THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
CREATE CHANGE

Discussion Paper Series

UQ School of Economics

Spillover Effects of Retirement: Does Health Vulnerability Matter?

2021-04-28

Dominic Byrne, Do Won Kwak, Kam Ki Tang, and Myra Yazbeck

Discussion Paper No.643

Spillover Effects of Retirement: Does Health Vulnerability Matter?

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

April 28, 2021

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 partner has long-term health conditions. Given the inherent identification challenges associated with entry into retirement, we exploit an exogenous variation of pension-qualifying age in Australia. Using this exogenous variation as an instrument and data from the Household Income and Labour Dynamics in Australia survey, we find that the husband's retirement 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 AUD7,000 and AUD21,000. We also identify redistribution of domestic workload as a key transmission mechanism of the spousal spillover effects. Women with LTHCs will see their QALY and health improves only if their husband devotes more time to domestic tasks after retirement.

JEL Code: I-31; I-19

Keywords: Spillovers; Retirement; Long-term Health Conditions; QALY

[*]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 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 in these countries 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 thus 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 beyond the individual spheres of retirees. Indeed, spillovers are ubiquitous within the family; thus, the impact of a reform that increases the qualifying age for pension may affect not only the eligible person but also their family members, particularly those who are in vulnerable health.

A large body of literature has investigated the health impacts of retirement; however, conclusions remain mixed. On one hand, some evidence indicates that retirement positively affects 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 et al., 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 reform. 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 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, Müller 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 social policy: spouses in vulnerable health. In addition, they do not stratify these effects by gender. Using a

[1] Clark and Etilé (2006); Clark and Etilé (2011); Cutler and Glaeser (2010); Davillas and Pudney (2017), among others.

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 Müller 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 has overlooked the unintended effects of retirement on spouses in vulnerable health namely, spouses with long-term health conditions (LTHCs).[2] 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 Müller 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 a non-negligible negative impact on the vulnerable population. In addition, this paper explores the mechanisms through which spillovers occur. There is growing interest in uncovering the mechanisms for the health effects of retirement. Most of this literature has pointed 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.[3] The medical literature has emphasised 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).

[2]In this paper, we use Household Income and Labour Dynamics in Australia (HILDA) data, which asks respondents, "Do you have any long-term health condition, impairment or disability that restricts you in your everyday activities, and has lasted or is likely to last, for 6 months or more?" LTHCs include arthritis, asthma, heart disease, Alzheimer's disease, dementia, stroke, back problems, migraines and chronic pain, as well as sight, hearing, speech and mobility problems.

[3]The Household Income and Labour Dynamics in Australia (HILDA) data 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.

Although our data do not provide direct information on informal care, our evidence can shed light on 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 Müller 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.[4] Second, given the population segment of interest, this paper focuses on a specific measure of health outcome not previously considered in the literature: quality adjusted life year (QALY), as compared with self-reported general health status or mental health. Although self-reported health status is widely used in the literature and has its merits, its limitations are well-known, and it does not account for the quality of life, a dimension of crucial importance when focusing on a population with LTHCs. Second, health policy deliberations and cost-benefit analyses are often cast in terms of qualify of life.[5] 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

[4] A more nuanced literature on spousal spillovers by Mavromaras et al. (2013) 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 because our analyses probe how retirement impacts change depending upon characteristics of the spouse (vulnerability) rather than focusing on the average effect.

[5] 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.

countries.[6] Specifically, we exploit the 2009 policy change that generated a variation in the pension eligibility cutoff age across cohorts. Given that the cutoff 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 2018, we assess whether the change in the qualifying age is a credible source of variation in retirement status. Our empirical evidence confirms that a change in the cutoff age is strongly correlated with 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 an instrumental variables (IV) method. Our results suggest that a husband's retirement has a positive impact on his wife's health outcome when the latter has LTHCs. Specifically, we find that for women who have LTHCs, their husband's retirement, on average, improves their QALY by 0.66 standard deviations, their physical health 0.59 standard deviations, and their mental health 0.75 standard deviations. To get a sense of the importance of these effects, we draw upon the literature on QALY and cost-effectiveness thresholds and estimate the dollar value of the husband-to-wife spillover effects. Our back-of-the-envelope calculations suggest that these effects could be worth somewhere between AUD7,000 and AUD21,000. We also find that intra-household reallocation of domestic tasks is a key transmission mechanism of the identified spousal spillover effects: women with LTHCs see their health improve only if their husbands spend more time on domestic tasks after retirement.

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 presents our conclusions.

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.

[6] 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 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.

2.1 Sample Construction

We employ data from waves 2–18 (2002–2018) 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.

The following sample restrictions are imposed on the data set. 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, we can infer that this person 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, it is not guaranteed that this person is retired. Such ambiguous 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 17,363 observations corresponding to 2,710 matched couples.

Table 13 in the Appendix presents the summary statistics for the females in our overall sample and two subsamples based on whether they have LTHCs. The summary statistics focus on females because, as explained later, our analysis focuses on how women are affected by their husbands' retirement and how their LTHCs matter in that regard. Approximately 30% of females in our sample 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.

2.2 Health Outcome Variables

Our health outcome measure is QALY 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, Peter J. Neumann, 2014; Nimdet et al., 2015; and Woods et al., 2016). 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 affect 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 (see Table 12 in the Appendix). 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 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 combined income for working couples and is higher for single-income couples; thus, the prospect of receiving the age pension upon reaching the eligibility cutoff age represents a financial incentive to enter into retirement. This reasoning is supported by the first-stage estimation results presented in Section 3.1.[7]

2.3.1 Pension Age and Eligibility

In Australia, age pension 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 eligibility cutoff age alters labour market incentives for a substantial portion of the population, potentially creating an exogenous variation in retirement rates. The cutoff age for the pension varies depending on the date of birth and gender (see Table 14 in the Appendix).

Pension is not a universal welfare in Australia. Besides reaching the cutoff age, people must also pass income and asset tests. 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 are collected only infrequently in the HILDA survey. Consequently, 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. First, we include income as a control variable. Second, we undertake sensitivity analysis to ascertain whether the responsiveness of retirement rates to the pension-qualifying age differs by income level.

[7] Furthermore, the F-test results for the first-stage estimation using the subsamples with low incomes confirm that low income people respond to this financial incentive to retire (Table 9).

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

3 Empirical Framework

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

$$H_{it} = \beta_0 + \beta_1 R_{jt} + \boldsymbol{X}_{it}\gamma_1 + \boldsymbol{X}_{jt}\gamma_2 + \boldsymbol{Z}_{ht}\gamma_3 + v_i + u_{it}, \tag{1}$$

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; $\boldsymbol{X}_{it}$ includes observed characteristics variables for i in year t; $\boldsymbol{X}_{jt}$ includes other observed characteristics variables for j in year t; $\boldsymbol{Z}_{ht}$ includes household-level observed characteristics variables for household h of the couple in year t; v_i is the unobserved time-invariant heterogeneity of i; and u_{it} is an an error term. We are interested in obtaining a causal estimate of β_1 and identifying the impact of the spouse's retirement on the respondent's health outcome.

Determining 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 who could catalyse the spouse's retirement (reverse causality). Second, it is possible that unobserved confounding factors (e.g., inheritance or lottery winnings) may 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 the couple's health. Given these identification issues, using a naive linear regression to estimate equation (1) 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. For our sample period, there have been four changes in the eligibility cutoff age for male (increasing incrementally from 65 to 67) and 14 changes for female (increasing incrementally from 60 to 67). By exploiting these 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 using an IV approach. It should be pointed out that with the estimated effect is local in the sense that we only estimate the effect of those people who are induced into later retirement by the exogenous increase in the cutoff age Imbens and Angrist (1994). A few studies have used the same IV approach and adopt eligible age as a source of exogenous variation in the retirement status to analyse different outcomes (e.g., Coe and Zamarro, 2011; Rohwedder and Willis, 2010; Insler, 2014).

To justify our approach, we must examine two critical issues. First, the validity of our IV relies on the determination of pension eligibility. As eligibility is determined by age (among other factors), it is required that the spouse's eligibility is not correlated with outcome variables in the second-stage (i.e., the respondent's QALY, mental health, and physical health status) conditional on the included control variables and fixed effects (FEs). Because the cutoff age is a policy decision, whether the spouse meets the cutoff age criterion should not be correlated with the respondent's health status.

Second, we need to establish that reaching the cutoff age strongly affects the actual retirement status. We thus test for weak IV through the first-stage estimation using the following equation:

$$R_{jt} = \alpha_0 + \delta_1 \cdot 1(A_{jt} \geq age\ cutoff) + \boldsymbol{X}_{it}\alpha_1 + \boldsymbol{X}_{jt}\alpha_2 + \mathbf{Z}_{ht}\alpha_3 + u_{jt}, \tag{2}$$

where A_{jt} is age; $1(A_{jt} \geq age\ cutoff)$ is an indicator of meeting the age criterion for a pension and is our IV; and R_{jt} is equal to 1 if spouse j is retired in year t, and 0 otherwise.

Table 1 presents the results for the first-stage estimations for four different samples. In the first three columns, we test how reaching the cutoff age affects husbands' retirement probability; in the last column, we do the same for the wives.

In the first column, where all husbands are considered, reaching the cutoff age increase the husbands' retirement probability by 12.6 percentage points and is statistically significant at the 1% level. Among the first three subgroups, the effect is strongest for the one in which the wives have LTHCs; the increase in the husbands' retirement rates is 18.0 percentage points.

The findings in the last column show that this estimation of the wives' retirement probability suffers from the weak IV problem (the F-statistics = 8.821 < 10), so the estimate is not reliable enough for interpretation. One explanation of this finding is that the prospect of receiving the age pension may not generate a strong retirement incentive for a person (i.e., the wives in this case) if there is another income earner (i.e., their husband) 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] Given that the wives' retirement probability estimation does not pass the weak IV test, for the rest of the paper, we focus only on how the husbands' retirement may affect their wives.

In the second-stage, we apply an IV approach followed by an FE-IV approach as a way to test the robustness of the results. As long as the assumptions on the IV, $1(A_{jt} \geq age\ cutoff)$, are valid, both the IV

[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.

and FE-IV estimators are consistent. The difference between estimates of the two approaches should go to zero as the sample size goes to infinity.

4 Results

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 LTHCs. 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 husband's retirement on the wife's QALY, for all women, women with with LTHC, and women without LTHC, are displayed in Table 2. For each sample, we present two specifications: IV and FE-IV. The F-statistics are above 10 in all six these cases in the table, reconfirming the absence of weak instrument issues. Furthermore, for all three subsamples, the differences between the IV and FE-IV estimates are not significant. This is evidence to support the validity of the IV because under the validity assumption, the outcome variable should be uncorrelated with unobserved factors and thus, additionally accounting for unobserved time-invariant factors using FEs should not affect the estimates.

Based on the IV results, the estimated spillover effects for the full sample suggest that a husband's entry into retirement leads to an improvement in the wife's QALY of approximately one-quarter of a standard deviation, but the observed impact is not statistically significant. On the other hand, there is a positive and statistically significant finding for the subsample of wives with LTHCs, for whom the spillover effects are approximately two-thirds of a standard deviation or 12% of the average QALY. The estimate for the subsample without LTHCs is also positive, but smaller in magnitude and statistically insignificant. This last finding does not appear to be a consequence of a reduced sample size, as this subsample is larger than the subsample with LTHCs.

The estimates from the alternative FE-IV specification are not statistically different from the IV one, although the level significance changes slightly for the full sample and the LTHC subsample. Given that additional control for FEs does not change our main coefficient estimates, for all other estimations presented in this paper we focus on the IV estimation results.

In summary, we obtain robust results that the retirement of a husband has a positive *causal* impact on his wife's QALY only when she has LTHCs, and the impact is very significant in both the statistical and economical sense.

4.2 Estimating the cost effectiveness of the intervention

The findings in section 4.1 allow us to consider the cost of a hypothetical 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 cutoff 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 AUD18,507 per year as in 2021.[10]

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 (World Health Organisation, 2001), and the "one to three times of the GDP per capita" threshold was subsequently set out in another WHO-affiliated publication (Hutubessy et al., 2003). In 2019, Australia's GDP per capita was equal to AUD76,985 (World Bank data), so one QALY is worth between AUD76,985 and AUD230,955. From Table 2, we know that the husband's retirement is expected to improve his wife's QALY by 0.09 (the average of IV and FE-IV estimates). Therefore, the estimated benefit from spousal health spillover resulting from the intervention is worth between AUD6,929 and AUD20,786, and the cost of intervention (AUD18,507) falls within, though near the higher end of, this range. We reiterate that AUD18,507 is the maximum pension payout, which does not apply to all retirees.

The literature on cost-effectiveness thresholds has stressed that one threshold size cannot fit all circumstances. Policymakers must 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 the 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.

[10]This is the pension per person. The data is drawn from the Australia Government's Services Australia website: https://www.servicesaustralia.gov.au/individuals/services/centrelink/age-pension/how-much-you-can-get

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. Although 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 that the impact of retirement on QALY is concentrated in the subsample of women experiencing LTHCs, we focus on this subsample in this section.[11]

The results in the upper panel of Table 3 show that the husband's retirement has a positive and statistically significant effect on the aggregate physical health of the wife. This appears to be driven primarily by an improvement in the Physical Functioning and General Health scores, both of which are significant at the 5% level. The husband's entry into retirement also improves the wife's Bodily Pain score (i.e., a reduction in Bodily Pain) at the 10% level of significance. Furthermore, the magnitudes of these coefficients are economically significant and are in the range of 0.5 to 0.8 standard deviations. 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 results in the lower panel of Table 3 show that the husband's retirement also has a broad impact on the wife's mental health. Social Functioning, Role-Emotional Health, and Vitality are positively affected at the 5% level of significance. In each instance, the improvement in health scores is around 0.7 standard deviations. The directions and magnitudes of these coefficients are comparable to the findings of Mavromaras et al. (2013), who use similar Australian data to ascertain the effects of own retirement, rather than spousal retirement.

Overall, the results in Table 3 indicate that no single component drives overall spousal health improvement. Rather, it appears that improvements across various health components are responsible for the positive spillover effects detected in both aggregate physical and mental health.

[11] For the subsample of wives without LTHCs, which has a valid first-stage estimation with no weak IV problem, no health outcomes significantly increase upon their husband's retirement.

4.4 Mechanisms

To understand why a husband's entry into retirement has discernible impacts on women with LTHCs, we investigate if redistribution of domestic workload is a transmission mechanism. Redistribution of domestic workload between couples is a candidate because women typically shoulder many more domestic responsibilities than men,[12] and taking a bigger share of domestic workload is an obvious way for men to reduce the burden on their wives.

First, we aim to identify which domestic activities are affected by husbands' retirement and investigate whether these changes differ in the presence of LTHCs. To measure domestic workload, we deploy the weekly time use data from HILDA. We focus on the time use for three types of domestic work: housework, household errands, and outdoor tasks. We define the combined time use for domestic work 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 4. Upon retiring, the husband's time allocated to domestic work increases by approximately 12 hours per week (significant at the 1% level) if the wife has LTHCs. If the wife does not have LTHCs, entry into retirement leads to an increase in the husband's domestic work of slightly less than 10 hours per week (significant at the 5% level). For the LTHC subsample, the time use for all three individual activities significantly increases, whereas only that for errands and outdoor tasks significantly increases in the no-LTHC sample.

Overall, Table 4 reveals that husbands engage more in domestic work after retirement, but there is yet evidence that it is the transmission mechanism of the identified spousal health spillovers. To verify whether husbands' greater involvement in domestic work is transmitting the health spillover effect, we partition the retiree sample into two subsamples according to how the husband's time use changes after retirement. The first subsample consists of the husbands who spend more time on domestic work upon retiring, whereas the second subsample contains the husbands who either decrease or maintain their time spent on domestic work after retirement.[13] We perform this analysis only for the relevant sample: wives with LTHCs.

Table 5 presents results for QALY, Aggregate Physical Health, and Aggregate Mental Health using these two subsamples. The first three columns relate to those husbands who spend more time on domestic work upon retiring. For their wives' QALY, there is a statistically significant improvement that is similar in magnitude to the baseline result in Table 2 (columns 3 and 4). Likewise, their wives' Aggregate Physical Health

[12] For our sample, on average, women spend 28.8 hours per week on domestic tasks, almost twice the 15.9 hours spent by men.

[13] For individuals who retired before the timeframe of the sample, it is not possible for us to observe how their time use changed in the year in which they retired. Therefore, such individuals were dropped from the sample when we conducted the mechanism analysis.

and Mental Health benefit from their retirement, and the benefits are similar in magnitude to the baseline estimate in Table 3. However, none of the effects are as precisely measured as in the baseline estimations, probably because of smaller sample sizes. The results in the last three columns are from the subsample where the retiring husbands spend less, or the same amount of time, on domestic work upon retirement. Although the instrument strength F-statistic remains above the conventional rule of thumb, the health effects are negative, albeit not statistically significant. Taken together, the findings of the mechanism analysis suggest that the positive health spillovers are confined to those couples in which the husbands devote more time to domestic work after retirement.

An increase in time spent by husbands on domestic work does not imply that they are taking up a bigger share of workload if both husbands and wives increase their time inputs. Such a scenario is possible because couples may change their domestic routines after retirement; for instance, they may prepare meals at home more often or make more home improvements. In fact, we find that wives also spend more time on domestic tasks upon their husbands' retirement, but wives without LTHCs increase their time allocated to domestic work more than those with LTHCs. Therefore, to get a deeper understanding of domestic workload redistribution as a transmission mechanism, we investigate if there is any within-household substitution effect as a result of the wife reallocating domestic tasks to her husband upon his retirement.

Table 6 presents the results. We consider one relative measure and one absolute measure of the substitution effect: the relative measure is the husband's share of the couple's total time use for domestic work, and the absolute measure is the husband's time use for domestic work minus the wife's as expressed in hours per week. The observed substitution effect differs markedly between the two subsamples based on LTHCs. In the case of the wife having a LTHC, the results indicate that the husband's entry into retirement leads to the husband's share of domestic workload increasing by about 13 percentage points, or almost 6 hours per week. On the contrary, if the wife does not have a LTHC, there is no evidence of within-household redistribution of domestic workload.

Overall, the evidence supports our hypothesis that the redistribution of domestic workload between couples is a key transmission mechanism of the positive spousal spillovers identified earlier.

4.5 Robustness Tests

In Section 3, we compare the IV and FE-IV estimations as a way to validate our identification strategy. This test and the weak IV test show no violation of the assumptions that justify the IV method. In this subsection, we provide further evidence from four more tests to validate our methodology. We first examine the influence of divorce and mortality on the spousal spillover effect through attrition. After that, we consider two falsification tests and a placebo test.

4.5.1 Attrition

A spousal interdependence dissolves if the couple divorces or one of them dies. A recent study by Doorley and Stancanelli (2019) indicates that retirement may increase the chance of divorce. As our sample excludes the couples who divorce during the same year the husbands retire, it may be overpopulated with couples who have stronger relationships and who are potentially more altruistic towards each other. If true, this could cause sample selection bias. Similarly, Fitzpatrick and Moore (2018) find that retirement increases mortality. In our case, if a husband passes away in the year of his retirement, he (and his wife) will be dropped from our sample, potentially causing another sample selection bias. Because the population we are interested in are elderly people, a separation caused by death is not a trivial issue.

Any separation caused by either of these two reasons will induce attrition problems for our panel data analyses. To examine how attrition may affect our earlier findings, we apply the inverse-probability-weighting (IPW) method Wooldridge (2007). To implement the IPW method, we first model the probability of a couple remaining in a marital relationship (i.e., effectively remaining in the sample) using a probit model.

For each time period t, we estimate the following model:

$$1(S_i = 1) = 1(\boldsymbol{X}_i\gamma_{1t} + \boldsymbol{X}_j\gamma_{2t} + \boldsymbol{Z}_h\gamma_{3t} \geq 0), \tag{3}$$

where $S_i = 1$ if the couple does not attrit at $t+1$ and 0 otherwise; and $\boldsymbol{X}_i$, $\boldsymbol{X}_j$, and $\boldsymbol{Z}_h$ are the characteristics of the respondent, the spouse, and the household, respectively, at time t. By estimating equation (3), we can obtain the predicted probability for each individual at each time period. The inverse of the predicted probability is then to be used as a weight in the IV estimations.

Table 7 reports the results. No estimation significantly differs statistically from their baseline counterparts in Table (2). This finding implies that women's QALY is not differently correlated with their husbands' retirement status across attriting and non-attriting couples.

4.5.2 Falsification tests

The first falsification test applies female cutoff ages 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 to the female cutoff ages. To conduct the test, we regenerate the instrument for males using the cutoff age for a female with the same date of birth. As Table 8 shows, the first-stage estimate becomes strongly insignificant (the F-statistic is equal to 0.86), indicating that males' retirement decisions are not influenced by the rule on females.

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 retirement incentive induced by age pension to be smaller (or perhaps indiscernible) for individuals with higher per capita incomes. Using per capita income, we perform two subsample 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 9 presents the results from the first-stage estimations. The results show that reaching the qualifying age generates a change in the retirement probability only for those on lower incomes.

Considering that the incentive to retire upon reaching the cutoff age appears to be weak for those on high incomes, it is prudent to re-estimate the main models of this paper without these observations. Table 10 presents the second-stage results for the effect of the husband's retirement on the wife's QALY, 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 (0.083 for QALY, significant at the 5% level in Table 2; 57.079 for Aggregate Physical Health, significant at the 10% level in Table 3; and 67.402 for Aggregate Mental Health, significant at the 5% level in Table 3).[14]

[14] As another 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 the income cutoff values. For instance, in the case of using the median income 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 with those from the 75th percentile cutoff, albeit statistically insignificant.

4.5.3 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 (44%). We implement the baseline specification with these false-retirement observations instead of the true-retirement observations to verify whether there are any statistically significant effects. Results reported in Table 11 show that the instrument strength is very weak under the placebo assignment. This supports the hypothesis that the actual eligibility age for the 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 an instrumental variable 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 physical and mental 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 husbands' time use on domestic tasks increases significantly upon retirement, especially if their wive have LTHCs. More importantly, we find that the positive health benefits for these women are present only when their husbands take up more of the domestic workload. On this basis, there is evidence to suggest that domestic task sharing is a key transmission mechanism of the positive health spillover effects that we observe.

The implication of these findings is that informal care and support, 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 must 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. and Barrett, G. F. (2014). The causal effect of retirement on health: New evidence from australian pension reform. *Economics Letters*, 125:392–295.

Atalay, K. and 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):71–87.

Atalay, K. and Zhu, R. (2018). The effect of a wife's retirement on her husband's mental health. *Applied Economics*, 50(43):4606–4616.

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

Bertoni, M. and 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., and Hill, S. R. (2016). Cost-effectiveness thresholds: pros and cons. *Bulletin of the World Health Organisation*, 94(12):925–930.

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

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

Clark, A. E. and Etilé, F. (2011). Happy house: Spousal weight and individual wellâbeing. *Journal of Health Economics*, 30(5):77–86.

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

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

Davillas, A. and 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.

Doorley, K. and Stancanelli, E. (2019). Marital status and retirement: An empirical study for france. techreport 12299, IZA Discussion Papers, Institute of Labor Economics (IZA).

Eibich, P. (2015). Understanding the effect of retirement on health: Mechanisms and heterogeneity. *Journal of Health Economics*, 56:87–102.

Eisenberger, N. I. (2012). The pain of social disconnection: Examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience*, 13(6):421–34.

Fitzpatrick, M. D. and Moore, T. J. (2018). The mortality effects of retirement: Evidence from social security eligibility at age 62. *Journal of Public Economics*, 157:121–137.

Godard, M. (2016). Gaining weight through retirement? results from the share survey. *Journal of Health Economics*, 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., and Edejer, T. (2003). Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. *Cost Effectiveness and Resource Allocation*, 1(1):8.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

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

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

Kämpfen, F. and Maurer, J. (2016). Time to burn (calories)? the impact of retirement on physical activity among mature americans. *Journal of Health Economic*, 45:91–102.

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

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., and Zhu, R. (2013). Age pension age eligibility, retirement and health outcomes in australia. techreport 201, NILS Working Paper Series, National Institute of Labour Studies, Flinders University, Adelaide, Australia.

Müller, T. and 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.

Nimdet, K., Chaiyakunapruk, N., Vichansavakul, K., and 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):1–16.

Peter J. Neumann, Joshua T. Cohen, M. C. W. (2014). Updating cost-effectiveness - the curious resilience of the $50,000-per-qaly threshold. *The New England Journal of Medicine*, 371(9):796–797.

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

Rohwedder, S. and 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. techreport 41/13, Melbourne Institute Working Paper, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne.

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

Wooldridge, J. M. (2007). Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics*, 141(2):1281–1301.

World Health Organisation (2001). Macroeconomics and health: Investing in health for economic development. report of the commission on macroeconomics and health. Technical report.

Younger, J., Aron, A., Parke, S., Chatterjee, N., and Mackey, S. (2010). Viewing pictures of a romantic partner reduces experimental pain: Involvement of neural reward systems. *PLOS ONE*, 5(10):1–7.

Table 1: First-stage estimations

	All husbands	Husbands (wife without LTHC)	Husbands (wife with LTHC)	All wives
Instrument: $1(A_{jt} \geq age\ cutoff)$	0.126***	0.094***	0.18***	0.051***
	(0.016)	(0.019)	(0.027)	(0.017)
F-statistic	64.964	25.806	46.240	8.821
Number of observations	17,363	12,139	5,224	13,577
Number of couples	2,710	2,381	1,470	2,201

Note: Robust standard errors clustered at the individual level are in parentheses. ***, **, * indicate significance at the 1%, 5%, and 10% level, respectively. The combined number of couples in the second and third columns exceeds the number of couples in the first column because some couples may appear in both subsamples in different years if the onset of the LTHC occurs during the period of analysis. An example would be a person who did not have any LTHC in the first few waves but developed one in a subsequent wave.

Table 2: The impact of the husband's retirement on the wife's QALY

	All wives		Wives with LTHCs		Wives without LTHCs	
	IV	FE-IV	IV	FE-IV	IV	FE-IV
Husband retirement	0.029	0.037*	0.083**	0.097***	0.006	0.016
	(0.028)	(0.020)	(0.038)	(0.037)	(0.038)	(0.027)
Change in standard deviations	*0.244*	*0.306*	*0.663*	*0.777*	*0.062*	*0.155*
Instrument strength (F-statistic)	64.964	87.984	46.104	37.230	25.806	42.641
Number of observations	17,363	16,957	5,224	4,713	12,139	11,634

Note: Robust standard errors clustered at the individual level are in parentheses.. 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. The combined number of observations in the subsamples for the FE-IV specifications is lower than the number of observations in the FE-IV specification for all wives due to the loss of singletons that emerge when the sample is split.

Table 3: The effect of the husband's retirement on the health of the wive with long-term health conditions

PHYSICAL HEALTH	Physical Functioning (PF)	Role-Physical (RP)	Bodily Pain (BP)	General Health (GH)	**Aggregate Physical Health**
Husband retirement	19.602**	12.358	13.443*	17.691**	57.079*
	(8.338)	(12.763)	(7.674)	(7.359)	(30.571)
Change in standard deviations	*0.763*	*0.285*	*0.536*	*0.771*	*0.585*
Instrument strength (F-statistic)	43.824	44.090	44.090	42.250	42.772
Number of observations	5,485	5,437	5,505	5,424	5,322

MENTAL HEALTH	Social Functioning (SF)	Role-Emotional (RE)	Vitality (VT)	Mental Health (MH)	**Aggregate Mental Health**
Retired	19.703**	30.484**	14.960**	4.854	67.402**
	(8.730)	(13.054)	(6.487)	(5.641)	(28.211)
Change in standard deviations	*0.708*	*0.738*	*0.720*	*0.260*	*0.748*
Instrument strength (F-statistic)	43.560	45.024	43.957	43.957	45.968
Number of observations	5,548	5,431	5,515	5,512	5,400

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, whereas 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 4: The impact of retirement on the husband's time use on domestic activities

Wives with LTHCs	Housework	Errands	Outdoor tasks	Combined
Husband retirement	3.725*	3.032**	5.773**	12.268***
	(2.008)	(1.471)	(2.463)	(4.202)
Change in standard deviations	*0.477*	*0.617*	*0.605*	*0.784*
Instrument strength (F-statistic)	47.472	43.165	46.376	45.698
Number of observations	5,485	5,434	5,509	5.363

Wives without LTHCs	Housework	Errands	Outdoor tasks	Combined
Husband retirement	0.889	2.965**	5.456*	9.743**
	(2.058)	(1.487)	(3.143)	(4.595)
Change in standard deviations	0.138	0.689	0.0597	0.681
Instrument strength (F-statistic)	35.760	33.989	33.178	36.603
Number of observations	12,724	12,661	12,810	12,470

Note: Robust standard errors clustered at the individual level are in parentheses. Combined time use is calculated as the sum 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 5: Retirement impacts disaggregated by the husband's time use on domestic activities

	Husbands with increased time use			Husbands with decreased/unchanged time use		
	QALY	Aggregate Physical Health	Aggregate Mental Health	QALY	Aggregate Physical Health	Aggregate Mental Health
Husband retirement	0.117*	45.637	83.719*	-0.056	-14.372	-17.336
	(0.063)	(46.909)	(44.300)	(0.072)	(56.465)	(53.016)
Instrument Strength (F-statistic)	18.749	17.140	18.490	12.461	11.156	12.532
Number of observations	1,611	1,649	1,665	877	889	901

Note: Robust standard errors clustered at the individual level are in parentheses. **, * indicate significance at the 5%, and 10% level respectively. For individuals who retire prior to the timeframe of the sample, it is not possible for us to observe how their time use changed in the year in which they retired. Therefore, such individuals are dropped from the sample in the estimations .

Table 6: Substitutions effects in time use for domestic work

	Husband's share of total time use for domestic work		Husband's time use minus wife's time use for domestic work	
	Wives with LTHC	Wives without LTHC	Wives with LTHC	Wives without LTHC
Husband retirement	0.132**	-0.004	5.875*	-0.202
	(0.062)	(0.072)	(3.181)	(3.918)
Instrument strength (F-statistic)	50.837	30.360	51.409	30.140
Number of observations	4,833	11,422	4,838	11,433

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, whereas a negative coefficient indicates that the husband's share of the household time use is decreasing. ***, **, * indicate significance at the 1%, 5%, and 10% level, respectively. The number of observations is slightly smaller for the relative estimates because total time use was reported to be zero for a very small number of couples, preventing the calculation of the ratio .

Table 7: The impact of the husband's retirement on the wife's QALY with attrition adjustment

	All wives	Wives with LTHCs	Wives without LTHCs
Husband retirement	0.033	0.092**	0.006
	(0.029)	(0.040)	(0.039)
Change in standard deviations	0.273	0.737	0.056
Instrument strength (F-statistic)	63.203	43.165	25.301
Number of observations	17,304	5,204	12,100

Note: Robust standard errors clustered at the individual level are in parentheses. 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. The number of observations available for the attrition-adjusted analysis is slightly smaller than for the baseline analysis because the attrition-adjusted analysis draws on a larger array of demographic covariates (e.g. country of birth and whether a person speaks a language other than English), which are missing for a small number of observations in the baseline sample.

Table 8: Falsification test applying female eligibility rules to males

	Gender falsification
$1(A_{jt} \geq age\ cutoff)$	0.018
	(0.020)
F-statistic	0.86
Number of observations	8,326

Note: Robust standard errors clustered at the individual level are in parentheses. The male and female cutoff ages are the same for individuals born on or after 1 January 1949. Therefore, this test includes only those individuals who were born before this date .

Table 9: First-stage regressions, comparison by income percentile

Change in retirement rates, by percentile	0~75th percentiles	75~100th percentiles
$1(A_{jt} \geq agec\, cutoff)$	0.147***	0.033
	(0.017)	(0.028)
F-statistic	71.234	1.323
Number of observations	12,898	5,072

Note: The subsample at or below 75th percentile contains households with per capita income less than or equal to AUD 68,741, and the subsample at or above 75th percentile contains households with per capita income greater than or equal to AUD 68,741. *** indicates significance at the 1% level .

Table 10: Retirement spillover effects from husbands to wives with long-term health conditions (0~75th income percentile)

	QALY	Aggregate Physical Health	Aggregate Mental Health
Husband retirement	0.074*	57.371*	67.449**
	(0.039)	(32.217)	(29.612)
Instrument strength (F-statistic)	43.296	39.564	42.772
Number of observations	4,323	4,408	4,479

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

Table 11: Falsification test assigning placebo retirement status

	QALY	Aggregate Physical Health	Aggregate Mental Health
Husband retirement	-4.404	-1,239.804	-1,916.991
	(32.796)	(4,137.258)	(7,364.628)
Instrument strength (F-statistic)	0.017	0.090	0.068
Number of observations	5,497	5,605	5,685

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

Online Appendix

Table 12: 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 13: 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	0.13	59.23	0.00
Physical Functioning (PF)	77.03	22.41	84.03	16.58	60.97	25.57	61.48	0.00
Role Physical (RP)	72.88	39.35	84.07	31.43	47.18	43.46	56.78	0.00
Bodily Pain (BP)	66.71	24.41	73.73	20.51	50.61	25.02	60.26	0.00
General Health (GH)	66.38	21.61	72.88	17.36	51.40	22.95	61.71	0.00
Aggregate Physical Health	283.24	91.70	314.80	68.09	210.49	97.55	71.50	0.00
Social Functioning (SF)	82.56	23.65	88.36	18.93	69.23	27.68	46.72	0.00
Role Emotional (RE)	84.12	32.47	90.09	25.64	70.39	41.19	32.79	0.00
Vitality (VT)	60.73	20.00	65.35	17.80	50.13	20.74	47.36	0.00
Mental Health (MH)	76.00	16.65	78.66	14.96	69.89	18.61	30.88	0.00
Aggregate Mental Health	303.44	77.77	322.48	62.96	259.68	89.96	46.97	0.00
Retired	0.48	0.50	0.41	0.49	0.64	0.48	-27.47	0.00
Age	60.94	7.28	60.29	7.27	62.42	7.07	-18.43	0.00
Household Income (AUD)	84,596	76,828	90,539	75,947	70,597	77,106	15.79	0.00
Persons in Household	2.44	0.91	2.48	0.95	2.35	0.81	9.60	0.00
Husband's time use on domestic work								
Housework	6.37	6.87	6.06	6.44	7.11	7.72	-8.69	0.00
Errands	4.43	4.54	4.22	4.34	4.90	4.95	-8.56	0.00
Outdoor Tasks	8.48	9.36	8.32	9.24	8.83	9.63	-3.26	0.05
Combined	19.24	14.85	18.58	14.43	20.75	15.69	-8.47	0.00
Number of Observations	17,363		12,139		5,224			

Table 14: 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

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 have full access to their superannuation when they reach the age of 65, regardless of whether or not they are retired. However, they could have access to superannuation before 65 if they have reached the preservation age and have retired. The preservation age differs 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 retirement. Therefore, we test whether reaching the preservation age cutoff has an impact on retirement probability. However, we find that the F-statistics are below 10 for all husband subsamples, meaning that reaching the preservation age does not provide husbands with a strong enough incentive to retire.