and Glaeser, 2011; Davillas and Pudney, 2017 among others.

social policy: spouses in vulnerable health. In addition, they do not stratify these effects by gender. Using a gender-based perspective, Bertoni and Brunello (2017) assess spillovers arising from husbands' retirement. They find evidence that the husband's retirement decreases his wife's mental health and increases her stress levels. In contrast, Atalay and Zhu (2018) find that the wife's retirement has a beneficial impact on the husband's mental health, suggesting the presence of gender-based heterogenous effects. Similarly to Muller and Shaikh (2018), the work of Atalay and Zhu (2018) and Bertoni and Brunello (2017) do not consider the potential for heterogeneous impacts for those who are in vulnerable health. Thus, the current literature investigating these spousal spillovers overlooks the unintended effects of retirement on spouses in vulnerable health namely, spouses with long-term health conditions (LTHCs). Indeed, retirement may allow potential caregivers to provide valuable (and needed) support to their frail partners that otherwise would have not been possible. Thus, if retirement has a positive effect on the health outcomes of spouses with LTHCs, and as long as having a spouse with LTHCs does not qualify the partner for early retirement, increasing the qualifying age for pension will only be detrimental for this part of the population.

Motivated by this critical policy issue and building on the work of Muller and Shaikh (2018), Bertoni and Brunello (2017), and Atalay and Zhu (2018), the objective of this paper is to fill this gap in the literature by focusing on a reference group for which spillovers are likely to matter most from a policy perspective: spouses with LTHCs. In such a context, the health impact of a policy reform that is deemed neutral for the average person, may reveal to have a non-negligible negative impact on the vulnerable population. In addition, this paper explores the mechanisms through which spillovers occur. There is a growing interest in uncovering the mechanisms for the health effects of retirement. Most of this literature points towards an increase in healthy habits such as increasing exercising and sleep duration, and reducing smoking and work-related stress (Insler, 2014; Kämpfen and Maurer, 2016; Eibich, 2015) as potential channels through which retirement influences health. This paper complements this literature by examining whether changes in couples' time use after one spouse retires can possibly explain any observed spousal health spillover effects. Furthermore, focusing on the impacts of retirement on spouses with LTHCs allows us to get some insights (or indirect evidence) into the extent to which informal support supplied by partners is valuable.[2] The medical literature emphasizes the importance of the presence of significant others on the well-being of patients as well as pain processing (Younger et al., 2010; Eisenberger, 2012; and Master et al., 2009 among others). While our data do not provide direct information on informal care, our evidence can shed light on

[2]In this paper, we use Household Income and Labour Dynamics in Australia (HILDA) data, which provides information on whether the person is a carer or not; however, it does not specify who is the recipient of the care, nor the type of care provided.

this as a possibility. If retirement has a positive effect on the health outcomes of spouses with LTHCs, then identifying and quantifying these effects are of significant policy relevance and value for future research. Therefore, from this perspective, we contribute to the debate on the importance of informal care.

This paper is closely related to the paper by Muller and Shaikh (2018), which also focuses on the impact of spillovers of one's retirement on the spouse's health outcomes. Nevertheless, this paper differs in three respects. First, this paper is interested in a specific population segment: people who have LTHCs.[3] Second, given the population segment of interest, this paper focuses on a specific measure of health outcome that was not previously considered in the literature: Quality Adjusted Life Year (QALY), as compared to self-reported general health status or mental health. While self-reported health status is widely used in the literature and has its merits, it has its well-known limitations and does not account for the quality of life–a dimension of crucial importance when focusing on a population with LTHCs. In addition, health policy deliberations and cost-benefit analyses are often cast in terms of qualify of life.[4] By focusing on the quality of life measure, we can estimate the dollar value of a specific policy intervention that changes the well-being of a targeted population. Third, using each component of health outcome as a distinctive outcome measure, this paper investigates the dimensions of health in which these spillover effects manifest most, providing insight on the aspects of health that retirement is most likely to affect.

Assessing the impact of an individual's retirement on their partner's health outcomes presents some identification challenges. First, retirement decisions are far from exogenous because a person's retirement decision can be affected by their spouse's health status. Second, spurious correlations between a person's retirement decision and their partner's health outcomes could result from either the presence of assortative matching or the couple's sharing the same environment. Specifically, if both spouses chose each other based on their preference for unhealthy lifestyles, then it is possible that the retirement decision and the partner's quality of life are affected by the couple's lifestyle choice. Further, living in the same house implies that they are exposed to the same physical environment. To address these issues, we exploit a quasi-experimental setting related to the Australian age pension program, which is equivalent to old-age social security in other countries.[5] Specifically, we exploit the 2009 policy change that generated a variation in the pension quali-

[3]A more nuanced literature on spousal spillovers by Mavromaras et al. (2014) and Atalay and Barrett (2014) shows that the effect of retirement on own health could differ significantly by gender and income. The angle from which this paper considers spillovers can be perceived as related to these two papers since our analyses probe upon how retirement impacts change depending upon characteristics of the spouse (vulnerability), rather than focusing on the average effect.

[4]In assessing whether a drug should be included in the Pharmaceutical Benefits Scheme and the received subsidy, the Department of Health (DoH) of the Australian government often referred to the impact of the assessed drug on quality of life.

[5]Australia's age pension system is very different from the pension systems in many other industrial countries. It is essentially an old-age social security program from which Australians are eligible to receive a regular retirement income (the 'age pension') from

fying age across cohorts. Given that the qualifying age is determined only by the year of birth, it provides an arbitrary cutoff that serves as a useful source of exogenous variation in the retirement status. Using data from the Household Income and Labour Dynamics in Australia (HILDA) survey from 2002 to 2014, we assess whether the change of the qualifying age is a credible source of variation in retirement status. Our empirical evidence confirms that a change in the pension-qualifying age induces a shift in people's retirement year. Building on this finding, we then identify spillover effects of an individual's retirement on the spouse's health outcome using a Fuzzy Regression Discontinuity Design (FRDD). Our results suggest that a husband's retirement has a positive impact on his wife's health outcome when the later has LTHCs. Specifically, we find that for women who have LTHCs, their husband's retirement will, on average, improve their QALY by 0.9 standard deviations, their physical health by 0.82 standard deviations, and their mental health by 0.88 standard deviations. To provide insights on the relative importance of the effects, we draw upon the literature on QALY and cost-effectiveness thresholds and estimate the dollar value of the husband-to-wife spillovers. Our back-of-the-envelope calculations suggest that these effects could be worth somewhere between AUD8,354 and AUD25,062.

The remainder of this paper is organized as follows. Section 2 describes the data and offers an overview of the Australian policy context. Section 3 introduces the empirical framework and the identification strategy. Section 4 reports estimation results and robustness checks, and Section 5 concludes.

2 Data

In this section we first provide a brief description of the data used and the sample restrictions. We then describe the health outcomes of interest as well as the treatment variable used in the estimation.

2.1 Sample Construction

We employ data from waves 2–14 (2002–2014) of the HILDA survey. The survey collects wide-ranging information on personal well-being, labour market status, family life and time use. To assess spillover effects, an individual can be included in our analysis only if information is available for their partner as well. Each observation in HILDA is assigned a household identifier, and a relationship grid records how each

the government as long as they have reached a certain age threshold and passed a means test. The amount of retirement income received by an eligible person depends on their current financial situation (and hence the means test) but is unrelated to the number of years they were in the labour force.

household member is related to the other members in their household (for example, partner, mother, father). This enables us to match partners with each other in order to investigate their spillover effects.

The following sample restrictions are imposed on the data set. We restrict our sample to individuals within +/- 10 years of their respective eligibility cutoff.[6] For some individuals, the HILDA survey does not explicitly enquire about their retirement status. If the respondent indicates elsewhere on the survey that they are in the workforce, then we can infer that the individual is not retired. On the other hand, if the respondent reports that they are not in the labour force, but their retirement status is unclear, then it is not guaranteed that this person is retired. Such observations have been discarded. Individuals who have never worked are excluded from the sample as well. After excluding observations with core data missing, we have a baseline sample of 11,943 observations corresponding to 2,266 matched couples.

2.2 Health Outcome Variables

For health outcome measures, we deploy the QALY measure as it accounts for both lifespan and life quality. It has been widely used to measure the value of health outcomes (Prieto and Sacristán, 2003), especially in cost-effectiveness analysis (see, for example, Neumann et al., 2014; Nimdet et al. , 2015; and Woods et al., 2015). Because we are focusing on a population with LTHCs, considering their quality of life is crucial. The measure takes a value between 0 and 1, with the value of 1 indicating perfect health, and the value of 0 indicating death. Throughout the paper, we also present the change in standard deviations for ease of interpretation.

We also investigate other physical and mental health indicators derived from the Short Form 36 (SF-36) survey, because entry into retirement could have impact physical health and mental health differently. The SF-36 health dimensions can be used to form standardized scores (between 0 and 100) over eight health categories, which Table 1 summarizes. The first four categories are summed to generate the aggregate physical health score and the last four categories are summed to generate the aggregate mental health score.

2.3 Treatment Variable

Our main treatment variable of interest is the indicator of retirement. In the HILDA survey, respondents are asked whether they have retired completely from the workforce. Possible responses are “yes,” “no”, or “never in workforce.” As we are concentrating on the impact of entry into retirement, we consider an

[6] The choice of this window is compatible with the works of Lavecchia (2018), Chen et al. (2017) and Battistin et al. (2009) among others. More importantly, the results for our parameter of interest are not sensitive to the choice of the window bandwidth albeit, the estimates have larger standard errors with a smaller bandwidth.

individual to be retired only if they respond with a "yes" to this question.

From a financial perspective, the age pension is of high importance in the retirement decision, especially for those with limited private savings. The age pension is up to about 57% of the average income for working couples and is higher for single-income couples; thus, the prospect of receiving the age pension upon reaching the eligible age represents a financial incentive to enter into retirement. This reasoning is supported by the first-stage estimation results presented in Section 3.1. Furthermore, the F-test results for the first-stage estimation using the subsamples with low incomes confirm that people respond to this financial incentive to retire (Table 10).

Table 2 presents summary statistics for our overall sample. Given the key role of LTHCs in retirement spillover health effects, we provide separate columns for females with and without LTHCs. Approximately 30% of females have LTHCs. As expected, females with LTHCs report lower health scores on average. Additionally, females with LTHCs are more likely to be retired and to come from households with substantially lower income. The average number of persons in the household is only modestly lower for the LTHC subsample but statistically significant.

2.3.1 Pension Age and Eligibility

In Australia, pension is not a universal welfare; only people who meet certain criteria are eligible for the full or partial pension.[7] However, it is a substantial social program that affects about 70% of the elderly population.[8] On this basis, it is reasonable to suspect that reaching the pension age alters labour market incentives for a substantial portion of the population, potentially creating an exogenous discontinuity in retirement rates. Thus, if we can demonstrate that there are no sudden changes in individual or household characteristics at this cutoff, and that there are no other relevant policies coming into effect, then we can associate any discontinuous changes in health outcomes with retirement induced by the exogenous eligibility milestone.

Table 3 shows how the qualifying age for the pension varies depending on the date of birth and gender. As mentioned earlier, to be eligible for pension, people must also pass both an income test and an asset test. There are different income and asset thresholds for people of different social and health conditions. The asset test is more complicated because of differential treatments of various types of assets, such as owner-occupied home, superannuation investments (see Appendix for details), business assets, and even funereal investments. Unfortunately, data on household assets is collected only infrequently in the HILDA survey. Consequently,

[7]The criteria include: age; residence status; income; and assets.

[8]As of 2013, there were about 2.4 million people receiving Age Pensions or Veterans Affairs Pensions.

partitioning the sample on the basis of household assets is difficult. On the other hand, household income is surveyed annually, and thus allows us to account for the impact of income in two ways. Firstly, we include income as a control variable, because it appears to change discontinuously at the pension-qualifying age. Secondly, we undertake sensitivity analysis to ascertain whether the responsiveness of retirement rates to the pension-qualifying age differs by income level. The findings corroborate the notion that pension eligibility is the source of the discontinuity in retirement rates at the cutoff.

3 Empirical Framework

Our empirical model for the impact of a spouse's retirement on the respondent's health outcome is given by

$$H_{it} = \beta_0 + \beta_1 R_{jt} + \gamma_0 A_{jt} + \boldsymbol{X}_{1it}\gamma_1 + \boldsymbol{X}_{2ht}\gamma_2 + u_{it}, \tag{1}$$

where H_{it} is the health status of respondent i in year t; R_{jt} is a retirement status indicator of spouse j in year t; A_{jt} is age; $\boldsymbol{X}_{1it}$ includes observed characteristics variables for i in year t; $\boldsymbol{X}_{2ht}$ includes household level observed characteristics variables for household h of the couple in year t; and u_{it} is an unobserved error term. We are mainly interested in obtaining a causal estimate of β_1 and identifying the impact of the spouse's retirement on the respondent's health outcome.

The identification of the impact of the spouse's retirement decision on the respondent's health presents identification issues. First, the respondent's poor health status may require a full-time carer that could catalyze the spouse's retirement (reverse causality). Second, it is possible that unobserved confounding factors (e.g., inheritance or lottery winnings) affect both a spouse's retirement as well as the respondent's health status, generating a spurious positive correlation. Third, partners often share common characteristics and tastes (i.e., assortative matching); thus, if both partners have the same taste for unhealthy lifestyles, this may affect both the retirement decision and health of the couple. Given these identification issues, estimating equation (1), using a naive linear regression framework will lead to a biased and inconsistent estimate of β_1 (the spousal spillover effects). To address these issues, we exploit exogenous variations obtained from *the age pension reforms* that lead to delays in retirement for younger cohorts. That is, by using the Regression Discontinuity Design (RDD) approach, which exploits the changes in the rule for the qualifying age for pension, we can estimate the causal effect of a person's retirement on their spouse's health outcomes.

3.1 Discontinuity in retirement by the Pension Age

The validity of the RDD method relies on the existence of an assignment/forcing variable (i.e., age) that is a main driver of treatment (i.e., spouse's retirement status) for a substantial proportion of individuals in the absence of simultaneous shocks. Our identification strategy relies on the assumption that individuals just above a certain age threshold are more likely to receive the treatment than individuals just below the threshold. Therefore, we can attribute a discontinuity in the outcome variable (i.e., the respondent's health outcome) at the threshold of the forcing variable to the change in the treatment status (Lee and Lemieux, 2010). That is, the discontinuity in the outcome variable is entirely induced by the arbitrary age differences between the treated and the non-treated.

We test our approach by examining whether the forcing variable is statistically relevant to the change in the treatment status. This aspect of the RDD approach is similar to the IV approach in earlier studies (e.g., Coe and Zamarro, 2011; Rohwedder and Willis, 2010; Insler, 2014) in the sense that both approaches use the eligible age as a source of exogenous variation in the retirement status. Therefore, we test for weak IV issues through the first-stage estimation using the following equation:

$$R_{jt} = \alpha_0 + \delta_1 \cdot 1(A_{jt} \geq agecutoff) + \delta A_{jt} + \boldsymbol{X}_{1it}\alpha_1 + \boldsymbol{X}_{2ht}\alpha_2 + u_{it}, \tag{2}$$

where A_{jt} is age and $1(A_{jt} \geq agecutoff)$ is an indicator of pension qualifying age and is an IV. Typically, a strong IV implies that the age threshold generates a significant discontinuous increase in the retirement probability; that is, δ_1 is statistically significant (F-statistic greater than 10). In our context, this implies that a large number of individuals delay their retirement because the pension-qualifying age has increased.

Compared with the standard IV approach typically implemented with a parametric specification on the eligible age variable, the RDD approach allows more flexibility for estimating specifications in various dimensions (Lee and Lemieux, 2010). For our sample period, women's eligible age increased from 60 to 67 and men's from 65 to 67. We thus calculate the centered age, which is equal to each individual's actual age minus the eligible age that corresponds to their date of birth. Using the centered age as the forcing variable means that all observations now have the same eligible age cutoff of zero. Accordingly, for the RDD approach, we replace the usual first stage estimation with the following equation:

$$\begin{aligned} R_{jt} = {} & \delta_1 \cdot 1(ctrA_{jt} \geq 0) + \delta_2 \cdot 1(ctrA_{jt} \geq 0) \cdot ctrA_{jt} + \delta_3 \cdot 1(ctrA_{jt} < 0) \cdot ctrA_{jt} \\ & + \boldsymbol{X}_{1it}\alpha_1 + \boldsymbol{X}_{2ht}\alpha_2 + u_{it} \end{aligned} \tag{3}$$

where R_{jt} is equal to one if spouse j is retired in year t, and zero otherwise; $ctrA_{jt}$ is the centered age (i.e. A_{jt} minus the cutoff age) for j in year t; $1(ctrA_{jt} \geq 0)$ is an indicator variable equal to one if j has reached the pension-qualifying age, and zero otherwise; and $1(ctrA_{jt} < 0)$ is an indicator variable equal to one if j has not reached the pension-qualifying age, and zero otherwise. The two interaction terms with the forcing variable, $ctrA_{jt}$, allow for different slopes around the cutoff. Table 4 presents the results of F-tests for the coefficient of δ_1 to verify if there is a statistically significant discontinuous increase in the retirement probability at the cutoff of the forcing variable.

There are two cases in which the qualifying age does not bring about a statistically significant discontinuity in the retirement status. Firstly, the change in the average retirement status at the eligibility cutoff for females is not sufficiently strong to serve as an IV (the F-statistic $= 5.76 < 10$). This suggests that the prospect of receiving the age pension may not generate the same incentive effect if there is another income earner in the household. This reasoning is supported by the findings of Atalay and Barrett (2015) that, following an increase in the female qualifying age for pension, the labour force participation rate of affected single females increased by 13 percentage points, whereas that of affected married and single females together increased by only 8 percentage points.[9] Secondly, reaching the age cutoff is not a sufficiently strong predictor when wives have not yet retired (see Table 4). This is in line with Warren's (2013) conclusion that "the most common pattern of retirement for Australian couples is one in which the wife retires before the husband."

Figure 1 graphically conveys the first-stage estimation results of equation (3). The left-hand panel shows a visible discontinuity using all males in the sample, whereas the right-hand panel for females fails to show a clear discontinuity at the centered cutoff age. The increase in retirement rates at the cutoff for men is estimated to be 10.6% and it is statistically significant at the 1% level, as Table 4 shows. The left-hand panel of Figure 1 shows that the quadratic specification for the forcing variable fits the local averages very well. A linear specification for the forcing variable seems reasonable within +/-10 years of the cutoff. Using the results from Figure 1, in the baseline specification we restrict our sample to observations who are within +/-10 years of being eligible for the age pension, and we impose the linearity of the forcing variable.[10] Within the centered age window, this leaves us with 11,943 observations for 2,266 unique couples.

[9] Likewise, Warren (2013) examines joint retirement decisions of couples in Australia and finds evidence that women were less likely to be influenced by monetary considerations compared to men.

[10] We have tested the baseline specification with 5 and 7 year windows around the cutoff. The coefficients are similar to the 10 year results in terms of direction and magnitude, but with lower precision. Therefore, given the relatively small sample size, we have retained a 10 year window. Results for the 5 and 7 year windows are available upon request.

3.2 Test of assumptions for RDD

3.2.1 Local Random Assignment

For the treatment of retirement to be locally random around the cutoff of the forcing variable, there should be no other shock simultaneously occurring and affecting individuals around the cutoff. This allows us to attribute any discrete jump in the outcomes at the cutoff to retirement. We can test this condition by assessing whether there are any discontinuities for the observed covariates around the eligibility cutoff. If this is the case, our identification strategy would be questionable, as any discontinuities in health outcomes could perhaps be attributed to discontinuities in these variables. Thus, we test if our data meets this requirement by re-estimating equation (3) but replacing R_{jt} with a number of observed covariates.

A visual inspection of Figures 2-4 shows that there are no discontinuities at the cutoff of the covariates examined with the exception of household income and male LTHCs. However, auxiliary regressions (available upon request) indicate that none of these discontinuities are statistically significant at the 10% level.[11] Furthermore, we note that any changes in household income and male LTHCs could potentially be the result of retirement. Because retirement precludes wage income, and because pension income is typically less than wage income, the income drop at the cutoff could be the result of retirement. However, we cannot confirm the source of this income drop in our dataset. Given this finding, we have included household income as a control variable in all of our estimates. Including income does not have any impact on the results. Similarly, males appear to report fewer LTHCs at the cutoff, implying that the average effect of retirement may be positive, but the observed decline is not statistically significant. Again, the inclusion of the husband's LTHC status as a control variable does not have any impact on our findings. Also, noteworthy is the lack of a discontinuity in female retirement rates at their male partner's eligibility cutoff. In other words, we do not observe simultaneous retirements of couples in our sample. This finding removes any concerns that the results observed in this paper are driven by the female partner's own entry into retirement.

3.2.2 Density of the Forcing Variable

Another test available to justify the RDD approach is that individuals must not be able to affect the realization of the forcing variable (i.e., no manipulation of the forcing variable). In our study, this takes the form of centered age (calculated as age less pension-qualifying age). Figure 5 shows no sudden change in the density of centered age around the qualifying age cutoff. The density of age (and centered age) is very

[11] These regressions consist of OLS estimations for equation (3) with the retirement outcome variable replaced with variables in Figures 2 to 4.

smooth across all ages in the sample.

3.3 Estimation

Our baseline is a two-stage RDD estimation, where the second stage is defined as follows:

$$H_{it} = \beta_0 + \beta_1 R_{jt} + \beta_2 \cdot 1(ctrage_{jt} \geq 0) \cdot ctrage_{jt} + \beta_3 \cdot 1(ctrage_{jt} < 0) \cdot ctrage_{jt} + \boldsymbol{X}_{1it}\gamma_1 + \boldsymbol{X}_{2ht}\gamma_2 + u_{it}, \tag{4}$$

where j indicates spouse, i indicates respondent, h indicates household, and t indicates year; $\boldsymbol{X}_{1it}$ includes observed characteristics variables for respondent i in year t; $\boldsymbol{X}_{2ht}$ includes household level observed characteristics variables for household h in year t; and, again, the two interaction terms allow for different slopes around the cutoff. As a supplement to this linear model, we also use a local linear non-parametric model and higher-order polynomial models as well.[12] All estimations in this paper use robust standard errors clustered at the individual level.

4 Results

This section presents the results for the spillover effects from husband to wife. As stated earlier, we focus on husband-to-wife spillovers because the strength of the instrument for wives was found to be statistically insignificant. To put things in perspective, we begin by evaluating the impact of the husband's retirement on the wife's QALY, using the full sample. Given that we are interested in spouses who are in vulnerable health, we then focus on wives who have an LTHC. We also provide results for a number of other physical and mental health constituents, which are components used to construct QALY.

4.1 Retirement Impact on QALY

The findings regarding the impact of a partner's retirement on the wife's QALY, for all women as well as women with and women without LHTC, are displayed in Table 5. For each sample, we have implemented two specifications: one parametric and one non-parametric.

Our baseline specification assumes a linear piecewise functional form for the forcing variable. In each estimated baseline equation, the F-statistic is significantly greater than 10, indicating that the eligibility for the age pension is a strong source of exogenous variation for entry into retirement. The estimated spillover effects for the full sample suggest that a husband's entry into retirement leads to an improvement in the

[12] This is implemented using *rdrobust* package in Stata, see Calonico et al. (2017) for details.

wife's QALY of approximately three quarters of a standard deviation or 12% of the average QALY, which is statistically significant at the 5% level. This positive and statistically significant finding for the full sample appears to be driven by the subsample of wives with LTHCs, for whom the spillover effects are slightly less than one standard deviation or approximately 17% of the average QALY. The estimate for the subsample without LTHCs is also positive, but smaller in magnitude and statistically insignificant. This finding does not appear to be a consequence of sample size, as this subsample is larger than the subsample with LTHCs.

To test whether the results obtained are an artifact of the linear specification we imposed, we experimented with a local linear non-parametric specification. This alternative specification shows that the effect on QALY for the full sample remains positive and comparable in magnitude to the baseline specification, but is now significant only at the 10% level. Meanwhile, the coefficient for the LTHC subsample is virtually unchanged and retains its significance at the 5% level. The effect on QALY for the subsample without LTHCs remains insignificant. It is important to note that F-statistics remain above 10 in all these cases, thus reflecting the absence of weak instrument issues.[13]

4.2 Estimating the cost effectiveness of the intervention

Our findings in section 4.1 allow us to consider the cost of an intervention that aims to improve the quality of life of partnered women with LTHCs. The hypothetical intervention is to offer the husband a pension before he reaches the qualifying age so that he can retire earlier to take better care of his wife. As explained before, pension rates in Australia are calculated based on a range of factors. For simplicity, we consider the scenario that the husband is eligible for the maximum basic pension rate and supplement and that he is living together with his wife. Under this scenario, the pension payout or the cost of the intervention is equal to AUD17,420 per year.

A widely cited "World Health Organisation recommended" cost-effectiveness threshold is one to three times of GDP per capita. The idea of using a GDP per capita based threshold originated from a report of the Commission on Macroeconomics and Health (WHO, 2001), and the "one to three times of GDP per capita" threshold was subsequently set out in another WHO-affiliated publication (Hutubessy, Chisholm, & Edejer, 2003). In 2018, Australia's GDP per capita was equal to AUD73,930 (World Bank data), so one QALY is worth between AUD73,930 and AUD221,790. From Table 5, we know that the husband's retirement is expected to improve his wife's QALY by 0.113. Therefore, the estimated benefit from spousal health

[13]Additionally, we have tested a quadratic functional form as well as higher-order functional forms for the forcing variable. The coefficients on retirement status in the second-stage estimation remain positive and are of similar magnitude, but the F-statistics for the forcing variable in the first-stage no longer exceed the traditional rule of thumb for a strong instrument. Results are available upon request.

spillover due to the intervention is worth between AUD8,354 and AUD25,062, and the cost of intervention (AUD17,420) falls well within this range.

The literature on cost-effectiveness thresholds stresses that one size of threshold cannot fit all circumstances. Policymakers need to consider not only the cost-to-benefit ratio of a given intervention but also the budget constraint and the opportunity cost (Bertram et al., 2016). Bearing this in mind, the current literature also suggests that an intervention that costs about two to three times of GDP per capita is at least worth considering. Based on this benchmark, the hypothetical policy intervention considered here is of an 'acceptable' cost-to-benefit ratio.

4.3 Retirement impact on the wife's other health constituents

In this section, we extend our analysis to examine the impact of the husband's retirement on various components of the SF-36 health measures for the wife. While components in the SF-36 are used to construct QALY, looking at the outcomes for individual health components can provide additional information about what aspects of health may be most responsive to spousal spillovers. Given the impact of retirement on QALY is concentrated in the subsample of women experiencing LTHCs, we focus exclusively on this subsample in this section.[14]

The results in Table 6 show the husband's retirement has a positive and statistically significant effect on the aggregate physical health of the wife. This appears to be driven by an improvement in the Physical Functioning score, and an improvement in the Bodily Pain score (i.e., a reduction in Bodily Pain), both of which are significant at the 5% level. The husband's entry into retirement also appears to benefit the wife's General Health score at the 10% level of significance. Furthermore, the magnitudes of these coefficients are economically significant and are in the range of one standard deviation. On the other hand, Role-Physical health does not appear to be significantly affected. The structure of Role-Physical questions could be responsible for this. Answers to these questions must be dichotomous; thus, it is possible that the husband's retirement is beneficial to the wife's Role-Physical health to a certain extent but insufficient to change the responses to these particular questions.

The husband's retirement also has a broad impact on the wife's mental health. Social Functioning is positively affected at the 5% level, while Role-Emotional Health and Vitality are positively affected at the 10% level. In each instance, the improvement in health scores is slightly less than one standard deviation. The

[14]For the subsample without LTHCs, which has valid first-stage with no weak IV problem, the only health outcome that significantly increased upon partner retirement was vitality, at the 10% level. Interestingly, the coefficient for vitality is virtually identical in both subsamples. This could suggest that a partner's vitality can benefit from their spouse's retirement, even if they do not have an LTHC. Results are available upon request.

directions and magnitudes of these coefficients are comparable to the findings of Mavromaras et al. (2014), who use similar Australian data to ascertain the effects of own retirement, rather than spousal retirement.

The results in Table 6 indicate that no single component drives overall spousal health improvement. Rather, it appears that improvements in various health components are responsible for the positive spillover effects detected in both aggregate physical and mental health.

4.4 Mechanisms

To understand why a husband's entry into retirement has discernible impacts on women with LTHCs, we have investigated a range of mechanisms that may transmit the retirement effect.

4.4.1 Time use changes

First, we must establish which activities are affected by a partner's retirement and consider why these changes may differ in the presence of LTHCs. For comparison we also include results for the subsample where wives are not experiencing any LTHCs. HILDA provides numerous time use variables that could potentially reveal some of the mechanisms for the positive health effects of partner retirement. We focus on three forms of time use that are most likely to be relevant to partner spillovers: housework, household errands, and outdoor tasks. We define combined time use as the sum of these three components. The results of the impacts of the husband's retirement on his time use for these activities are presented in Table 7. Upon retiring, an increase in time allocated to domestic work of approximately 10 hours is observed at the 10% level of significance, but only if the wife has an LTHC; otherwise, entry into retirement does not lead to any significant changes in the allocation of domestic work time. Individually, only household errands are significantly affected.

To verify whether this mechanism is transmitting the health spillover effect, we attempt a mediation analysis. Having revealed that combined time use changes discontinuously at the cutoff, we now control for it in our original specification. If its inclusion materially reduces the observed benefits of retirement, this implies that time use is responsible for the positive effect of retirement on the wife's health. We perform this only for regressions of the relevant sample: namely, wives with LTHCs. Table 8 presents a side-by-side comparison of the results with and without controlling for husband time use.

Although a decrease of almost 20% is observed for the coefficient on Physical Functioning, the remaining coefficients are largely unchanged after controlling for time use. Even slight increases occur for the coefficients on QALY, Bodily Pain, General Health and Social Functioning. Thus, only the difference in coefficients for the wife's Physical Functioning could potentially be viewed as an improvement arising from their husband's time allocation. This suggests that these forms of time use do not capture most of the health benefits from the husband's retirement. There may be other channels such as the intangible benefits of companionship that time use variable does not capture.

4.4.2 Possible substitution in time use between husband and wife

Another possible mechanism behind the spillovers is that the division of household labour changes upon retirement. By calculating the total time spent on housework, errands, and outdoor tasks for both the husband and the wife, we can measure the time-use shares between them. This allows us to investigate whether there is any substitution effect, in which the wife reallocates tasks to her husband. Table 9 presents the results of our evaluation of the substitution effect both as a percentage change and as a change in absolute values. None of the observed effects are precisely measured but, interestingly, the direction of the substitution effect differs between the two subsamples. In the case of the wife experiencing an LTHC, the imprecise results hint that the husband's entry into retirement may potentially lead to reallocation in the share of tasks away from the wife and towards the husband. Otherwise, if the wife is not experiencing an LTHC, the results hint that the wife accounts for a larger share of household tasks following her husband's retirement. The substitution effects have also been estimated as absolute hourly changes. However, given that the effects are not precisely measured, there is no point in discussing the quantitative results.

Overall, there is no evidence indicating that the redistribution of domestic work between the couple could be a mechanism behind the positive spillovers of the husband's retirement on the wife with an LTHC.

4.5 Robustness Tests

In Section 3, we provided two sets of tests to examine the validity of our identification strategy: (i) the tests of discrete jumps around the cutoff for observed covariates, and (ii) the test of density manipulation around the cutoff. These tests show no violation of the assumptions that justify the RDD method. In this subsection, we provide three more tests as further evidence to justify our methodology. More specifically, we have considered two falsification tests and a placebo test.

4.5.1 Falsification tests

The first test applies female age eligibility rules to the male sample. Given that males and females have historically been subject to different pension eligibility regimes, it would be a violation of the identification strategy if males responded significantly to the female eligibility rules. To achieve this, we calculate male centered age using the eligible age for a female with the same date of birth. The first stage becomes strongly insignificant (the F-statistic is equal to 1.66), indicating that males' retirement decisions are not influenced by the rule on females (see Appendix).

The second falsification test considers various income groups. The age pension acts as a financial incentive to enter retirement. However, people with high incomes, and thus large private savings, are not eligible for the public pension (part or full). If this is the case, then we would expect the discontinuity to be smaller (or perhaps indiscernible) for individuals with higher per capita incomes. Using per capita income, we perform two subsamples estimations. One sample consists of observations from the lowest three income quartiles (i.e., 75th percentile and below), and another consists of observations from the highest quartile (i.e., at or above the 75th percentile). In the latter case, we acknowledge that the subsample is quite small, and that the lack of observations could have a bearing on the results. That being said, the size of this subsample is only marginally smaller than the subsample of wives with LTHCs in Tables 5 and 6, where the F-statistics are strongly significant. Table 10 presents the results from the first-stage estimations. It is apparent that reaching the qualifying age generates a discontinuity in retirement rates only for those on lower incomes.

Considering that the incentive to retire upon reaching the pension qualifying age appears weak for those on high incomes, it is prudent to re-estimate the main models of this paper with these observations omitted. Table 11 presents the second-stage results for the effect of retirement on health, restricting the sample to those observations at or below the 75th percentile. As per the earlier analysis, we focus on the subsample of wives with LTHCs. The coefficients remain statistically significant and similar in magnitude to those in the baseline specification (0.113 for QALY, significant at the 5% level in Table 5; 79.456 for Aggregate Physical Health and 78.971 for Aggregate Mental Health, significant at the 10% level in Table 6).[15]

4.5.2 Placebo test

Another robustness check consists of dropping the retired males in the sample and assigns placebo-retirement at random such that the proportion of observations assigned the retirement status is equal to the observed proportion of retirement during the period of analysis (48%). We implement the baseline specification with

[15] As a robustness check, we also performed the same exercise, but partitioned the sample at the median (i.e. at or below the 50th percentile) as well as other percentiles (80th, 90th) as partition cutoffs. We still get strong effects for low income groups regardless of cutoff values. For instance, in the case of using the median cutoff, the first stage F-statistic for the below-median remained above the conventional rule of thumb of 10. For all health outcomes considered, the coefficients remained similar in magnitude compared to those from the 75th cutoff, but the coefficients remained statistically significant for only a few health outcomes such as for Physical Functioning and Bodily Pain. For other outcomes, the coefficients remained similar in magnitude, albeit insignificant. For the above-median subsample, the first stage F-statistic is 7.08, indicating a weak IV problem.

these false-retirement observations instead of the true-retirement observations, to verify whether there are any statistically significant effects. Table 12 shows that the instrument strength is very weak under the placebo assignment. This supports the hypothesis that the actual eligibility for the age pension serves as the incentive for retirement.

5 Conclusion

This paper contributes to the literature on retirement health impacts by expanding the evidence on intra-household spillovers with a focus on the husbands who have a wife with LTHCs. It does so by employing a Fuzzy Regression Discontinuity Design framework, which exploits exogenous variation in the retirement status induced by the Australian pension eligibility rules. Positive health effects of a husband's retirement are observed only if the wife is experiencing an LTHC. These benefits are broad and include improvement in quality adjusted life years as well as a range of other health constituents. The QALY results allow us to generate an indicative dollar value of a retirement intervention, which is comparable to the cost of the intervention in the Australian context.

In an attempt to explain how these effects are transmitted, we uncover evidence that a husband's household time use increases significantly upon retirement, especially if his wife has an LTHC. However, controlling for these mechanisms in our specification does not remove the significance of the partner retirement effect for any health outcomes, nor does it lead to any material decreases in the coefficients. Thus, while there is some evidence that husbands spend more time on household tasks upon retiring, there is insufficient evidence to conclude that time use plays a major role on the positive health effects.

The implication of these findings is that informal care, even if not explicitly identified as such, has a role to play in positive intra-household health externalities, which are likely to ease prospective public healthcare expenditure burdens. Formal long-term health care is costly, but informal care provided by family members can potentially serve as a partial substitute. In Australia, the current carer allowance is less than 15% that of the retirement pension. For people who are capable of providing informal care to their family members but need to work for financial reasons, it may be cost-effective for the government to provide a much higher carer allowance to encourage them to become full-time carers.

References

Atalay, K., & Barrett, G. F. (2014). The causal effect of retirement on health: New evidence from Australian pension reform. *Economics Letters, 125*, 392–395.

Atalay, K., & Barrett, G. F. (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), 71–7(1), 87.

Battistin, E., Brugiavini, A., Rettore, E., & Weber, G. (2009). The retirement consumption puzzle: Evidence from a regression discontinuity approach. American Economic Review, *American Economic Review , 99*(5), 2209–2226.

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal*, 75 (299), 493–517.

Behncke, S. (2012). Does retirement trigger ill health? *Health Economics, 21*, 282–300.

Bertoni, M., & Brunello, G. (2017). *Pappa Ante Portas*: The effect of the husband's retirement on the wife's mental health in Japan. *Social Science and Medicine,175*, 135–142.

Bertram, M. Y., Lauer, J. A., De Joncheere, K., Edejer, T., Hutubessy, R., Kieny, M. P., & Hill, S. R. (2016). Cost-effectiveness thresholds: pros and cons. *Bulletin of the World Health Organisation,* 94(12), 925–930.

Bonsang, E., Adam, S., & Perelman, S. (2012). Does retirement affect cognitive functioning. *Journal of Health Economics, 31*, 490–501.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). rdrobust: Software for regressionÐdiscontinuity designs. *The Stata Journal 17*(2), 372–404.

Caltabiano, M, Campolo, M. G., & Di Pino, A. (2016). Retirement and intra-household labour division of Italian couples: A new simultaneous equation approach *Social Indicators Research 128*, 1217–1238.

Chen, Q., Deng, T., Bai, J., & He, X. (2017). Understanding the retirement–consumption puzzle through the lens of food consumption–Fuzzy regression discontinuity evidence from urban China. *Food Policy, 73*, 45–61.

Clark, A., & Etile, F. (2006). Don't give up on me baby: Spousal correlation in smoking behaviour. *Journal of Health Economics, 25*(5), 958–978

Clark, A., & Etile, F. (2011). Happy house: Spousal weight and individual well–being. *Journal of Health Economics, 30*(5),1124–1136

Coe, N. B., & Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of Health Economics, 30*(1), 77–86.

Cutler, D. M., & Glaeser, E. L. (2010). Social interactions and smoking. In Research findings in the economics of aging (pp. 123–141). University of Chicago Press..

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

Eibich, P., (2015). Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of Health Economics, 43*, 1–12.

Eisenberger, N. I., (2012). The pain of social disconnection: examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience, 13*(6), 421.

Godard, M. (2016). Gaining weight through retirement? Results from the SHARE survey. *Journal of Health Economics, vol. 45*, 27–46.

Gorry, A., Gorry, D., and Slavov, S. N. (2018). Does retirement improve health and life satisfaction? *Health Economics, 27*(12), 2067–2086.

Hutubessy, R., Chisholm, D., & Edejer, T. (2003). Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. *Cost effectiveness and resource allocation, 1*(1), 8.

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

Johnston, D. W., & Lee, W (2009). Retiring to the good life? The short-term effects of retirement on health.*Economics Letters, 103*, 8–11.

Kämpfen, F., & Maurer, J. (2016), Time to burn (calories)? The impact of retirement on physical activity among mature Americans. *Journal of Health Economics, 45*, 91–102.

Latif, E. (2011). The impact of retirement on psychological well-being in Canada. *The Journal of Socio-Economics, 40*, 373–380.

Lavecchia, A. M. (2018). Do“catch-up limit” raise retirement saving? Evidence from a regression discontinuity design. *National Tax Journal, 71*(1), 121–154.

Lee, D. S. & Lemieux, T. (2010). Regression discontinuity designs in economics.*Journal of Economic Literature, 48*, 281–355.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies, 60*, 531–542.

Master, S. L., Eisenberger, N. I., Taylor, S. E., Naliboff, B. D., Shirinyan, D., and Lieberman, M.D. (2009). A picture's worth: Partner photographs reduce experimentally induced pain. *Psychological Science, 20*(11), 1316–1318.

Mavromaras, K., Richardson, S., & Zhu, R. (2014). *Age pension age eligibility, retirement and health outcomes in Australia* (NILS Working Paper Series No. 201). National Institute of Labour Studies, Flinders University, Adelaide, Australia.

Muller, T., & Shaikh, M. (2018). Your retirement and my health behaviour: Evidence on retirement externalities from a fuzzy regression discontinuity design *Journal of Health Economics, 57*, 45–59.

Neumann, P. J., Cohen, J. T., & Weinstein, M. C. (2014). Updating cost-effectiveness–the curious resilience of the $50,000-per-QALY threshold. *New England Journal of Medicine,371*,(9) 796–797.

Nimdet, K., Chaiyakunapruk, N., Vichansavakul, K., & Ngorsuraches, S. (2015). A systematic review of studies eliciting willingness-to-pay per quality-adjusted life year: Does it justify CE threshold? *PLOS One, 10*(4), Art. no. e0122760.

Prieto, L., & Sacristán, J.A. (2003). Problems and solutions in calculating quality-adjusted life years (QALYs). *Health and Quality of Life Outcomes 1*, 80.

Rohwedder, S., & Willis, R. J., (2010). Mental retirement. *Journal of Economic Perspectives 24,*(1), 119–138.

Warren, D. (2013). Retirement decisions of couples: The impact of spousal characteristics and preferences on the timing of retirement. Melbourne Institute Working Paper No. 41/13.

Woods, B., Revill, P., Sculpher, M., & Claxton, K. (2016). Country-level cost-effectiveness thresholds: Initial estimates and the need for further research. *Value in Health, 19*(8), 929–935.

World Health Organization (2001). Macroeconomics and health: investing in health for economic development: report of the Commission on Macroeconomics and Health.

Younger, J., Aron, A., Parke, S., Chatterjee, N., & Mackey, S. (2010). Viewing Pictures of a Romantic Partner Reduces Experimental Pain: Involvement of Neural Reward Systems. *PLoS ONE 5*(10), Art. no. e13309.

Figure 1: Discontinuity in male and female retirement rates at eligibility cutoff

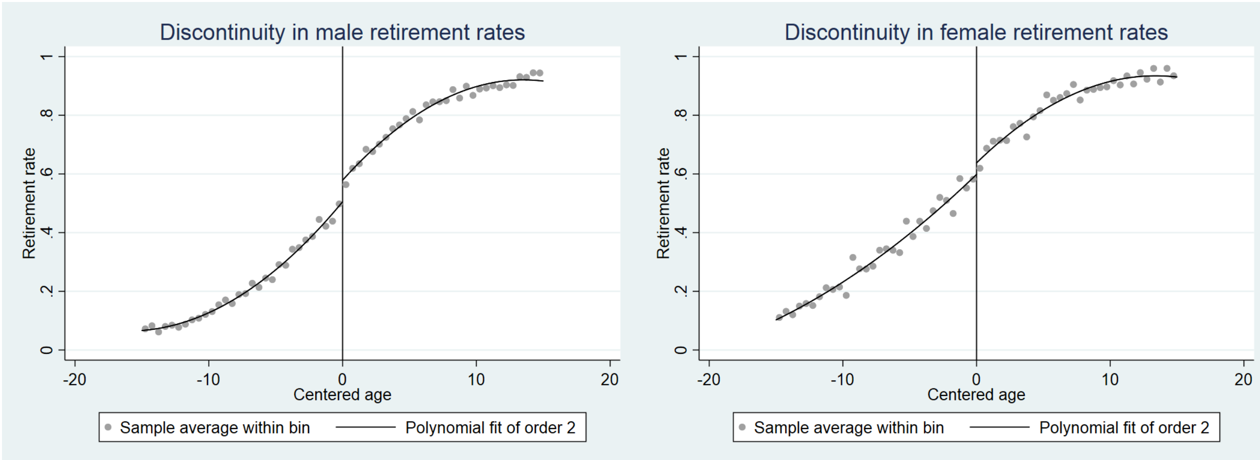


Figure 2: Continuity in household covariates at eligibility cutoff

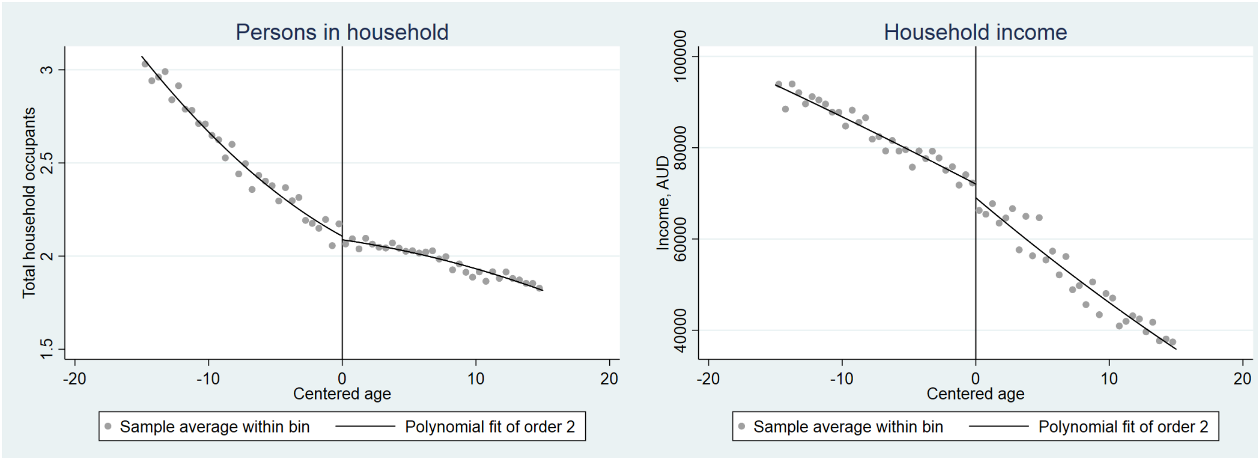


Figure 3: Continuity in long-term health status at eligibility cutoff

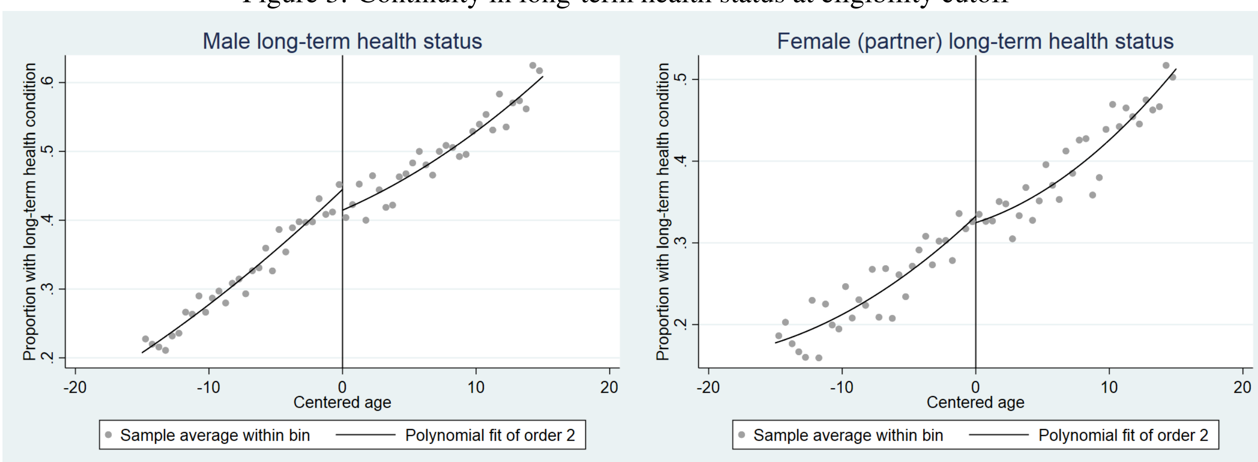


Figure 4: Continuity in female partner covariates at male eligibility cutoff

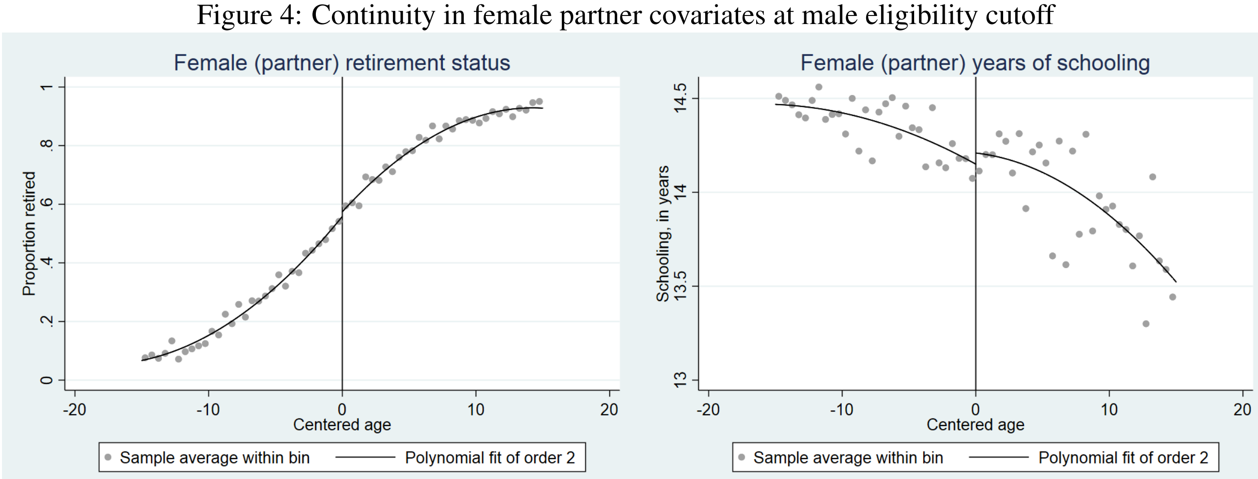


Figure 5: Density of the forcing variable

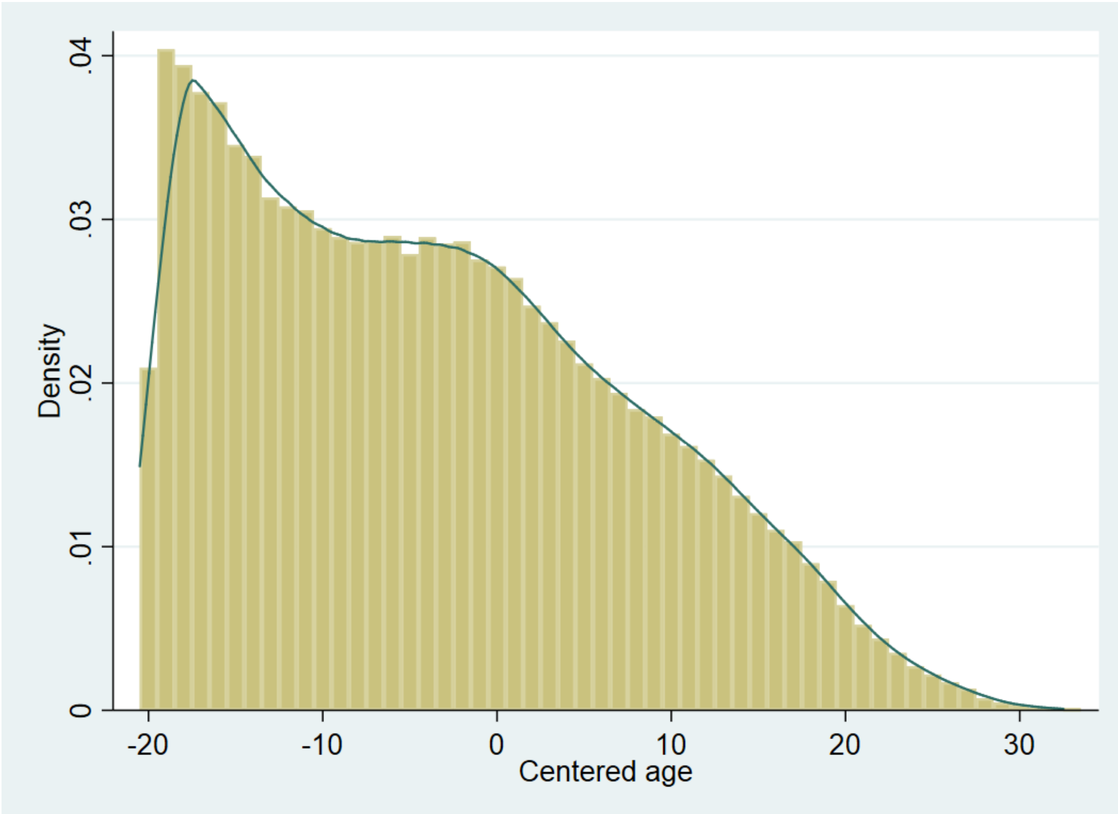


Table 1: SF-36 health measures

Health Category	Description
Physical Functioning (PF)	Measures the extent to which health limits a list of activities
Role-Physical (RP)	Examines how physical health problems affect regular daily activities
Bodily Pain (BP)	Measures the severity of pain and the extent to which it interferes with normal activities
General Health (GH)	Measures health perceptions
Social Functioning (SF)	Measures the extent to which health interferes with social activities
Role-Emotional (RE)	Examines whether emotional problems have limited regular daily activities
Vitality (VT)	Measures how often the respondent feels full of life and energy, and how often they feel tired or worn out
Mental Health (MH)	Measures frequency of nervousness, feeling down, calmness, peacefulness and happiness

Table 2: Summary statistics

	All Females		Females without LTHC		Females with LTHC		Difference in means	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	t-stat	p-value
QALY	0.75	0.12	0.78	0.10	0.67	0.12	46.82	0.00
Physical Functioning (PF)	76.43	22.78	83.70	16.85	60.09	25.72	50.72	0.00
Role Physical (RP)	72.89	39.22	84.04	31.40	47.66	43.26	45.45	0.00
Bodily Pain (BP)	66.80	24.68	74.02	20.68	50.56	25.19	49.38	0.00
General Health (GH)	66.40	21.76	73.06	17.37	51.38	23.14	50.19	0.00
Aggregate Physical Health	283.17	91.98	315.14	68.18	210.45	97.55	57.71	0.00
Social Functioning (SF)	82.43	23.78	88.37	18.97	69.10	27.76	38.31	0.00
Role Emotional (RE)	84.31	32.40	90.19	25.58	70.94	41.12	25.94	0.00
Vitality (VT)	60.95	19.78	65.62	17.48	50.47	20.63	38.68	0.00
Mental Health (MH)	75.88	16.67	78.56	14.99	69.84	18.57	25.03	0.00
Aggregate Mental Health	304.14	76.98	323.10	62.00	261.09	89.36	37.68	0.00
Retired	0.53	0.50	0.46	0.50	0.67	0.47	-22.04	0.00
Age	60.86	7.33	60.16	7.31	62.41	7.14	-16.42	0.00
Household Income (AUD)	70,236	71,595	83,065	69,724	65,767	83,493	11.43	0.00
Persons in Household	2.20	0.99	2.46	0.93	2.35	0.80	7.27	0.00
Housework	6.29	6.94	5.91	6.40	7.16	7.96	-8.32	0.00
Errands	4.44	4.56	4.21	4.28	4.98	5.11	-7.90	0.00
Outdoor tasks	8.76	9.71	8.65	9.70	9.02	9.73	-1.94	0.05
Combined time use	19.44	15.04	18.73	14.67	21.04	15.74	-7.44	0.00
Number of Observations	11,943		8,259		3,684			

Table 3: Eligibility for the age pension in Australia

Birth cohort	Qualifying age	
	Male	Female
Born before June 30, 1935	65	60
July 1, 1935~Dec 31, 1936	65	60.5
Jan 1, 1937~June 30, 1938	65	61
July 1, 1938~Dec 31, 1939	65	61.5
Jan 1, 1940~June 30, 1941	65	62
July 1, 1941~Dec 31, 1942	65	62.5
Jan 1, 1943~June 30, 1944	65	63
July 1, 1944~Dec 31, 1945	65	63.5
Jan 1, 1946~June 30, 1947	65	64
July 1, 1947~Dec 31, 1948	65	64.5
Jan 1, 1949~June 30, 1952	65	65
July 1, 1952~Dec 31, 1953	65.5	65.5
Jan 1, 1954~June 30, 1955	66	66
July 1, 1955~Dec 31, 1956	66.5	66.5
Born after Jan 1, 1957	67	67

Table 4: First stage estimations

Sample	All husbands	Husband (wife without LTHC)	Husband (wife with LTHC)	Husband (wife retired)	Husband (wife non-retired)	All wives
Instrument $1(ctrage_{jt} \geq 0)$	0.106***	0.085***	0.149***	0.094***	0.028	0.050**
	(0.018)	(0.021)	(0.031)	(0.026)	(0.027)	(0.021)
F-statistic	36.240	16.322	22.658	13.542	1.082	5.760
Number of observations	11,943	8,259	3,684	5,818	5,218	9,241

Note: Robust standard errors clustered at the individual level are in parentheses. ***, **, * indicate significance at the 1%, 5%, and 10% level respectively.

Table 5: The impact of the husband's retirement on the wife's QALY

	Full sample		Subsample: wives with LTHCs		Subsample: wives without LTHCs	
	Baseline	Non-parametric	Baseline	Non-parametric	Baseline	Non-parametric
Retired	0.092**	0.079*	0.113**	0.109**	0.054	0.043
	(0.043)	(0.041)	(0.057)	(0.055)	(0.055)	(0.050)
Change in standard deviations	*0.761*	*0.655*	*0.907*	*0.877*	*0.538*	*0.428*
Number of observations	11,314	11,314	3,455	3,455	7,859	7,859
Instrument Strength (F-statistic)	35.760	46.458	24.305	29.709	14.823	19.653

Note: Robust standard errors clustered at the individual level are in parentheses. Non-parametric estimates have been generated using the *rdrobust* command in Stata. Raw QALY values range between 0 and 1 and standardised QALY values have a standard deviation of 1. **, * indicate significance at the 5%, and 10% level respectively.

Table 6: Retirement spillover effects from male to female spouse with long-term health conditions

PHYSICAL HEALTH	Physical Functioning (PF)	Role-Physical (RP)	Bodily Pain (BP)	General Health (GH)	**Aggregate Physical Health**
Retired	28.899**	16.067	23.736**	21.021*	79.456*
	(12.574)	(18.563)	(11.844)	(11.023)	(45.327)
Change in standard deviations	*1.124*	*0.371*	*0.942*	*0.908*	*0.815*
Number of Observations	3,635	3,600	3,653	3,586	3,513
Instrument Strength (F-statistic)	23.232	22.848	23.136	21.716	21.996

MENTAL HEALTH	Social Functioning (SF)	Role-Emotional (RE)	Vitality (VT)	Mental Health (MH)	**Aggregate Mental Health**
Retired	25.659**	35.126*	16.650*	2.547	78.971*
	(12.893)	(19.203)	(9.182)	(7.648)	(40.829)
Change in standard deviations	*0.924*	*0.854*	*0.807*	*0.137*	*0.884*
Number of Observations	3,684	3,593	3,661	3,660	3,574
Instrument Strength (F-statistic)	22.656	23.136	23.136	23.136	23.814

Note: Robust standard errors clustered at the individual level are in parentheses. Aggregate Physical Health is calculated as the sum of the four physical health constituents, while Aggregate Mental Health is calculated as the sum of the four mental health constituents. Standardised health scores have a mean of 50 and a standard deviation of 1. ***, **, * indicate significance at the 1%, 5%, and 10% level respectively.

Table 7: The impact of male partner retirement on activities with potential spillover effects

Wives with LTHCs	Housework	Errands	Outdoor tasks	Combined time use
Retired	2.353	4.005*	3.367	10.038*
	(2.856)	(2.129)	(3.574)	(5.837)
Change in standard deviations	0.296	0.783	0.346	0.638
Number of Observations	3,628	3,591	3,646	3,544
Instrument Strength (F-statistic)	25.503	23.523	26.112	25.806

Wives without LTHCs	Housework	Errands	Outdoor tasks	Combined time use
Retired	0.755	3.495*	-0.510	3.895
	(2.894)	(2.117)	(4.709)	(6.472)
Change in standard deviations	0.118	0.817	-0.053	0.266
Number of Observations	8,205	8,168	8,269	8,039
Instrument Strength (F-statistic)	20.430	18.576	18.576	21.068

Note: Robust standard errors clustered at the individual level are in parentheses. Combined time use is calculated as the some of the three time use variables. Time use is measured on a weekly basis. ***, **, * indicate significance at the 1%, 5%, and 10% level respectively.

Table 8: Side-by-side comparison of coefficients

Health Outcome	Original Equation	Controlling for husband time use
QALY	0.113**	0.131**
Physical Functioning (PF)	28.899**	23.308**
Role Physical (RP)	16.067	20.036
Bodily Pain (BP)	23.736**	27.076**
General Health (GH)	21.021*	23.182**
Aggregate Physical Health	79.456*	85.177*
Social Functioning (SF)	25.659**	27.606**
Role-Emotional (RE)	35.126*	35.145*
Vitality (VT)	16.650*	16.803*
Mental Health (MH)	2.547	4.440
Aggregate Mental Health	78.971*	83.587**

Note: ***, **, * indicate significance at the 1%, 5%, and 10% level respectively.

Table 9: Substitution effect between husband and wife household time use

	Subsample with LTHC	Subsample without LTHC
Substitution effect (percentage point change)	0.067	-0.102
	(0.165)	(0.205)
Substitution effect (absolute values in hours)	*5.767*	*-2.898*
Number of observations	3,185	7,390
Instrument Strength (F-statistic)	32.149	17.724

Note: Robust standard errors clustered at the individual level are in parentheses. A positive coefficient indicates that the husband's share of the household time use is increasing, while a negative coefficient indicates that the husband's share of the household time use is increasing. ***, **, * indicate significance at the 1%, 5%, and 10% level respectively.

Table 10: First stage regressions, comparison by income percentile

Change in retirement rates, by percentile	0~75th percentiles	75 ~100th percentiles
$1(ctrage_{jt} \geq 0)$	0.130***	0.025
	(0.019)	(0.033)
Observations	9,659	3,221
F-statistic	45.024	0.548

Note: The subsample at or below 75th percentile contains households with per capita income less than or equal to AUD 62,804, and the subsample at or above 75th percentile contains households with per capita income greater than or equal to AUD 62,804. *** indicates significance at the 1% level.

Table 11: Retirement spillover effects from husbands to wives with long-term health conditions (0~75th percentile)

	QALY	Aggregate Physical Health	Aggregate Mental Health
Retired	0.092*	71.304*	69.200*
	(0.051)	(42.561)	(38.684)
Number of observations	2,879	2,931	2,984
Instrument Strength (F-statistic)	27.248	24.305	26.112

Note: Robust standard errors clustered at the individual level are in parentheses.

Table 12: Falsification test assigning placebo-retirement

	QALY	Aggregate Physical Health	Aggregate Mental Health
Retired	-4.280	1140.011	1089.420
	(34.754)	(3449.021)	(2885.717)
Number of observations	7,674	7,788	7,906
Instrument Strength (F-statistic)	0.014	0.116	0.144

Note: Robust standard errors clustered at the individual level are in parentheses.

Online Appendix

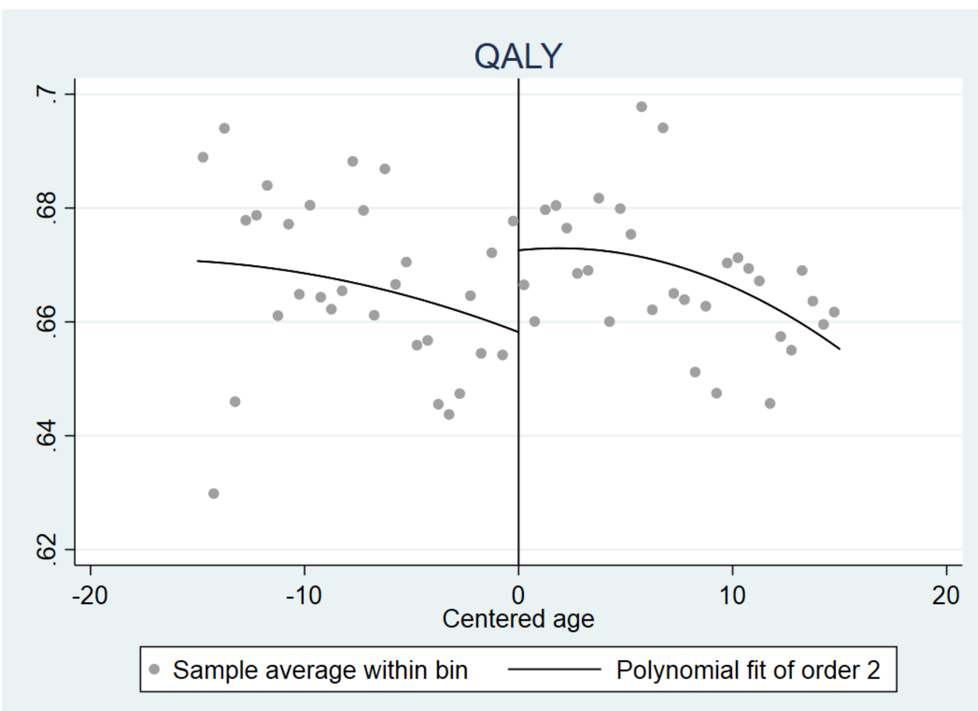
Graphs of Outcome Variables

The discontinuity in health outcomes at the eligibility cutoff can be observed visually. Figures A.1-A.11 show changes in various health outcomes by the centered age for the subsample of wives with LTHCs.

The Wald estimate is equivalent to the IV estimate; $\beta_{1,y|x}$ can be constructed as the ratio of the OLS estimate of $\beta_{1,y|z}$ and the OLS estimate of $\beta_{1,x|z}$ where y denotes the health outcome, z denotes the indicator of centered age (i.e., $z = 1(centerage >= 0)$), and x denotes the retirement status. Thus, $\beta_{1,y|x}$ is the effect of the retirement status on health, $\beta_{1,y|z}$ is the effect of pension eligibility on a health outcome, and $\beta_{1,x|z}$ is the effect of pension eligibility on the retirement status.

Consequently, the effect of the retirement eligibility on those who retired, $\beta_{1,y|x}$, is much stronger than the effect indicated by Figures A.1–A.11. These figures effectively show $\beta_{1,y|z}$ rather than $\beta_{1,y|x}$. Because $\beta_{1,x|z}$ is strictly less than 1 (not every individual retires upon becoming eligible for a pension), $\beta_{1,y|x}$ will be significantly larger than $\beta_{1,y|z}$.[16]

Figure A.1 QALY



[16]These results are comparable to several other RDD papers that remark on the Treatment on Treated (TOT) effect, which is represented by $\beta_{1,y|x}$. For more on the relationship between TOT effect and Intent to Treat (ITT), see a useful discussion in Eibich (2015).

Figure A.2 Aggregate Physical Health

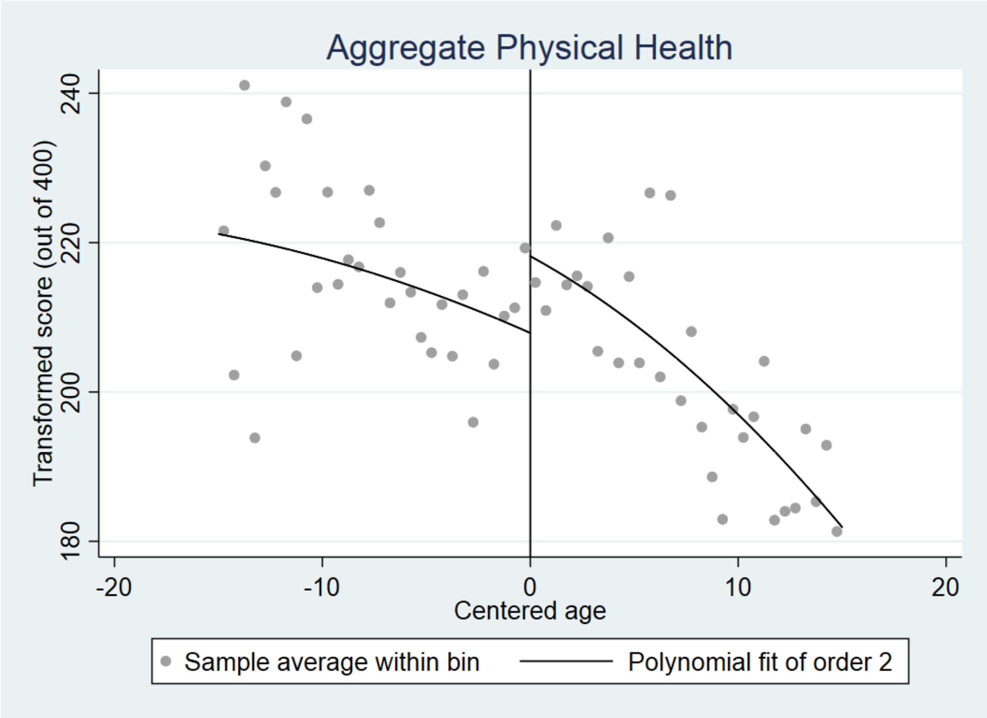


Figure A.3 Aggregate Mental Health

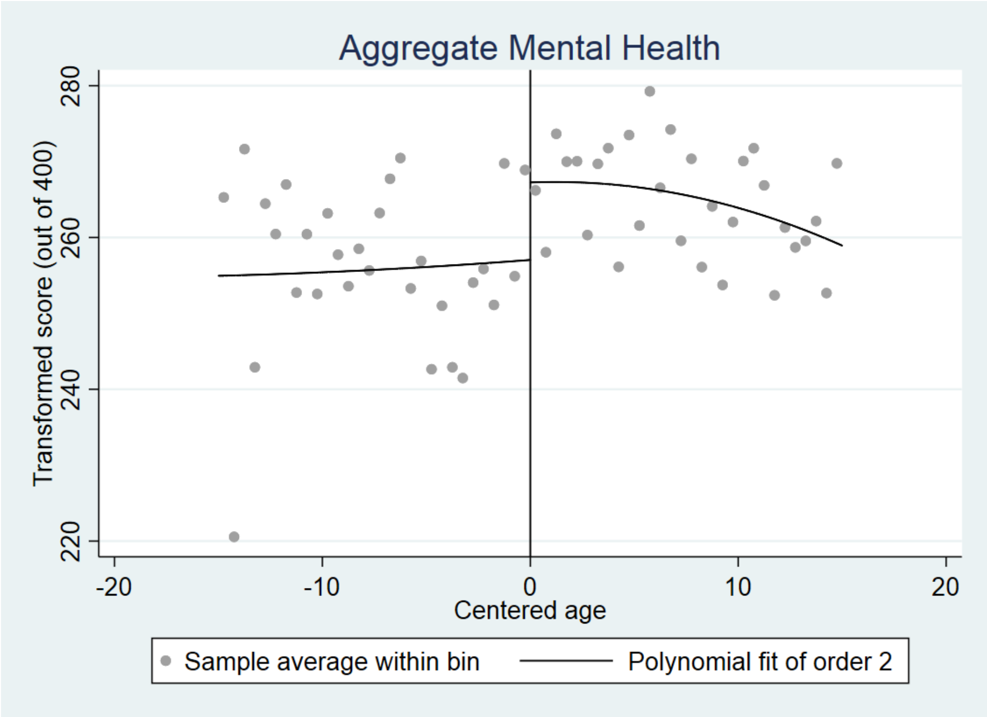


Figure A.4 Physical Functioning

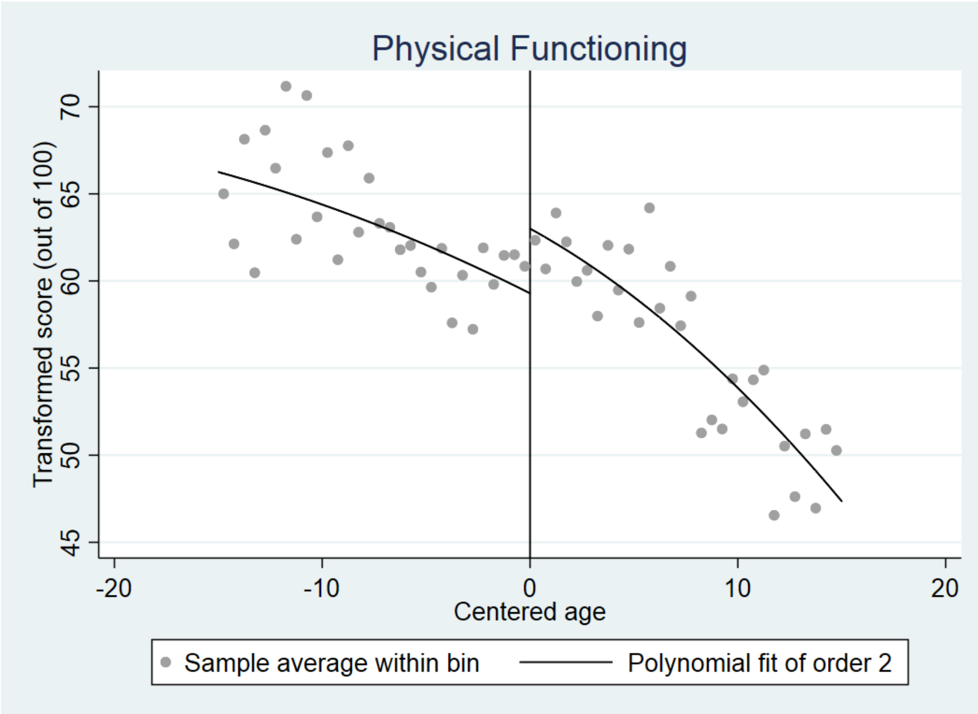


Figure A.5 Role-Physical

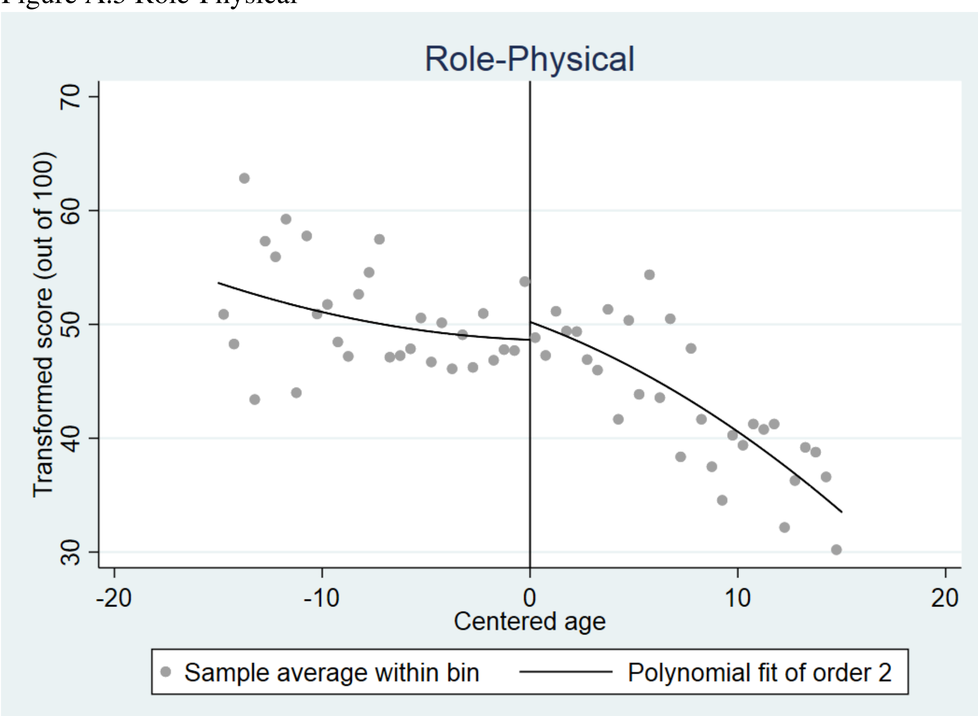


Figure A.6 Bodily Pain

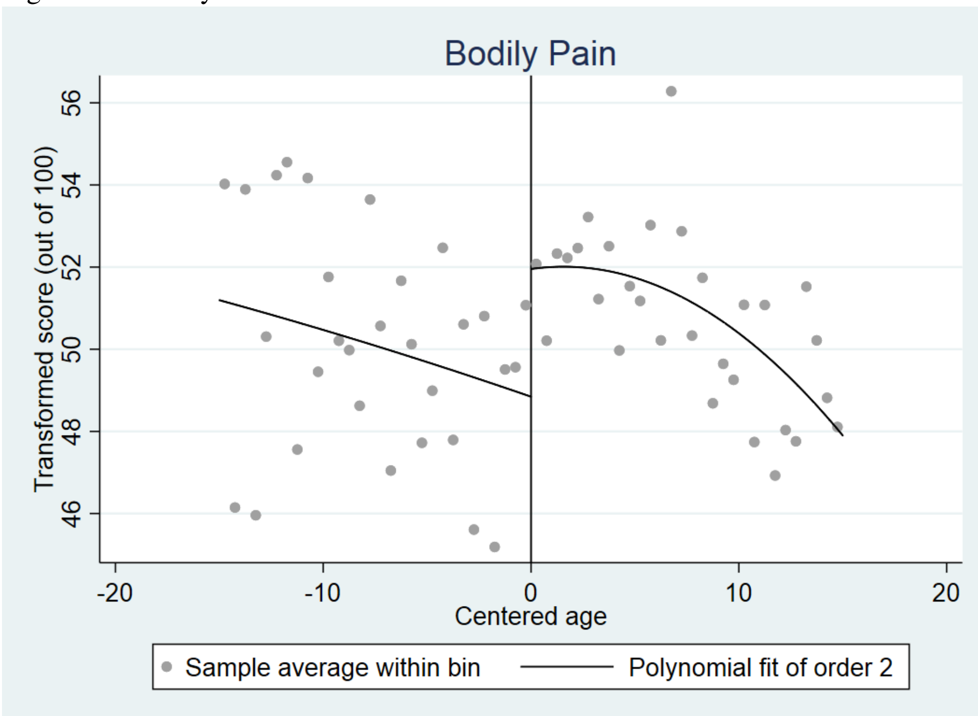


Figure A.7 General Health

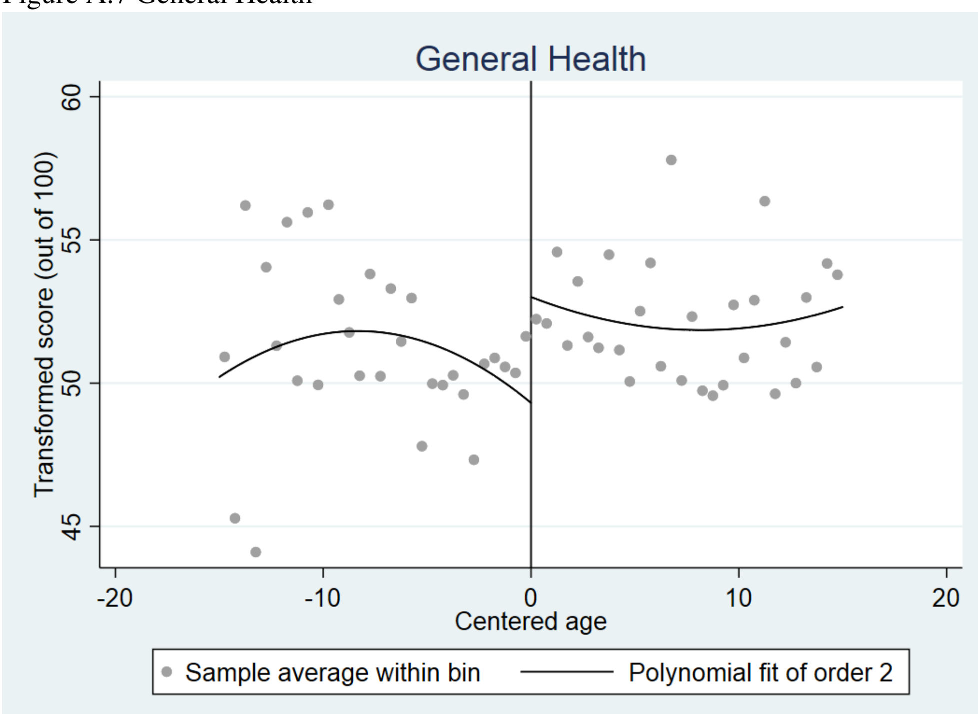


Figure A.8 Social Functioning

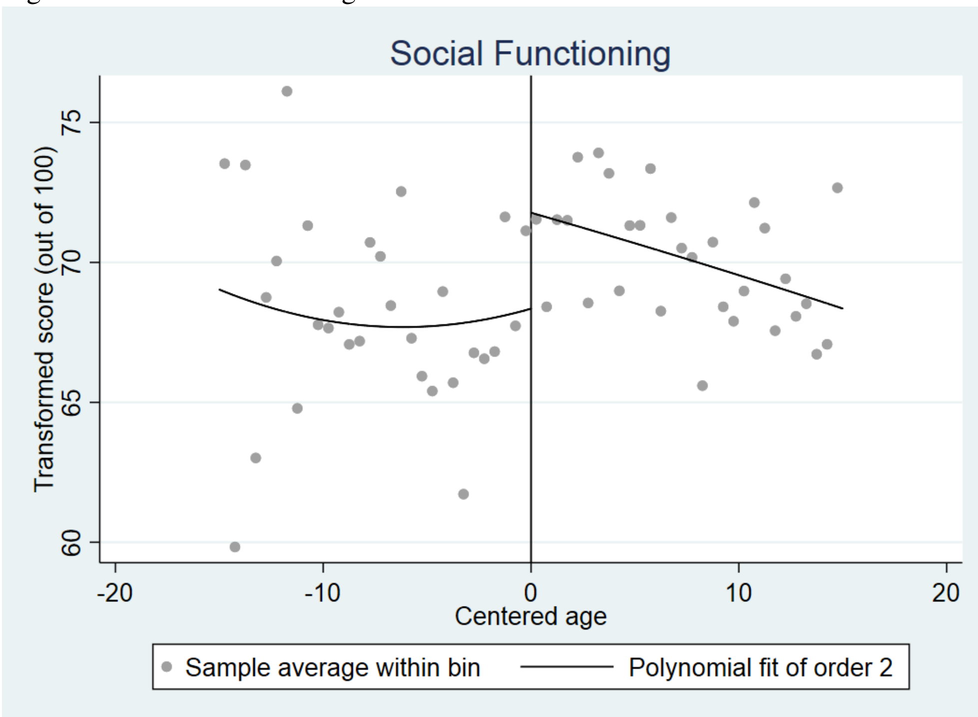


Figure A.9 Role Emotional

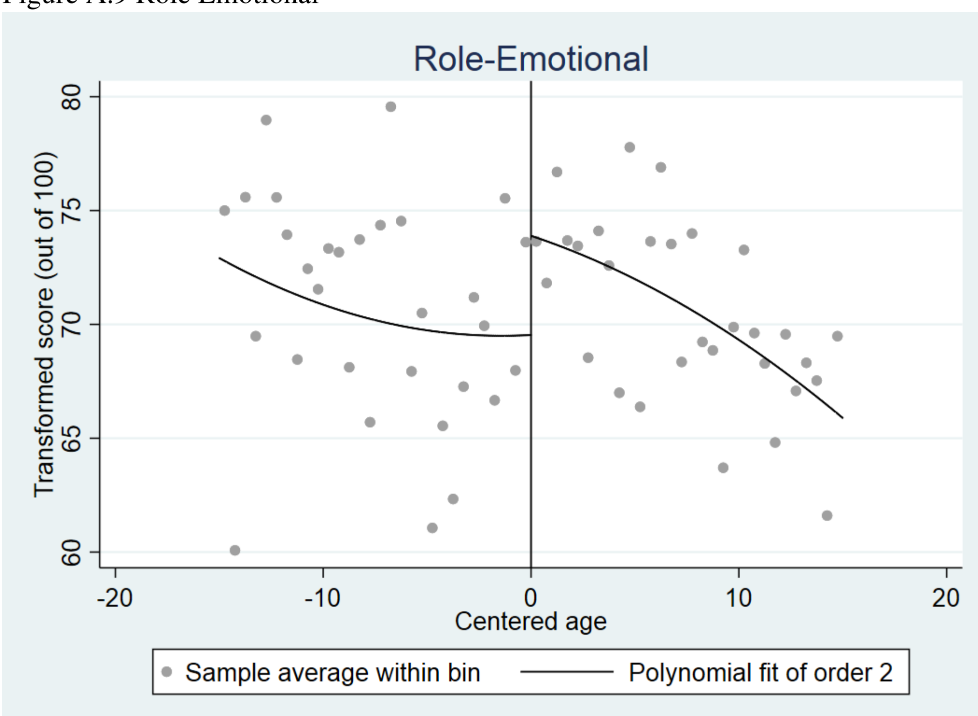


Figure A.10 Vitality

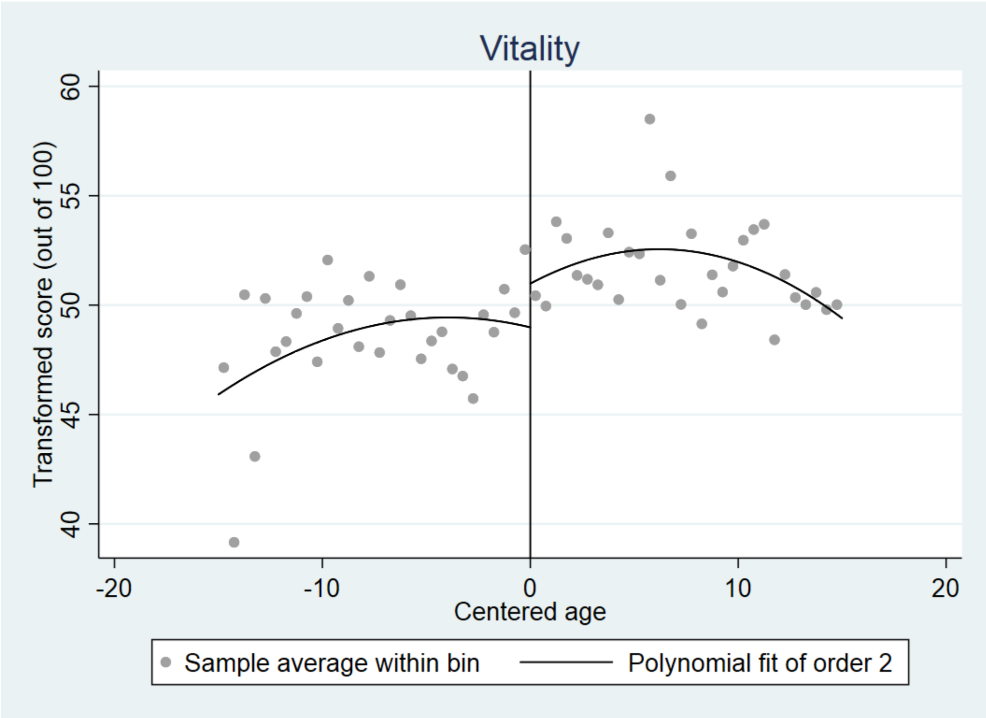
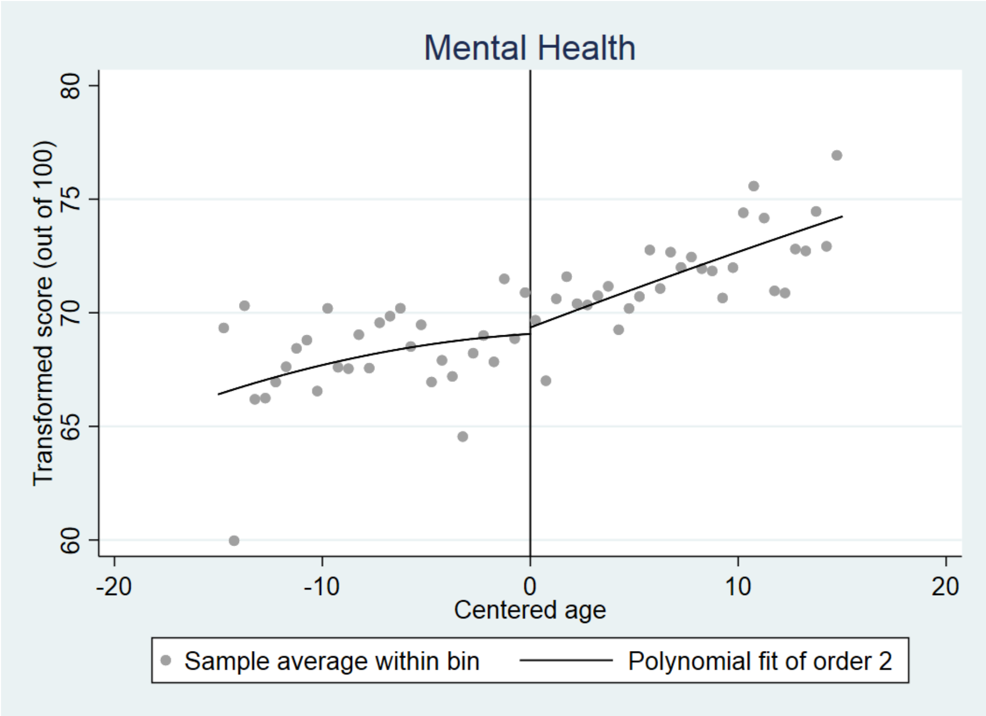


Figure A.11 Mental Health



document document

Gender falsification test

This section presents the falsification test applying female eligibility rules to the male sample (referred to in Section 4.5.1). It shows that males' retirement decisions are not influenced by the eligibility rules applicable to females.

Table 13: Falsification test applying female eligibility rules to males

	Gender falsification
$1(ctrage_{jt} \geq 0)$	0.025
	(0.020)
Observations	10,289
F-statistic	1.66

Note: Robust standard errors clustered at the individual level are in parentheses.

Alternative Income Sources

Some Australians have alternative sources of retirement income such as superannuation an employer -and employee- co-funded retirement scheme similar to the 401(k) scheme in the United States. Superannuation is not compulsory for self-employed people. People with superannuation savings will have full access to their superannuation when they reach the age of 65, regardless of whether they are retired or not. However, they could have access to superannuation before 65 if they have reached the preservation age and have retired. The preservation age is different from the pension qualifying age, but it also changes with the date of birth. For instance, the preservation age is 55 for those born before 1 July 1960, and the threshold gradually increases to 60 for those born on or after 1 July 1964. It could be the case that reaching these age thresholds also induces a discontinuity in retirement rates. However, our tests for discontinuities at the preservation age cutoff show that the F-statistics are below 10 for all samples. The lack of any meaningful discontinuity is corroborated by Figure 5 (shown only for the full sample).

Figure A.12 Discontinuity in retirement rates at preservation age (all husbands)