Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

The unequal burden of retirement reform

Evidence from Australia

Todd Morris

DP 12/2019-046

The Unequal Burden of Retirement Reform: Evidence from Australia*

Todd Morris

Munich Center for the Economics of Aging, Max Planck Institute for Social Law and Social Policy

December 16, 2019

Abstract

As many governments attempt to contain rising expenditure on retirement pensions by increasing eligibility ages, there are concerns that such reforms disproportionately affect poorer households. In this paper, I examine this trade-off in the context of a 1994 Australian reform that increased women's pension-eligibility age from 60 to 65. Using detailed longitudinal data, I estimate that each one-year increase in the pension age reduced net government expenditure by 800 million dollars per annum. However, the responses to the reform — increased employment and higher receipt of other transfers — were concentrated among poorer households, as were the negative effects on household incomes. These unequal impacts meant that, among affected cohorts, the reform increased relative poverty rates by 33 to 39 percent and inequality measures by 12 to 15 percent. The results also highlight the importance of other government transfers in attenuating the fiscal and regressive impacts. Back-of-the-envelope calculations indicate that, without other transfers, the fiscal savings and regressive impacts would have been two- and four-times larger respectively.

Keywords: Retirement age; distributional effects; poverty; inequality

JEL Classification: H55, I38, J26

*Acknowledgements: I wish to especially thank my advisors — John Haisken-Denew, Timothy Moore and Jenny Williams — for their excellent comments and supervision, and I am grateful for comments from Mary Kate Batistich, Jillian Carr, Bill Evans, Greg Gilpin, Fedor Iskhakov, Leslie Martin, Cain Polidano, Stefan Staubli, Leran Wang and Roger Wilkins. Useful feedback was also provided by participants at a number of seminars, workshops and conferences. This paper uses unconfidentialised unit-record data from the Household Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services and is managed by the Melbourne Institute of Applied Economic and Social Research. I gratefully acknowledge the support of the Faculty of Business and Economics and the Department of Economics at the University of Melbourne. **Correspondence:** Munich Center for the Economics of Aging, Amalienstrasse 33, Munich 80799, Germany. Email: toddstuartmorris@gmail.com

1 Introduction

Aging populations are causing serious concerns about the fiscal sustainability of public-pension systems. In response, many governments are introducing policies that aim to extend working lives and reduce pension outlays. The most common policy has been to increase the eligibility age for retirement benefits. Such policies have occurred in many countries, including Australia, the United Kingdom, Germany, France, Italy and many other European countries. There have also been calls in the United States to raise the eligibility ages for Social Security and Medicare.[1] Existing studies find that increases in pension-eligibility ages are effective in reducing net fiscal expenditures (Staubli and Zweimüller, 2013; Oguzoglu, Polidano and Vu, 2016) and increasing the employment rates of older individuals (Staubli and Zweimüller, 2013; Vestad, 2013; Atalay and Barrett, 2015; Cribb, Emmerson and Tetlow, 2016; Manoli and Weber, 2016; Geyer and Welteke, forthcoming).

However, these reforms are not without controversy, partly due to concerns about equity. One concern is that increases in pension-eligibility ages are unfair on poorer households due to a large and increasing gap between the life expectancy of individuals with high and low socioeconomic status (Cristia, 2009; Waldron, 2007; Bound et al., 2014). Another concern is that individuals with physically demanding jobs, who are also more likely to have low levels of wealth, may find it harder to delay retirement than those with more sedentary jobs. A further concern is that these reforms may be regressive; since low-income households generally receive a larger share of their income from public pensions, delaying eligibility for such pensions may increase poverty and inequality among older households. Despite these concerns, there is remarkably little empirical evidence on the distributional effects of retirement reforms.

This paper helps address this gap; I show that increases in the pension-eligibility age can be significantly regressive. I study the effects of a 1994 Australian reform that gradually increased women's eligibility age for the public retirement pension from 60 to 65. I draw on three strengths of this environment. First, the impact of the reform was large. As Australia's retirement pension is received by nearly three-quarters of women above the eligibility age, the five-year increase in the eligibility age significantly delayed many women from being eligible for this pension. Second, a detailed longitudinal survey from 2001 to 2015 allows me to study the effects of the reform using information on incomes, labor supply, wealth and other characteristics at both the individual and household level. Third, the reform was phased in based on women's date of birth over the period from 1995 to 2013, providing plausibly exogenous variation in women's eligibility for the retirement pension during the sample period.

Using a differences-in-differences approach that exploits this variation, I estimate the causal effects of

[1] These proposals are distinct from the legislated increases in the Normal Retirement Age in the U.S., which correspond to a reduction in benefits rather than a delay in eligibility.

the reform on households' labor supply and income. I start my analysis by estimating the average effects of the reform on women's labor supply and income from government transfers. Consistent with previous studies of the reform (Atalay and Barrett, 2015; Oguzoglu, Polidano and Vu, 2016), I find that the largest response was women remaining on other government transfers for longer; on average, women offset 63% of their income losses from the retirement pension through other government transfers. In contrast, the estimated effect on female labor supply is relatively modest; on average, women offset 21% of their income losses from the retirement pension through an increase in earnings.

As my focus is on the distributional effects of the reform, I consider how the effects on households' labor supply and income are expected to vary across the income distribution using a basic model of labor supply. Since the retirement pension is means-tested based on current household income, whereby payments decrease with household income and are zero for high-income households, we expect the reform to cause (i) the largest reduction in pension payments for women in low-income households and (ii) no reduction in payments for women in high-income households. In addition, due to other payment rules, we expect larger reductions in payments among single women and women in poorer households. Therefore, we expect heterogeneity in the labor supply response, with larger responses from single women and households with less income and wealth. The model also indicates that the reform may have a regressive impact overall; assuming that households only partially offset their income losses through an increase in earnings, we expect a reduction in total income for low-income households and no effect for high-income households.

My empirical results support these predictions. The income losses from the retirement pension are largest among single women, renters and women in poorer households. In addition, these groups explain most of the increase in labor supply and higher receipt of other transfers, and the labor supply response is confined to part-time and relatively low-earning women. Interestingly, I find only small and statistically insignificant effects on the labor supply of partnered women and their spouses.

To examine the distributional effects on household incomes, I estimate the effects on household income at many points in the income distribution using distribution regression methods, an increasingly common alternative to quantile regression (e.g., Foresi and Peracchi, 1995; Rothe, 2012; Chernozhukov, Fernández-Val and Melly, 2013; Dube, 2019). Here, distribution regression involves estimating a sequence of binary regressions, where in each regression the dependent variable is an indicator for household income being below a given cut-off, and the effects are estimated at a grid of cut-offs across the distribution. My measure of household income is household disposable income (income net of taxes and transfers), with an adjustment for housing costs and the value of in-kind benefits.

Using this approach, I find that the reform had a regressive impact overall. In line with expectations, I find strong negative effects on the incomes of low-to-middle-income households but little impact on households in the top half of the income distribution. To quantify the overall regressivity of the reform, I use the distribution regression estimates to estimate the counterfactual distribution (i.e., what the income distribution would have looked like if not for the reform) and then compare inequality measures for the actual and counterfactual distributions. Using this approach, I estimate that the reform increased inequality measures by 12–15% among affected households (households containing women in the affected cohorts). This increase is substantial; the same increase could be achieved by transferring around $5,000 per annum ($3,500 US) from households in the bottom quintile of the income distribution to households in the top quintile.

Focusing on the effects near the bottom of the income distribution, I find a meaningful increase in relative poverty rates. Using the standard approach of setting the poverty line at 50% or 60% of the median household income, I find that the reform increased relative poverty rates among affected households by 2.1–4.4 percentage points (33–39%), with single women and renters explaining most of this increase.

As a final step, I estimate the net fiscal savings from the reform, considering (i) the change in government expenditure on the retirement pension, other transfers and in-kind benefits and (ii) the change in government revenue from income tax. I estimate that each one-year increase in women's eligibility age saved approximately 800 million dollars per annum, a saving which is equivalent to 1.8% of the Australian Government's current expenditure on the retirement pension.

Overall, these findings highlight a significant trade-off for policymakers: while widespread increases in pension-eligibility ages reduce government expenditure, these savings can result in higher poverty and inequality among older households. The results also highlight the importance of other government transfers in attenuating these impacts. Back-of-the-envelope calculations indicate that the fiscal savings would have been twice as large (1.76 billion dollars) if women were unable to substitute to other transfers. However, as low-income households disproportionately relied on such transfers, the increase in inequality measures would have been around four-times larger (42–88%), and the increase in relative poverty rates around six-times larger (20.4–24.1 percentage points). While these calculations ignore behavioral responses, they demonstrate the importance of other government transfers as a buffer for low-income households.

Related literature. This paper contributes to a large literature examining behavioral responses to changes in retirement policy. Within this literature, the main focus has been on the effects on employment. With limited variation in statutory retirement ages, earlier studies mainly examined changes in the level of retirement benefits (Burtless, 1986; Krueger and Pischke, 1992; Coile and Gruber, 2007; Liebman, Luttmer and Seif, 2009) or used cross-sectional variation in retirement incentives to predict labor supply responses to changes in retirement ages (Rust and Phelan, 1997; Panis et al., 2002; Gruber and Wise, 2004). However, with more and more variation in retirement ages, the number of ex-post evaluations has grown. Many of these studies initially focused on increases in the Normal Retirement Age in the U.S.,[2] a policy that is distinct

[2]For example, see Duggan, Singleton and Song (2007); Song and Manchester (2007); Mastrobuoni (2009); Blau and Goodstein (2010); Behaghel and Blau (2012). Other studies have examined a similar reform for women in Switzerland (Hanel and Riphahn,

from the Australian reform because it reduces the generosity of retirement benefits but does not affect the age at which individuals can start claiming. The Australian reform is more like an increase in the Early Retirement Age (ERA), a policy that is occurring in many European countries. Several recent papers have examined changes in the ERA and found large effects on employment (Börsch-Supan and Schnabel, 1998; Vestad, 2013; Staubli and Zweimüller, 2013; Manoli and Weber, 2016; Cribb, Emmerson and Tetlow, 2016; Geyer and Welteke, forthcoming), while others document spillover effects on other government programs (Staubli and Zweimüller, 2013; Geyer et al., 2018). Relative to these studies, I find smaller labor force responses and larger substitution to other government programs, and these effects are more concentrated among single and poorer households than in the small subset of studies examining heterogeneity (Staubli and Zweimüller, 2013; Cribb, Emmerson and Tetlow, 2016; Geyer et al., 2018).[3]

A second and more important contribution of this paper relates to a literature on the distributional effects of old-age social security programs. In the U.S., several studies estimate the anti-poverty effects of Social Security and other welfare programs (e.g., Scholz, Moffitt and Cowan, 2009; Meyer and Wu, 2018; Fox, 2019). This literature consistently finds that Social Security is by far the most important welfare program in reducing poverty. For example, Meyer and Wu (2018) estimate that Social Security cuts the poverty rate by a third, more than twice the size of the combined effect of five large means-tested transfers. However, these studies are purely descriptive and there have been few attempts to estimate the distributional effects of changes in retirement policy. An early exception is Engelhardt and Gruber (2004), who examine how changes in the level of Social Security income affect poverty. Using cohort variation from the Social Security notch, they estimate an elasticity of poverty with respect to changes in Social Security income close to one — that is, a 1% increase in Social Security income decreases old-age poverty by around 1%. More closely related to my paper, two recent studies examine the impact of increasing the ERA on household incomes (Geyer et al., 2018; Cribb and Emmerson, 2019). Geyer et al. (2018) study a reform in Germany that increased the ERA of women from 60 to 63 and find no effect on average household incomes, both overall and for various subgroups. However, they do not study changes in the distribution of household income or estimate the effects on measures of poverty or inequality. Cribb and Emmerson (2019) study a similar reform affecting women in the U.K. and find negative effects on household incomes and an increase in relative poverty rates by 6.4 percentage points (43%). Their estimates show proportionately larger effects

2012; Lalive, Magesan and Staubli, 2017) or exploited variation in the level of retirement benefits across cohorts in the U.S. from the Social Security Notch (e.g., Krueger and Pischke, 1992; Snyder and Evans, 2006; Gelber, Isen and Song, 2016).

[3] Staubli and Zweimüller (2013) study gradual increases in the ERA for men and women in Austria and find larger employment effects and larger spillovers into the unemployment insurance system among individuals with higher lifetime earnings. Cribb, Emmerson and Tetlow (2016) examine the labor supply response to a similar reform affecting women in the U.K. and find statistically indistinguishable responses with respect to household composition and home-ownership. Geyer et al. (2018) study a sharp increase in the ERA from 60 to 63 in Germany and find slightly larger responses from single women and women with a low-income partner. They also estimate spillovers into the unemployment insurance system, though these effects are relatively small and generally statistically indistinguishable across groups.

on low-income households, but they do not quantify the impact on inequality measures.

This paper builds on these studies in several ways. First, to my knowledge, this paper is the first to quantify the impact of these reforms on inequality measures. Given the central role of old-age pensions as a source of redistribution, it is crucial to understand how widespread increases in eligibility ages affect inequality. Moreover, the distribution regression methods applied here provide a useful framework for analyzing the impact of other treatments on measures of inequality (or other distributional statistics of interest). Second, my results highlight the importance of other government transfers in attenuating the regressive impacts of increases in pension-eligibility ages. While several studies highlight the fiscal implications of spillovers onto other government programs (Duggan, Singleton and Song, 2007; Staubli and Zweimüller, 2013; Atalay and Barrett, 2015; Oguzoglu, Polidano and Vu, 2016; Geyer and Welteke, forthcoming), there has been little consideration of the importance of these effects in mitigating the impacts on poverty and inequality. Third, the Australian context is distinct from the reforms in Germany and Britain. In Germany, the old-age pension does not provide a strong form of redistribution, as benefit amounts are essentially proportional to lifetime earnings. This contrasts with most pension systems, including contributory systems like Social Security, where the replacement rates are much higher for individuals with lower levels of past earnings. Moreover, the German reform only affected women with a relatively strong attachment to the labor market. The U.K. state pension provides a similar level of income for all households and thus provides a significant form of redistribution, though to a lesser extent than Australia's means-tested pension. The means-tested nature of Australia's pension also meant that the reform affected a group of women with a relatively weak attachment to the labor force and a relatively high reliance on other government transfers, which at least partly explains the smaller labor force responses and larger benefit-substitution effects than found in other countries.

The paper proceeds as follows. Section 2 describes Australia's retirement pension, the reform, and the expected heterogeneity in the effects on households. Section 3 describes the data and presents graphical evidence of the effects of the reform. Section 4 describes the identification strategy. Section 5 presents the estimated effects on households' labor supply and income from government transfers. Section 6 presents the estimates of the distributional effects of the reform and the effects on poverty and inequality measures. Section 7 presents the estimated fiscal savings. Section 8 concludes and discusses the implications for policy.

2 Institutional background

This section describes Australia's public retirement pension and the details of the reform. In Section 2.1, I discuss the importance of this pension, the Age Pension, in the context of Australia's retirement system. In Section 2.2, I explain the reform, which increased women's Age Pension Age (APA) from 60 to 65, and discuss the expected effects on households' labor supply and income.

2.1 The Age Pension

Australians fund their retirement through a combination of public and private sources. The public sources consist of the Age Pension and other transfer programs provided by the Australian Government. The private sources consist of voluntary savings and a mandatory defined-contribution scheme known as superannuation, which could be accessed from the age of 55 for all cohorts in my sample.[4] Private sources of retirement funding are becoming increasingly important in Australia, but most elderly Australians still partially or fully fund their retirement via the Age Pension; in 2013, the Age Pension was received by around 73% of women above the APA and 68% of men (Oguzoglu, Polidano and Vu, 2016). Age Pensioners receive payments every two weeks (fortnightly). Payment rates do not depend on the age at which an individual starts claiming, which creates a strong incentive to start claiming at the APA. Eligibility for the Age Pension is not contingent on an individual's history of employment. Instead, eligibility is based on the following conditions: (i) an age condition — recipients must have reached the APA; (ii) a residency condition — recipients must have lived in Australia for at least ten years; and (iii) a means test.

The means test determines both eligibility for the Age Pension and payment rates. It consists of two tests, an income test and an assets test, based on the current income and assets of the household. Pensioners receive the payment specified by either the income test or the assets test, whichever is lower. Each of these tests has an initial phase-out threshold. If a household's income and assets are below the relevant phase-out thresholds, pensioners receive the maximum payment, which is set at 27.7% of Male Average Weekly Total Earnings for single households and 20.88% for each eligible member of a couple (i.e., 25% less than single households). In 2016, around 60% of Age Pensioners received the maximum payment (Department of Social Services, 2016). This payment was $877.10 per fortnight for single pensioners ($22,883 per annum) and $661.20 for each eligible member of a couple ($17,250 per annum).[5] These income levels mean that a single household receiving the maximum payment, or a couple receiving two payments, has Age Pension income equal to 51% of the median household disposable income in the Australian population.[6] For households who are not eligible for the maximum payment, payments gradually decrease with income/assets above the relevant phase-out threshold. In 2016, each dollar of household income per fortnight above the income threshold reduced payments to pensioners by 50 cents (25 cents each for couples), while each $1,000 of household assets — excluding the value of the primary residence — above the asset threshold reduced payments by $1.50.[7] Households are ineligible for the Age Pension if their income or assets are above the

[4]The superannuation access age is rising to 60 for younger cohorts.

[5]All incomes are in 2016 Australian dollars (AUD). At current exchanges rates, 1 AUD $\approx$ 0.70 USD.

[6]This percentage is calculated using estimates of the median household disposable income (after adjustments for household size) in 2015–16 from the Australian Bureau of Statistics (2018*b*).

[7]These taper rates changed during the sample period. Prior to September 20, 2007, the taper rate for the assets test was $3 per $1,000 of assets. Prior to September 20, 2009, the taper rate for the income test was 40 cents per $1 of income.

income (assets) limit, set at the point where payments are reduced to zero.[8]

While the Age Pension is the main source of publicly funded retirement benefits, around one in three women are already receiving a regular transfer from the Australian Government by the time they reach the APA (Oguzoglu, Polidano and Vu, 2016). These transfers have stricter eligibility conditions than the Age Pension but are similar in many other respects. Like the Age Pension, other transfers are non-contributory, subject to similar means tests, and two of the most common transfers — the Disability Support Pension and Carer Payment — provide the same level of income, although others, such as the transfer for the unemployed (Newstart Allowance), provide significantly less. Generally, recipients of other transfers are moved onto the Age Pension at the APA, which results in either an increase in income or no change.

Older Australians also receive valuable in-kind benefits through meaningful discounts on healthcare and other expenses. Eligibility for these discounts is based on concession card ownership. There are two main types of concession cards: the Pensioner Concession Card and the Commonwealth Seniors Health Card. The Pensioner Concession Card is available to older Australians who receive a government transfer; Age Pensioners qualify automatically and nearly all recipients of other transfers qualify if they are over 60. The Commonwealth Seniors Health Card is available to those who (i) have reached the APA, (ii) are not receiving a government transfer, and (iii) satisfy an income test that is similar to the income limit for the Age Pension. Both cards provide large benefits by considerably reducing out-of-pocket medical expenses and providing other discounts.[9] For Age Pensioners, in-kind benefits are almost as valuable as the income the Age Pension provides; according to Harmer's (2008) estimates, the total value of in-kind benefits for Age Pension households is 82% as large as the amount of income these households receive from the Age Pension. Of course, the value of in-kind benefits will depend on household expenditure, especially on healthcare, with around three-quarters of the value of these benefits coming from cheaper healthcare (Harmer, 2008).

All things considered, the Age Pension is an especially important source of retirement funding for women. First of all, women are more likely to receive the Age Pension than men; in 2013, 73% of pension-age women received the Age Pension compared to 68% of pension-age men (Oguzoglu, Polidano and Vu,

[8] In 2016, the phase-out threshold for the income test was \$162 per fortnight for singles and \$292 for couples. The income limit was \$1,918.20 for singles and \$2,936.80 for couples. For homeowners, the phase-out threshold for the assets test was \$209,000 for singles and \$296,500 for couples, while for non-homeowners it was \$360,500 for singles and \$448,000 for couples. For homeowners, the asset limit was \$788,250 for singles and \$1,170,000 for couples, while for non-homeowners it was \$937,250 for singles and \$1,319,000 for couples.

[9] For example, both cards offer heavily subsidized doctor visits and large discounts on prescription medicines covered by the Pharmaceutical Benefits Scheme (PBS). The PBS subsidizes prescription medication for Australian residents. Most PBS medicines require a co-payment, which is lower for concession card holders. In 2016, the co-payment was up to \$6.20 for concession card holders and up to \$38.30 for other patients. Above a safety net threshold for each family unit of \$372 for concession card holders and \$1,475.70 for other patients, concession card holders are not charged for additional prescriptions, while other patients are charged at the concessional rate. Concession card holders also receive higher rebates from Medicare — Australia's publicly funded universal healthcare system — if they have high out-of-pocket medical expenses. Concession card holders also receive discounts on the following: public transport fares; council and water rates; electricity; and vehicle registration charges (Harmer, 2008).

2016). Second, women generally receive the Age Pension for longer; Australian women not only have a higher life expectancy than men but also had a lower APA up until July 2013. Third, women are more likely to be single in old age than men — due to their higher life expectancy and greater likelihood of having an older spouse — and singles receive more income from the Age Pension per person. Finally, women have fewer assets to self-fund their retirement, especially in superannuation. According to Clare (2015), the mean superannuation balance of men aged 60–64 was more than double that of women of the same age in 2013–14 ($292,510 compared to $138,154), and men's median balance was almost four-times larger ($100,000 compared to $28,000). These large gender gaps in superannuation balances are primarily due to older women's much lower lifetime earnings (Smith and Hetherington, 2018).

Next, I describe the 1994 reform that gradually increased women's APA from 60 to 65 and, in turn, delayed women's eligibility for the Age Pension.

2.2 The 1994 reform: Raising women's pension age from 60 to 65

Prior to the 1994 reform — introduced as part of the Social Security Legislation Amendment Act (No. 2) 1994 — the APA was 60 for women and 65 for men. The reform legislated that women's APA would gradually increase to 65, increasing by six months every two years from July 1, 1995 to July 1, 2013. Initially proposed in 1993, the reform aimed to foster employment among older women, reduce government expenditure on the Age Pension and ultimately harmonize the APA of men and women. As Figure 1 shows, the impact of the reform varied based on women's date of birth; while the reform had no effect on women born before July 1, 1935, women born on or after this date faced an increase in the APA, with larger increases for later cohorts. For example, the APA increased to 60.5 for women born between July 1935 and December 1936, and it increased to 61.0 for women born between January 1937 and June 1938. The APA continued increasing in this manner — that is, by an extra six months for each eighteen-month cohort — until the APA reached 65 for women born after December 1948.[10]

For the rest of this section, I discuss the expected effects of increasing women's APA on the labor supply and total income of affected households. In line with the empirical analysis later in the paper, I focus on the effect of women being delayed from reaching the APA, abstracting from any effects of the reform at ages where women's eligibility for the Age Pension is not affected. This approach is consistent with the majority of studies in the literature examining changes in the eligibility age for retirement benefits (e.g., Staubli and Zweimüller, 2013; Cribb, Emmerson and Tetlow, 2016; Oguzoglu, Polidano and Vu, 2016).[11]

[10] Under a reform in 2009, the APA is increasing further towards 67 for men and women. However, the estimates in this paper do not capture any effects of this reform as these increases only started affecting people's eligibility for the Age Pension in 2017, two years after the end of the sample period.

[11] Later, I investigate whether the reform affected outcomes at ages where women's eligibility for the Age Pension is not affected. However, the estimates in these regressions are relatively imprecise, and so I focus on the effects at ages where women's eligibility for the Age Pension is directly affected.

Expected effects of an increase in the Age Pension Age. Basic labor supply theory makes strong and heterogeneous predictions about the impact of delaying women from reaching the APA on their households' labor supply and total income. To explain these predictions, Figure 2 depicts the change in the household budget constraint, in terms of the set of feasible combinations of total income and leisure, when women are delayed from reaching the APA. Figure 2 shows the respective budget constraints of households with women who are above the APA (in black) and below the APA (in gray), assuming that households are eligible for the maximum Age Pension payment under the assets test and ignoring the presence of in-kind benefits and other government transfers. We can think about the effect of delaying women from reaching the APA as shifting the household budget constraint down from the black line to the gray line.

The shift in the budget constraint varies with household income. Because of the income test for the Age Pension, the shift is largest at the bottom of the income distribution, and there is no shift at the top of the distribution. At zero or low levels of hours worked (to the right of H_1 on Figure 2), where households earn less than the phase-out threshold for the income test, the budget constraint shifts down by the amount of the maximum pension payment. For a slightly higher level of hours (between H_1 and H_2), where households earn more than the phase-out threshold but less than the income limit, the shift in the budget constraint is smaller as households are not eligible for the maximum payment. In addition, the budget constraint becomes twice as steep in this region — that is, the effective wage rate doubles — because the income test reduces pension payments by 50 cents for each dollar of earnings above the phase-out threshold. Finally, for hours above H_2, where households earn more than the income limit, there is no effect on the budget constraint.

We can use this framework to make predictions about the impact of the reform on the earnings and total income of households. For simplicity, I assume a unitary model of the household, where women and their spouses make a joint labor supply decision. Otherwise, I consider a similar static model of labor supply to Bitler, Gelbach and Hoynes (2006): I assume households can freely choose hours of work at the given wage, which is constant, and I ignore any human capital or search-theoretic issues.

Consider first the effects on a household who would have located at point **A** (i.e., worked zero hours) if not for the delay in women reaching the APA. These households face a negative income effect, with their budget constraint shifting down by the amount of the maximum pension payment. As these households are at a corner solution, they may either continue working zero hours or enter the labor force. We would expect a decrease in total income for both types of households, as long as consumption is a normal good.

Next, consider the effects on a household who would have located at point **B**. These households, who have positive earnings but earn less than the phase-out threshold, also face a negative income effect equal to the amount of the maximum pension payment. Assuming that both leisure and consumption are normal goods, we would expect an increase in labor supply and a decrease in total income.

Now consider the effects on a household who would have located at point **C**. These households, who earn more than the phase-out threshold but less than the income limit, face a negative income effect and a substitution effect that doubles the effective wage rate. We would expect both of these effects to encourage households to increase their labor supply. The effect on total income is ambiguous.

Finally, consider the effects on a household who would have located at point **D**. As there is no effect on the budget constraint of these relatively high-income households, we would not expect the reform to affect their labor supply or total income.

The set of points $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ exhausts all of the qualitatively different combinations of total income and leisure for households in the counterfactual scenario, where women are not delayed from reaching the APA. For each of these households, the table below Figure 2 summarizes the expected effects on earnings and total income (in columns 4 and 5). Column 4 shows that we expect no effect on earnings for (i) all high-earning households and (ii) some non-working households, and we expect an increase in earnings for households in between. Column 5 shows that we expect a decrease in total income at the bottom of the income distribution, no effect at the top of the distribution, and the effect is ambiguous in between.

Three additional insights inform the predictions. First, because of other payment rules, the impact on the budget constraint is largest for single women and poorer households, and there is no impact for households with assets above the asset limit. As such, we would expect the effects on labor supply and total income to be larger among single women and poorer households. Second, previous studies of the same reform by Atalay and Barrett (2015) and Oguzoglu, Polidano and Vu (2016) find a significant increase in women's receipt of other transfers, which may attenuate the effects of the reform on households' labor supply and income. Third, the reform is also likely to delay many women's eligibility for the generous in-kind benefits associated with concession card ownership. Eligibility for concession cards is either directly linked to being above the APA, in the case of the Commonwealth Seniors Health Card, or indirectly linked to being above the APA in the case of the Pensioner Concession Card.

3 Data and descriptives

In this section, I describe the sample and present graphical evidence of the effects of the reform. In Section 3.1, I discuss the sample, which comes from the Household Income and Labour Dynamics in Australia (HILDA) Survey, and describe how the key variables are defined. In Section 3.2, I show graphical evidence of the effects of the reform on women's labor supply and income from transfers.

3.1 Sample construction and descriptive statistics

The HILDA Survey. My analysis is based on an unbalanced panel of respondents to waves 1–14 of the HILDA Survey who are female and aged 60–66 years old at the survey date. HILDA is Australia's only large, nationally representative, household-based longitudinal survey. HILDA is an annual survey that began in 2001 and has since spawned a large body of research (Watson and Wooden, 2012). Several features of HILDA make it ideal for this analysis. First, HILDA has exact date of birth and survey date information. This precision is crucial for my identification strategy, which exploits quasi-experimental variation in whether women have reached the APA, because it allows me to precisely identify women's APA and establish whether they have reached the APA at the survey date. Second, HILDA contains detailed information on the income and labor supply of each member of the household, allowing me to estimate the effects of the reform on the incomes and labor supply of women and their spouses. Third, HILDA contains detailed information on household characteristics — such as home-ownership status and wealth — allowing me to examine how the effects of the reform varied with respect to these characteristics. No other Australian dataset has all three of these features.

Key outcome variables. To start, I examine the effects of the reform on women's receipt of the Age Pension and other government transfers. I construct indicator variables specifying whether women receive the following transfers at the survey date: (i) the Age Pension; (ii) any other government transfer; and (iii) any government transfer.[12] Then, I examine the effects on women's labor supply using the following measures: (i) an indicator for labor force participation (being employed or unemployed at the survey date); (ii) an indicator for employment; and (iii) the number of hours of paid work in an average week.

Next, I estimate the effects on women's income from each source, using the following annualized measures of income: (i) income from the Age Pension; (ii) income from other government transfers; and (iii) earnings from wages and salaries.[13] I convert incomes to 2016 dollars using the Male Average Weekly Total Earnings index, the benchmark index for the Age Pension's payment rates. The estimates are similar using the Consumer Price Index.

Finally, I use all of the above variables to examine the effects of the reform on women's spouses.

Sample restrictions and descriptives. To select the sample, I start by including all women aged 60–66 years old at the survey date, which results in 8,748 observations and 2,112 individuals. Then, I make four additional restrictions. First, I exclude women who have lived in Australia for less than ten years as these women will not satisfy the residency condition for the Age Pension. Second, I exclude observations with missing information on women's receipt of government transfers, employment status, income or years of schooling (which is used as a control variable). Third, I exclude observations on partnered women if their

[12] I only include government transfers that are classified as income support payments. Including other transfers results in similar estimates as these transfers are (i) much less generous and (ii) not an obvious substitute for the Age Pension. Individuals are not eligible to receive multiple income support payments at the same time, so variable (iii) is equal to the sum of variables (i) and (ii).

[13] For (iii), I exclude individuals who ever had earnings in the top 1% of workers in the sample to reduce the sensitivity of the estimates to changes at the top of the earnings distribution. The estimated effect on women's earnings is 18% smaller without this restriction.

spouse is female or, more often, when information on their spouse is missing. Finally, I exclude a handful of women who ever reported receiving the Age Pension before they had reached the APA (as this is not possible). The final sample consists of 7,999 observations and 1,945 individuals.

Table 1 summarizes the characteristics of the sample. As shown in column 1, the sample consists of women with an average age of 63.4 years. Around two-thirds are partnered (66.3%) and a high proportion are homeowners (84.2%). Only a quarter completed year 12 of high school (25.5%), while three-quarters completed year 10 (74.7%). Nearly 50% receive a transfer from the Australian Government (48.2%), with 27.6% receiving the Age Pension, while only 36.1% participate in the labor force. Conditional on receiving a transfer, women receive \$14,831 per annum from transfers; conditional on employment, women earn \$45,102 per annum from wages and salaries. Table 1 also summarizes the characteristics of women's spouses. Spouses are slightly older, with an average age of 65.7 years. Nonetheless, a higher proportion of spouses participate in the labor force (42.6%) and, conditional on employment, spouses have higher average earnings (\$63,614 per annum).

Columns 2 and 3 show how the characteristics of the sample has changed over time, presenting the means of these variables in waves 1 and 14 respectively. Although the average age of women has not changed, the proportion of women receiving the Age Pension has fallen markedly, from 46.8% to 15.5%. Women's labor force participation has also risen significantly, from 23.0% to 41.4%, while their overall receipt of government transfers has fallen, from 60.3% to 40.5%. Note that while the reform may partially explain these trends, there is also a strong increase in women's level of education, which indicates that cohort factors are likely to be important.

3.2 Graphical evidence: Trends around the Age Pension Age

This section presents graphical evidence of the effects of the reform on women's labor supply and income from transfers. Figure 3 shows the age-related trends in women's labor force participation and transfer receipt for different cohorts. Figure 3 divides women into three cohorts based on their APA: (i) APA $\in [61.0, 62.0]$; (ii) APA $\in [62.5, 63.5]$; and (iii) APA $\in [64.0, 65.0]$. Figure 3a shows large differences across these cohorts in women's Age Pension receipt at ages 62–64. For example, at age 63, Age Pension receipt is around 60% for women in the first cohort, 40% for women in the second cohort, and 0% for women in the third cohort. However, these differences largely disappear once each cohort has reached the APA, indicating that the reform has delayed Age Pension receipt for many women with a later APA.

However, Figure 3b shows that many women offset their income losses from the Age Pension by extending their receipt of other transfers to later ages. There are large differences in women's receipt of other transfers at ages 62–64, with higher rates among women with a higher APA, but these differences disappear at age 65 when each cohort has reached the APA.

Nonetheless, Figure 3c indicates that the reform delayed transfer receipt for a large proportion of women. While each cohort has similar rates of transfer receipt at age 65, there are considerable differences at ages 62–64. For example, at age 63, women with an APA between 64.0 and 65.0 are around 35 percentage points less likely to receive a transfer than women with an APA between 61.0 and 62.0, but the gap is less than 10 percentage points at age 65.

Despite these visually apparent effects on women's transfer receipt, Figure 3d shows little change in women's labor force participation as they reach the APA. In fact, a more striking feature of Figure 3d is the difference in labor force participation rates across cohorts at ages 60 and 61, with higher participation rates for women with a higher APA. While these differences could reflect anticipatory responses to the increase in the APA, there is strong evidence that other factors explain most of these effects. Morris (2019) documents a strong trend prior to the reform in female labor force participation rates across the relevant cohorts. This trend is consistent with trends in education, birth rates and attitudes. Figures A2 and A3 show that women with a higher APA have higher levels of education, fewer children and more progressive attitudes about the intra-household allocation of paid work and child-rearing responsibilities.

To show the effects around the APA more clearly, Figure 4 presents women's mean outcomes with respect to the number of years between their age and their APA, i.e. $\lfloor age - APA \rfloor$, where $\lfloor \rfloor$ is the floor function. Figure 4a shows that women's receipt of the Age Pension increases by nearly 50 percentage points in the year they reach the APA, but the increase in the proportion of women receiving any transfer is much lower, less than 20 percentage points, as many women move onto the Age Pension from other transfers. Figure 4b shows analogous changes in women's income; on average, women's income from the Age Pension increases by around $7,000 per annum at the APA, while their total income from transfers increases by around $2,500. There remains only weak evidence of a decrease in female labor supply when women reach the APA; while Figures 4c and 4d show a decrease in women's labor force participation and earnings at the APA, the decrease is consistent with the trend prior to the APA (as women age).

In addition to showing preliminary evidence of the effects of the reform, Figures 3 and 4 highlight the importance of age and cohort factors on women's outcomes near the APA. As such, I carefully control for these factors. The next section describes my identification strategy in detail.

4 Identification strategy

I use quasi-experimental variation in whether women have reached the APA to estimate the causal effects of increasing women's APA on households' income and labor supply. Specifically, I estimate the causal effects of women being delayed from reaching the APA — and thus remaining ineligible for the Age Pension

— because of the phased increases in women's APA from 61.5 to 65.0 under the 1994 reform.[14] I use a differences-in-differences approach that is similar to the approach of recent studies evaluating the effects of increases in the eligibility age for retirement benefits (Staubli and Zweimüller, 2013; Cribb, Emmerson and Tetlow, 2016; Oguzoglu, Polidano and Vu, 2016). My estimates compare the outcomes of women who remain below the APA because of the reform to the outcomes of women in earlier cohorts who had reached the APA when they were the same age. In this context, the treatment is women being below the APA because of the phased increases in the APA from 61.5 to 65.0. Table A1 shows that all women in the sample who have an APA between 62.0 and 65.0, and some women with an APA of 65.5, are treated. The age at which these women are treated ranges from 61.5 to 64.5 (in half years), with later cohorts treated for a longer period of time and at older ages. For example, women born between January 1940 and June 1941 are treated at age 61.5 because of the increase in their APA to 62.0, while women born between January 1949 and June 1952 are treated at ages 61.5–64.5 because of the increase in their APA to 65.0.

I implement my approach using the following type of regressions:

$$y_{it} = \beta x_{it} + \delta \mathbf{1}(age_{it} < APA_i) + FE\_age\_0.5yrs_{it} + FE\_APA\_cohort_i + \varepsilon_{it} \tag{1}$$

where i denotes a female individual (or her spouse/household); t denotes the survey date; y_{it} is the outcome variable of interest; $FE\_age\_0.5yrs_{it}$ is a set of fixed effects for each woman's age in half years at time t to control for constant age-related factors affecting y_{it}; $FE\_APA\_cohort_i$ is a set of fixed effects for each woman's APA cohort to control for constant unobservable differences in y_{it} across cohorts, which is important given the strong trends across female cohorts in levels of education, birth rates, attitudes and labor force participation; and x_{it} includes controls for each woman's household size, years of schooling, number of children, marital status, state of residence and the monthly state-level unemployment rate.[15]

The key explanatory variable identifying the treatment effect is $\mathbf{1}(age_{it} < APA_i)$, which is equal to one for women who are below the APA and zero otherwise. This variable is fully explained by the combination of women's APA cohort and age in half years. As equation (1) includes fixed effects for these variables, δ identifies the average effect on y_{it} of women being below the APA at ages 61.5–64.5, the age range where women's eligibility for the Age Pension varies across cohorts in the sample.

We can interpret δ as a causal effect if the following parallel-trends assumption holds: different cohorts of women would have had identical age-related trends in y_{it} if not for the differences in their APA. Two features of the Australian context add credibility to this assumption. First, different cohorts faced similar

[14] I cannot estimate the effects of the increases in women's APA from 60.0 to 61.5. As the sample starts in 2001, there is no variation in whether women have reached the APA below the age of 61.5.

[15] I estimate all regressions by OLS. I cluster standard errors at the level of the female individual to allow for an arbitrary correlation between each woman's errors (or those of her household/spouse). Standard errors are similar without clustering.

macroeconomic conditions at each age. Over the sample period, Australia's economy has been very stable. Australia has not had a recession since 1991, and its economy was relatively unaffected by the 2008 financial crisis. Second, besides the changes to the APA, female cohorts faced a similar set of retirement policies at each age, as other changes to retirement policy during the sample period were relatively minor. Nonetheless, I examine the validity of the parallel-trends assumption empirically. This analysis, discussed in Section 5.1, provides no evidence against the parallel-trends assumption.

Finally, note that estimates of δ are conditional on any responses before women reached age 61.5. As the reform was announced well before women in the sample reached this age, households had time to adjust their behavior in advance. For most outcomes, we would expect any anticipatory responses to be in the same direction as the effects of women being delayed from reaching the APA. For example, households may have increased their labor supply before women reached age 61.5 and when women were just below the APA. This type of anticipatory behavior may attenuate estimates of δ as these responses are likely to be absorbed by the cohort fixed effects. However, the findings of Geyer and Welteke (forthcoming) indicate that such effects may be small, at least in terms of labor supply. Using a regression discontinuity design, Geyer and Welteke (forthcoming) examine a sharp increase in the pension-eligibility age of women in Germany from 60 to 63. They find a precisely estimated zero effect on employment prior to age 60 despite (i) very large effects at ages 60–61 and (ii) the fact that the reform was well anticipated. Nonetheless, I test for anticipatory responses at ages 55–61 in Appendix A.1. I do not find any statistically significant evidence of such responses. However, as the estimates are relatively imprecise, I cannot rule out sizeable effects.

5 Effects on labor supply and government transfers

In this section, I present the estimated effects of the reform on households' labor supply and income from government transfers. In Section 5.1, I start by discussing the estimates of the average impacts on women. The estimates indicate that many women remained on other government transfers for longer but show relatively little increase in labor supply. Then, I verify the robustness of the estimates, examine heterogeneity in the effects and present the estimated effects on spouses. The estimates show larger effects on single women and women in poorer households, and there is little evidence of any effect on the labor supply or income of women's spouses. In Section 5.2, I show that the labor supply response was confined to part-time and low-earning women, consistent with the predictions from theory.

5.1 Average impacts

Table 2 presents the estimated effects of the reform on women's receipt of the Age Pension and other government transfers (in columns 1–3). Consistent with previous studies of the reform (Atalay and Barrett, 2015; Oguzoglu, Polidano and Vu, 2016), the estimates indicate that many women offset their delayed receipt of the Age Pension by extending their receipt of other transfers. While the reform decreased Age Pension receipt among women by an estimated 48.3 percentage points ($p < 0.01$), their receipt of other transfers increased by an estimated 30.0 percentage points ($p < 0.01$). Overall, the proportion of women receiving any transfer decreased by an estimated 18.2 percentage points ($p < 0.01$), a relatively small effect compared to the decrease in Age Pension receipt.

Table 2 also presents the estimated effects on women's labor supply (in columns 4–6). Overall, the estimates indicate that the reform caused only a modest increase in female labor supply overall. In terms of the effects on women's employment status, the estimated increase in women's labor force participation is 3.1 percentage points ($p < 0.1$), with most of this increase explained by a 2.7-percentage-point increase in employment ($p = 0.121$). In terms of the effects on hours worked, the estimate in column 6 indicates that women worked an additional 0.6 hours per week on average ($p \geq 0.1$).

Overall, these estimates are considerably smaller than previous estimates of the reform by Atalay and Barrett (2015), who estimate a 12.0-percentage-point increase in female labor force participation. However, Morris (2019) shows that this estimate may be biased upwards by a pre-existing trend in female labor force participation rates across the relevant cohorts. As I directly control for differences in labor force participation rates across cohorts, my estimates are considerably smaller. If I exclude the fixed effects for women's APA cohort, the estimated increase in labor force participation is four-times larger — 12.9 percentage points — similar to Atalay and Barrett's (2015) estimate (not shown).

Table 3 presents the estimated effects on women's income. While women's income from the Age Pension decreased by an estimated \$6,957 per annum ($p < 0.01$), their income from other transfers increased by an estimated \$4,365 per annum. Hence, on average, women offset 63% of their income losses from the Age Pension through other transfers and lost an estimated \$2,592 per annum from transfers overall. Column 3 presents the estimated effect on earnings. On average, women's earnings increased by an estimated \$1,458 per annum ($p = 0.102$). This marginally insignificant estimate indicates that, on average, women earned an extra 21 cents for each dollar of income lost from the Age Pension. Finally, column 4 shows that women's combined income from transfers and earnings decreased by an estimated \$1,177 per annum, though this estimate is not statistically significant at the 10% level.

Robustness checks. The main way I verify the robustness of the estimates is to examine precisely when each outcome changes relative to women's APA. If the delay in women reaching the APA *causes* changes in an outcome, y_{it}, we would expect to see two patterns in y_{it}. First, after controlling for age and cohort factors, we would not expect to see any trend in y_{it} prior to women's APA. For example, the effect of women being one year below the APA should be similar to the effect of them being two years below the APA. Second, we would expect to see sharp changes in y_{it} at the APA. To verify that these two patterns are present for the main outcomes, I estimate the following regressions:

$$y_{it} = \beta x_{it} + \delta_{-5}\mathbf{1}(age_{it} < APA_i - 4) + \sum_{\substack{j=-4 \\ j \neq -1}}^{0} \delta_j \mathbf{1}(age_{it} \in [APA_i + j, APA_i + j + 1))$$
$$+ \delta_1 \mathbf{1}(age_{it} \geq APA_i + 1) + FE\_age\_0.5yrs_{it} + FE\_APA\_cohort_i + \varepsilon_{it} \quad (2)$$

where these regressions modify equation (1) by replacing the key right-hand-side variable, $\mathbf{1}(age_{it} < APA_i)$, with six indicator variables that each specify that a woman is a certain number of years above or below the APA. Each of the δ coefficients on these indicators should be interpreted as an effect relative to the omitted category, which is women being less than or equal to one year below the APA. For example, for $j \in \{-4, -3, -2, 0\}$, δ_j estimates the effect on y_{it} from women being j to $j+1$ years above the APA, i.e. $age_{it} \in [APA_i + j, APA_i + j + 1)$, relative to them being less than or equal to one year below the APA, i.e. $age_{it} \in [APA_i - 1, APA_i)$. If women being delayed from reaching the APA causes changes in y_{it}, we would expect δ_j to be statistically significant for $j \geq 0$ but not for $j < -1$.[16]

Figure A4 presents the estimated δ coefficients with 95% confidence intervals. Overall, the estimates are consistent with a causal interpretation of the estimates in Tables 2 and 3. First of all, the estimates show no trend in women's outcomes in the five years prior to the APA; none of the δ coefficients prior to the APA are statistically significant at the 5% level, either individually or jointly. In addition, the estimated effects on women's outcomes at the APA are consistent with the estimates in Tables 2 and 3. For example, at the APA, the estimates show the following: (i) a sharp and statistically significant increase in women's income from the Age Pension ($p < 0.01$); (ii) a sharp but slightly smaller decrease in women's income from other transfers ($p < 0.01$); (iii) a relatively small and statistically insignificant decrease in women's earnings ($p \geq 0.1$); and (iv) a modest increase in women's combined income from transfers and earnings ($p < 0.1$).

In Table A2, I verify the robustness of the estimates further. First, I show that the estimates are similar when I augment the regressions with a set of fixed effects for each survey wave, indicating that the estimates are unlikely to be strongly affected by temporal factors, such as macroeconomic shocks or other policy changes, affecting cohorts at different ages. Second, I show that the estimates are similar with controls for women's physical and mental health at the time of each survey, indicating that the estimates are unlikely to be strongly affected by different age-related trends in health across cohorts.

Heterogeneity by household characteristics. As the Age Pension provides higher payments to single and low-asset households, we expect a stronger impact on these households. The estimates in Tables A5 and A6 support these predictions. First, the estimates indicate that the income losses from the Age Pension

[16]To enhance the precision of these estimates, I expand the age sample to 55–69 for these regressions (but only include women who are ever in the main sample).

were considerably larger among single women and women in poorer households. The estimates in Table A5 indicate that single women lost 89% more Age Pension income than partnered women. This is partly due to the 33% higher maximum rate for singles, and partly due to single women's lower levels of household wealth (see Table A4). The estimates in Table A6 indicate that (i) renters lost more than twice as much Age Pension income as homeowners and (ii) women in the poorest third of households lost around twice as much as women in the middle third and more than three-times as much as women in the wealthiest third.[17] Second, the estimated increase in income from other transfers is much larger among single and poorer groups of women. Moreover, the estimates suggest that poorer women offset a larger proportion of their income losses through other transfers. For example, women in the poorest third of households offset 74% of their Age Pension losses with other transfers, while women in the top third offset just 34%. Third, the estimated increase in labor supply is larger among single and poorer groups of women, though these differences are generally not statistically significant. Interestingly, there is little evidence of any increase in labor supply among partnered women and women in the top two thirds of the wealth distribution.

Spousal responses. Table A5 presents the estimated effects on women's spouses in Panel B. All of the estimates are fairly close to zero and none are statistically significant at the 10% level.[18] For example, the estimated increase in the labor force participation of spouses is just 0.3 percentage points ($p \geq 0.1$). Thus, there is no evidence of an 'added worker effect' (Lundberg, 1985), whereby spouses offset a decrease in their partner's income by increasing their own earnings. However, this result is not surprising given the small and statistically insignificant estimates on the labor supply of partnered women. Moreover, this result is consistent with the literature, which generally finds little evidence that male spouses respond to changes in the pension incentives of their partner (Selin, 2017; Lalive and Parrotta, 2017; Geyer et al., 2018).

Nonetheless, this finding contrasts with the recent work of Atalay, Barrett and Siminski (2018), who study the same reform and estimate an increase in the participation rates of spouses of at least 6.8 percentage points (which is outside the 99% confidence interval for my estimate). This difference appears to stem from the fact that my regressions include fixed effects for the birth cohorts of men *and* their wives. Without the fixed effects for the birth cohort of wives, the estimated increase in spousal labor force participation increases from 0.3 to 7.5 percentage points (not shown). In this context, it is important to include fixed effects for the birth cohorts of wives. As shown in Morris (2019), there is a strong trend in female labor force participation rates across the relevant cohorts prior to the reform, with much higher participation rates

[17] I divide women into thirds based on their households' net worth in wave 2, the first wave containing information on wealth. To compare the household wealth of single women and couples, I divide household wealth by 1.5 for couples, the equivalence scale implied by the Age Pension payment rules. This adjustment reflects both the additional expenditure needs of couples and the returns to scale from living with a partner, in terms of the ability to share housing and other costs.

[18] Figure A5 demonstrates the robustness of the estimated effects on spouses, presenting estimates of equation (2) for the relevant spousal variables. The estimates show no trend in the outcomes of male spouses prior to the APA of their wives and no change at the APA.

for later cohorts, and the higher participation rates of later cohorts could directly increase the participation rates of their husbands due to preferences for joint retirement. Such preferences have been emphasized in several studies to explain the tendency of couples to retire close together regardless of their difference in age (e.g., Gustman and Steinmeier, 2000, 2004; Coile, 2004).

5.2 Labor supply responses across the distribution of hours worked/earnings

All of the labor supply analysis so far has estimated mean effects on employment, hours worked and earnings. However, because of the income test for the Age Pension, we expect systematic heterogeneity in the labor supply response across the distributions of hours worked and earnings. Therefore, in this section, I examine the effects on the *distributions* of hours worked and earnings using distribution regression, an attractive and increasingly common alternative to quantile regression (e.g., Rothe, 2012; Chernozhukov, Fernández-Val and Melly, 2013; Dube, 2019). First proposed by Foresi and Peracchi (1995), distribution regression involves estimating a sequence of binary regressions, where in each regression the dependent variable is an indicator for the variable of interest — in this case hours or earnings — being above/below a given threshold, and the effects are estimated at a range of thresholds. This is analogous to the common practice of estimating quantile regressions at a range of quantiles.

In this context, I favor distribution regression over quantile regression for two reasons (though quantile regression estimates are consistent with the results). First, Chernozhukov, Fernández-Val and Melly (2013) show that distribution regression outperforms quantile regression in the presence of mass points, which is relevant here as around 65% of women in the sample do not work. Second, distribution regression allows the response to vary with the level of earnings, rather than a woman's rank in the conditional distribution, which is desirable because the incentives to work vary with earnings due to the income test for the Age Pension.

To implement this approach, I estimate equation (1) many times via OLS.[19] In each regression, the dependent variable is an indicator for hours or earnings being above a given threshold, say x. For hours worked, I set x at two-hour intervals from zero to 50 hours per week; for earnings, I set x at \$2,500 intervals from \$0 to \$100,000 per annum.

Figure 5 presents the estimates along with 95% confidence intervals. Overall, the estimates are modest in size but consistent with the theoretical predictions. The estimates show the strongest evidence of an increase in labor supply from part-time and relatively low-earning women, and there is little evidence of any effects on the labor supply of full-time and high-earning women. For example, for hours-thresholds up to 20 hours per week, the estimates are consistently positive, equal to around 3.0 percentage points and either statistically significant at the 10% level or marginally insignificant; above 20 hours per week, the estimates are smaller, close to zero and statistically insignificant at the 10% level. There is a similar pattern in the ef-

[19] The average marginal effects are similar using a logit model (see Figure A6).

fects for earnings. For earnings-thresholds up to $40,000 per annum, the estimates are consistently positive, equal to around 3.0 percentage points, and either statistically significant at the 10% level or marginally insignificant; above $40,000 per annum, the estimates are smaller and mostly insignificant at the 10% level.[20]

6 Distributional effects on household incomes

In this section, I examine the distributional effects of the reform on household incomes. In Section 6.1, I explain the key variable used for analysis: household disposable income after adjustments for housing costs and the value of in-kind benefits (called "adjusted household income"). In Section 6.2, I show how the effects on household incomes varied across the distribution and quantify the effects on inequality measures. In Section 6.3, I estimate the effects on relative poverty rates. Finally, in Section 6.4, I examine the sensitivity of the estimates to different assumptions about the value of in-kind benefits.

6.1 Accounting for the value of in-kind benefits and housing costs

While measures of inequality and relative poverty are usually defined solely based on household income, I also use information on housing costs and the value of in-kind benefits for different households. Considering in-kind benefits is important here because many Australians become eligible for generous concessions on healthcare and other expenses at the APA, and these benefits are disproportionately received by lower income households (Harmer, 2008). Considering housing costs is important for the measurement of relative poverty. Relative poverty lines are typically set based on a benchmark level of income in the entire population (usually 50% or 60% of the median). However, older Australians generally have much lower housing costs than the rest of the population and thus require less income to sustain a given level of consumption. Moreover, housing costs vary considerably among older Australians, especially between homeowners and renters, and accounting for this variation results in more sensible comparison of the effects on these groups.

HILDA contains comprehensive information on two of the three variables required for this analysis: household disposable income and housing costs. I define household disposable income as the combined disposable income of women and their spouses in the financial year prior to the survey date (as disposable income is not measured in the data on a current basis). I define housing costs as the sum of mortgage repayments and rent, plus an additional amount for homeowners to reflect the fact that some costs, such as council rates, are only paid by homeowners.[21]

Unfortunately, HILDA lacks information on the third variable: the value of in-kind benefits for different households. I address this issue using the 2003–04 and 2009–10 cross-sections of the Australian Bureau

[20] There is no evidence of any effects on spousal labor supply at any point in the distribution (see Figure A7).

[21] Specifically, I increase the annual cost of housing for homeowners by $2,474, which is the average cost of housing after mortgage repayments for homeowners in the 2015–16 Australian Bureau of Statistics' Income and Housing Costs Survey.

of Statistics' Fiscal Incidence Study. The Fiscal Incidence Study is a nationally representative survey conducted every six years. It estimates the value of in-kind benefits available to individual households from the provision of the following services: health; education; housing; social security and welfare services; and electricity concessions and rebates. The estimates are based on (i) the total cost of providing these services for governments at the local, state and federal level and (ii) the reported expenditure of households on a comprehensive list of items asked in the accompanying Household Expenditure Survey.

I use these estimates to construct an imputed value of in-kind benefits for different types of households in HILDA. I allow the value of in-kind benefits to vary with household composition and concession card ownership. Specifically, I pool the 2003–04 and 2009–10 surveys and calculate, in 2016 dollars, the average total value of in-kind benefits for the following types of households containing women aged 60–64: (i) single women who do not own a concession card; (ii) single women who own a concession card; (iii) couples in which neither spouse owns a concession card; (iv) couples in which one spouse owns a concession card; and (v) couples in which both spouses own a concession card. As expected, the value of in-kind benefits is higher for concession card owners (see Table A7). For example, the value of in-kind benefits is 66% higher for single women who own a concession card than those who do not (\$15,559 compared to \$9,365).

To allocate women in HILDA into these categories, I first have to impute concession card ownership for women and their spouses for the previous financial year (to align with the time frame of the information on disposable income). Fortunately, it can be imputed accurately based on respondents' age, receipt of qualifying transfers and household income (see Appendix A.2). Then, for the baseline estimates, I set the total value of in-kind benefits for each household in HILDA equal to the average of households in their category from the Fiscal Incidence Study (see Appendix A.3 for more details).

Finally, I define the key variable, adjusted household income, as household disposable income *plus* the total value of in-kind benefits *less* housing costs, adjusted for household size.[22]

Figure A8 shows how accounting for in-kind benefits and housing costs affects the distribution of household income among women in the sample. Evidently, accounting for in-kind benefits shifts the income distribution to the right, and the shift is largest near the bottom of the distribution as in-kind benefits provide a significant form of redistribution among older households. In contrast, accounting for housing costs has an opposing (though relatively modest) effect, shifting the income distribution to the left, and the shift is larger towards the bottom of the distribution as low-income households are less likely to own their home outright. Figure A8 also shows the impact of these adjustments on the level of inequality in each distribution, as measured by the Gini coefficient — a standard measure of inequality ranging from zero to one. Based on household disposable income alone, the Gini coefficient is equal to 0.433, around 40% higher than the

[22]To adjust for household size, I divide household income by 1.5 for couples, the equivalence scale implied by the payment rules for the Age Pension.

coefficient for the Australian population overall, indicating that there is a relatively high level of income inequality among households in the sample. However, the Gini coefficient for adjusted household income is considerably lower (0.335) because of the redistributive nature of in-kind benefits.

6.2 Effect on inequality measures

In this section, I estimate the impact of the reform on inequality measures. I start with an overview of my four-step approach. *Step 1*: I estimate distribution regressions to estimate the effect of the reform on the distribution of adjusted household income at a grid of points across the distribution. *Step 2*: I use the distribution regression estimates from *Step 1* to estimate the counterfactual distribution of household income for treated women. *Step 3*: I estimate the effect of the reform on inequality measures by comparing such measures for the actual distribution among treated women (the "treated distribution") and the counterfactual distribution. *Step 4*: I use bootstrapping to determine the statistical significance of the estimated effects.

To implement the distribution regressions (*Step 1*), I estimate similar regressions to those described by equation (1). In each regression, the dependent variable is an indicator variable for adjusted household income being *less* than a given cut-off, say x, and I set x at $2,500 intervals up to one million dollars. As the dependent variable here contains information for the financial year prior to the survey date, rather than information that is current at the survey date, I modify the right-hand-side of equation (1) in three ways. First, I replace the key variable — the indicator for women being below the APA at the survey date — with the fraction of time they were below the APA in the previous financial year. Second, I replace the fixed effects for women's age in half years at the survey date with fixed effects for their age in half years at the start of the previous financial year. Third, I replace the state-level unemployment rate in the survey month, a control variable, with the average state-level rate in the previous financial year.

Figure 6a presents the distribution regression estimates with 95% confidence intervals. For clarity, I only present the estimates up to $125,000, which captures around 95% of the sample; above this threshold, the estimates are small and statistically insignificant at conventional levels (as expected).

The estimates indicate that the reform had a strong negative impact on the distribution for household incomes below $47,500 per annum, but the effect is negligible at higher levels of income. The estimates are positive and statistically significant at the 5% level between $25,000 and $47,500 (corresponding to the 9$^{\text{th}}$ and 54$^{\text{th}}$ percentiles among treated women), indicating that the reform increased the proportion of households with household incomes below these relatively modest cut-offs. In contrast, the estimates are close to zero and statistically insignificant at higher levels of income. Overall, these effects are broadly consistent with the theoretical predictions in Figure 2, in that there is a strong negative effect on the incomes of low-to-middle-income households but no effect on higher income households.

To construct the counterfactual distribution (*Step 2*), I start by estimating the distribution for treated women (i.e., the proportion of treated women with household incomes less than x at each value of x in the income grid). Here, and for the rest of Section 6, I define treated women as women who (i) had an APA above 61.5 and (ii) were below the APA and aged at least 61.5 for at least 50% of the previous financial year. Then, I construct the counterfactual distribution by subtracting the distribution regression estimates in Figure 6a from the treated distribution at each value of x.[23] Figure 6b presents the treated and counterfactual distributions. Below the sample median, the counterfactual distribution is consistently to the right of the treated distribution, and there is relatively little difference in the distributions above the median.

To quantify the overall regressivity of the reform (*Step 3*), I compare inequality measures for the treated and counterfactual distributions. I use two measures of inequality: the Gini coefficient and the 80:20 ratio, which is equal to the ratio of incomes at the 80$^{\text{th}}$ percentile to those at the 20$^{\text{th}}$ percentile. Both measures are higher in the treated distribution; the Gini coefficient is 12% higher (0.346 compared to 0.310), and the 80:20 ratio is 15% higher (2.312 compared to 2.004). This implies that the reform increased inequality measures by 12–15% among treated women. This is a significant increase. A 12% increase in the Gini coefficient is equivalent to an annual transfer of $8,100 from households in the bottom quintile of the sample to households in the top quintile, while a 15% increase in the 80:20 ratio requires a transfer of $3,100. These transfers amount to 13–34% of the average income of households in the bottom quintile.

To determine whether these estimates are statistically significant (*Step 4*), I use a pairs cluster bootstrap that is similar to the approach described by Cameron and Miller (2015). Specifically, from my sample, I re-sample individuals with replacement 999 times. For each replication, I repeat *Steps 1–3* and calculate the estimated effect on each measure of inequality. Then, I calculate p-values for the estimated effects under a two-tailed test as $p = 2n/(999+1)$, where n is the number of replications in which the estimated effect on the inequality measure is non-positive. The p-value is equal to 0.048 for the Gini coefficient and $<$0.002 for the 80:20 ratio (as none of the replications produced a non-positive estimate). Hence, both estimates are statistically significant at the 5% level.

6.3 Effect on relative poverty rates

In this section, I focus on the effects towards the bottom of the income distribution, estimating the impact of the reform on relative poverty rates among affected households. For developed countries such as Australia, where extreme poverty is rare, using a relative measure of poverty is the standard approach taken by the OECD and the EU (OECD, 2015; Bradshaw and Mayhew, 2011). I follow the standard approach, which dating back to Fuchs (1969), sets the poverty line equal to a percentage — generally 50% or 60% — of median household income in the entire population. Namely, I consider households to be in poverty if their

[23] I also restrict the counterfactual distribution to be monotonic (or non-decreasing in income) by replacing any 'valleys' in the counterfactual distribution with flat segments (see Appendix A.4 for the details). The impact of this restriction on the distribution is minor.

adjusted household income is below a certain percentage of the median in the Australian population.

To adjust for growth in real incomes over the sample period, I allow the poverty line to change in line with changes in median household incomes. I calculate the median household income in two steps. First, I calculate the median level of household disposable income in each wave using the Australian Bureau of Statistics' Income and Housing Costs surveys.[24] Second, I use the Fiscal Incidence Study in 2009–10 to calculate the ratio between the median level of adjusted household income and the median level of household disposable income. I find a ratio of 1.164 — that is, the median adjusted household income is 16.4% higher than the median household disposable income. I assume that this ratio is constant over the sample period to estimate the median level of adjusted household income in each wave.[25]

To estimate the effects on relative poverty rates, I estimate similar regressions to the distribution regressions described in Section 6.2. Here, though, the dependent variables are indicators for adjusted household income being less than a given percentage, say x%, of the median in the Australian population. As the choice of x is somewhat arbitrary, I estimate the effects for several values of x ranging from 20 to 70 (at intervals of 10). Figure 7 presents the estimates (with 95% confidence intervals), showing two patterns of note. First, regardless of which poverty line is used, all of the estimates are positive, implying an increase in relative poverty, though the estimates are only statistically significant at the 5% level at the 60% and 70% poverty lines. Second, although the size and statistical significance of the estimates depends on the choice of poverty line, with larger and more significant estimates at higher poverty lines, all of the point estimates imply similar increases in relative poverty rates in percentage terms (20–42%).

Figure 7 highlights the estimates at two common choices of the poverty line: 50% and 60% of the median, the poverty lines used by the OECD and the EU. At the 50% poverty line, the estimated increase in poverty is 2.1 percentage points (33%, $p = 0.119$). At the 60% poverty line, the estimated increase in poverty is larger — 4.4 percentage points (39%) — and statistically significant at the 5% level ($p = 0.010$). Therefore, under conventional definitions of relative poverty, the estimated increase in poverty among treated women is 2.1–4.4 percentage points (33–39%).

These effects are concentrated among single women and renters (see Figure A9). The estimates are considerably larger for these groups at every poverty line, and the estimates are close to zero and highly statistically insignificant for partnered women and homeowners. Concerningly, there is also evidence of an increase in the proportion of renters with household incomes below 20% of the median income ($p < 0.01$) and 30% of the median income ($p < 0.05$). These effects likely result from the much higher housing costs of renters and suggest that the reform pushed some renters deeper into poverty.

[24] As these surveys are conducted every second year, I use linear interpolation to estimate the median in intermediate years.

[25] The ratios are similar in the 2003–04 and 2015–16 surveys (1.139 and 1.157 respectively).

6.4 Sensitivity of the estimates to assumptions about in-kind benefit values

In this section, I examine the sensitivity of the estimates to different assumptions about the value of in-kind benefits, as the true value of in-kind benefits may differ from the values estimated in the Fiscal Incidence Study. On the one hand, the Fiscal Incidence Study assumes that there is no moral hazard — that is, that expenditure is not higher on goods that are subsidized or provided in kind. As moral hazard effects can be significant (e.g., see Aron-Dine et al., 2015; Einav, Finkelstein and Schrimpf, 2015; Finkelstein, Hendren and Luttmer, 2019; Lieber and Lockwood, 2019), the Fiscal Incidence Study may overestimate the value of in-kind benefits. On the other hand, the Fiscal Incidence Study ignores any value provided by in-kind benefits as a form of insurance against health shocks, which may be large for elderly, risk-averse households.

In Table A9, I present the estimated effects on poverty and inequality measures for different values of in-kind benefits. Column 2 presents the baseline estimates, where in-kind benefits are valued in the same way as the Fiscal Incidence Study (i.e., $1 spent by governments on in-kind benefits is worth $1 to households). Column 1 presents the estimates when in-kind benefits are valued *higher*, at 125% of expenditure, while columns 3–6 present the estimates when in-kind benefits are valued *lower*, at 75%, 50%, 25% and 0% of expenditure. Unsurprisingly, the estimates vary across columns, but the estimates are consistently positive, implying an increase in poverty and inequality measures, and even the most conservative estimates imply material effects. The estimated increase in inequality measures ranges from 4.8% to 13.0% for the Gini coefficient and 14.3% to 21.0% for the 80:20 ratio, and the estimated increase in relative poverty rates ranges from 1.8 to 4.7 percentage points.

7 Fiscal effects

The reform aimed to improve the fiscal sustainability of Australia's retirement system. In this section, I quantify its net fiscal impact. To do so, I estimate the average net fiscal impact on treated women and then multiply this estimate by the size of the affected population. Using equation (1), I estimate the average effect of the reform on the following: (i) women's income from the Age Pension; (ii) women's income from other transfers; (iii) the amount of in-kind benefits provided to households (using the information on in-kind benefits from the Fiscal Incidence Study); and (iv) the total amount of income tax paid by women.[26]

Table 4 presents the estimates. The estimates in columns 1 and 2 indicate that, on average, treated women lost $6,957 per annum in income from the Age Pension but gained $4,365 in income from other transfers (both $p < 0.01$). Column 3 presents the estimated effect of the reform on the amount of in-kind benefits provided to households; the estimate indicates that treated households received $3,341 less on average in in-kind benefits per annum ($p < 0.01$). It may seem surprising that this estimate is larger than the

[26] The information on in-kind benefits and income tax corresponds to the financial year prior to the survey date rather than being current as of the survey date. As such, I make the same modifications to equation (1) as described in Section 6.2.

estimated net income loss from transfers among women, but a relatively large proportion of women were delayed from accessing the in-kind benefits associated with concession card ownership. Table A8 shows that the reform decreased concession card ownership among treated women by an estimated 44.6 percentage points ($p < 0.01$), with over half of this decrease resulting from lower ownership of the Commonwealth Seniors Health Card, a concession card for pension-age Australians who do not receive the Age Pension.

In column 4, I present the estimated effect of the reform on income tax receipts. The estimate indicates that the reform increased the average amount of income tax paid by treated women by $1,001 per annum ($p < 0.01$).[27] This estimate may seem surprisingly high given the modest effects on women's earnings. However, as there are tax concessions for people above the APA, women would have paid more tax even without any labor supply response.[28]

Finally, in column 5, I calculate the average net fiscal impact on treated women. This estimate, of -$6,934 per annum, is equal to the estimated change in Age Pension income *plus* the estimated change in income from other transfers *plus* the estimated change in in-kind benefits *less* the estimated change in income tax (columns $1+2+3-4$). To estimate the aggregate impact of the reform, I multiply this estimate by the size of the affected population using annual population counts by single year of age and gender from the Australian Bureau of Statistics. During my sample period, the average number of treated women at each age is approximately 115,000, indicating that each one-year increase in women's APA has resulted in an average net fiscal saving of approximately 800 million dollars per annum. This saving is equivalent to 1.8% of current government expenditure on cash transfers to Age Pensioners (or 0.2% of total expenditure).

8 Conclusion

I examine the effects of a 1994 Australian reform that increased women's eligibility age for the public retirement pension from 60 to 65. Overall, I find that the largest effect was women extending their receipt of other government transfers to later ages; on average, women offset 63% of their lost income from the retirement pension through other transfers. In contrast, the estimated effect on female labor supply is relatively modest; on average, women offset 21% of their lost income from the retirement pension through an increase in earnings. Consistent with the predictions from a basic labor supply model, I find that these responses are concentrated among poorer households, as are the negative effects on household incomes. These unequal impacts meant that the reform was significantly regressive; my baseline estimates indicate that the reform increased relative poverty rates by 2.1–4.4 percentage points (33–39%) and inequality measures by 12–15%.

[27] I exclude women who were ever in the top 1% of taxpayers in the sample to reduce the sensitivity of the estimates to very high levels of tax. The estimated increase in tax is comparable without this restriction ($663 per annum, $p < 0.1$).

[28] The main tax concession for pension-age Australians is the Seniors and Pensioners Tax Offset (SAPTO). The SAPTO increases the tax-free threshold for people who have reached the APA and satisfy an income test (based on household income). In 2016, the maximum tax offset for a couple was $3,204, approximately 10% of the maximum Age Pension payment.

However, the reform significantly reduced government expenditure, with each one-year increase in women's pension age saving around 800 million dollars per annum. These findings highlight a clear trade-off for policymakers: while raising eligibility ages reduces government expenditure, these savings can result in higher poverty and inequality among older households.

The results also highlight the importance of other government transfers in attenuating these impacts. Back-of-the-envelope calculations indicate that the fiscal savings would have been around twice as large (1.76 billion dollars) if women were unable to claim other transfers prior to the pension age. However, as low-income households disproportionately relied on such transfers, the estimated increase in inequality measures would have been around four-times larger (42–88%), and the estimated increase in relative poverty rates around six-times larger (20.4–24.1 percentage points). While these estimates ignore behavioral responses, they suggest that governments should exhibit caution when considering simultaneous policies that tighten eligibility for other government transfers. Such responses may be a tempting way for governments to maximize the fiscal savings from higher pension-eligibility ages, but these policies are likely to magnify the effects on poverty and inequality.

The results in this paper have clear international relevance. As the population ages, policymakers are raising pension-eligibility ages in many countries, including the U.K., Germany, France, Italy and Spain. When considering the broader relevance of my results, it is important to remember that Australia's retirement pension is means-tested. As such, it is not surprising that the reform had a regressive impact. Similar reforms may be less regressive in other countries, especially where retirement income is based on contributions from employment. However, many contributory pensions still provide a significant form of redistribution. For example, the formula that links past earnings to Social Security income in the U.S. is highly progressive — that is, the replacement rate is higher for people with lower earnings. Hence, similar reforms may be significantly regressive in other countries as well. Moreover, as discussed above, most of the regressive impact of the Australian reform is averted by substitution to other government transfers. While such spillovers have been found in other countries, these effects are especially large in the Australian context. Thus, the estimated effects on poverty and inequality measures are not necessarily an upper bound for the effects in countries with less targeted retirement pensions, especially countries where other aspects of the social safety net are relatively weak.

Overall, this paper sheds light on the fiscal and distributional effects of raising the pension-eligibility age, a widespread response to population aging. The results show that although these reforms are fiscally effective, they can significantly diminish the social safety net for older households, especially older women who are single and renting. This highlights a potential role for complementary policies like targeted housing assistance that strengthen the social safety net for vulnerable elderly households.

References

Aron-Dine, Aviva, Liran Einav, Amy Finkelstein, and Mark Cullen. 2015. "Moral hazard in health insurance: Do dynamic incentives matter?" *Review of Economics and Statistics*, 97(4): 725–741.

Atalay, Kadir, and Garry F. Barrett. 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, Kadir, Garry F. Barrett, and Peter Siminski. 2018. "Pension incentives and the joint retirement of couples: Evidence from two natural experiments." *Journal of Population Economics*, 1–33.

Australian Bureau of Statistics. 2018*a*. "Average weekly earnings, Australia, May 2018." https://bit.ly/2ApdmiQ (accessed on October 23, 2018).

Australian Bureau of Statistics. 2018*b*. "Household income and income distribution, Australia, 2015-16." https://bit.ly/2yXnEEQ (accessed on October 23, 2018).

Behaghel, Luc, and David M. Blau. 2012. "Framing social security reform: Behavioral responses to changes in the full retirement age." *American Economic Journal: Economic Policy*, 4(4): 41–67.

Bitler, Marianne P., Jonah B. Gelbach, and Hilary W. Hoynes. 2006. "What mean impacts miss: Distributional effects of welfare reform experiments." *American Economic Review*, 96(4): 988–1012.

Blau, David M., and Ryan M. Goodstein. 2010. "Can social security explain trends in labor force participation of older men in the United States?" *Journal of Human Resources*, 45(2): 328–363.

Börsch-Supan, Axel, and Reinhold Schnabel. 1998. "Social security and declining labor-force participation in Germany." *American Economic Review*, 88(2): 173–178.

Bound, John, Arline Geronimus, Javier Rodriguez, and Timothy Waidmann. 2014. "The implications of differential trends in mortality for social security policy."

Bradshaw, Jonathan, and Emese Mayhew. 2011. "The measurement of extreme poverty in the European Union." European Commission, https://bit.ly/2SdvR0r (accessed on January 24, 2018).

Burtless, Gary. 1986. "Social security, unanticipated benefit increases, and the timing of retirement." *Review of Economic Studies*, 53(5): 781–805.

Cameron, A. Colin, and Douglas L. Miller. 2015. "A practitioner's guide to cluster-robust inference." *Journal of Human Resources*, 50(2): 317–372.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2013. "Inference on counterfactual distributions." *Econometrica*, 81(6): 2205–2268.

Clare, Ross. 2015. "Superannuation account balances by age and gender." The Association of Superannuation Funds in Australia (ASFA), https://www.superannuation.asn.au/policy/reports (accessed on October 23, 2018).

Coile, Courtney. 2004. "Retirement incentives and couples' retirement decisions." *Topics in Economic Analysis & Policy*, 4(1).

Coile, Courtney, and Jonathan Gruber. 2007. "Future social security entitlements and the retirement decision." *Review of Economics and Statistics*, 89(2): 234–246.

Cribb, Jonathan, and Carl Emmerson. 2019. "Can't wait to get my pension: the effect of raising the female early retirement age on income, poverty and deprivation." *Journal of Pension Economics & Finance*, 18(3): 450–472.

Cribb, Jonathan, Carl Emmerson, and Gemma Tetlow. 2016. "Signals matter? Large retirement responses to limited financial incentives." *Labour Economics*, 42: 203–212.

Cristia, Julian P. 2009. "Rising mortality and life expectancy differentials by lifetime earnings in the United States." *Journal of Health Economics*, 28(5): 984–995.

Department of Social Services. 2016. "DSS demographics December 2016." https://bit.ly/2ODfP15 (accessed on October 7, 2018).

Dube, Arindrajit. 2019. "Minimum wages and the distribution of family incomes." *American Economic Journal: Applied Economics*, 11(4): 268–304.

Duggan, Mark, Perry Singleton, and Jae Song. 2007. "Aching to retire? The rise in the full retirement age and its impact on the social security disability rolls." *Journal of Public Economics*, 91(7): 1327–1350.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf. 2015. "The response of drug expenditure to nonlinear contract design: Evidence from Medicare Part D." *Quarterly Journal of Economics*, 130(2): 841–899.

Engelhardt, Gary V, and Jonathan Gruber. 2004. "Social security and the evolution of elderly poverty."

Finkelstein, Amy, Nathaniel Hendren, and Erzo F.P. Luttmer. 2019. "The value of Medicaid: Interpreting results from the Oregon health insurance experiment." *Journal of Political Economy*, 127(6): 2836–2874.

Foresi, Silverio, and Franco Peracchi. 1995. "The conditional distribution of excess returns: An empirical analysis." *Journal of the American Statistical Association*, 90(430): 451–466.

Fox, Liana. 2019. "The supplemental poverty measure: 2016." Current population report P60-268. Washington DC: U.S. Census Bureau.

Fuchs, Victor. 1969. "Comment on measuring the size of the low-income population." *Six papers on the size distribution of wealth and income*, 198–202. New York: National Bureau of Economic Research.

Gelber, Alexander M, Adam Isen, and Jae Song. 2016. "The effect of pension income on elderly earnings: Evidence from social security and full population data."

Geyer, Johannes, and Clara Welteke. forthcoming. "Closing routes to retirement for women: How do they respond?" *Journal of Human Resources*.

Geyer, Johannes, Peter Haan, Anna Hammerschmid, and Michael Peters. 2018. "Labor market and distributional effects of an increase in the retirement age."

Gruber, Jonathan, and David A Wise. 2004. *Social security programs and retirement around the world: Micro estimation.* University of Chicago Press.

Gustman, Alan L., and Thomas L. Steinmeier. 2000. "Retirement in dual-career families: A structural model." *Journal of Labor Economics*, 18(3): 503–545.

Gustman, Alan L., and Thomas L. Steinmeier. 2004. "Social security, pensions and retirement behaviour within the family." *Journal of Applied Econometrics*, 19(6): 723–737.

Hanel, Barbara, and Regina T. Riphahn. 2012. "The timing of retirement—New evidence from Swiss female workers." *Labour Economics*, 19(5): 718–728.

Harmer, Jeff. 2008. "Pension review background paper." Department of Families, Housing, Community Services and Indigenous Affairs, https://bit.ly/2yZLi3L (accessed on October 23, 2018).

Krueger, Alan B., and Jörn-Steffen Pischke. 1992. "The effect of social security on labor supply: A cohort analysis of the notch generation." *Journal of Labor Economics*, 10(4): 412–437.

Lalive, Rafael, and Pierpaolo Parrotta. 2017. "How does pension eligibility affect labor supply in couples?" *Labour Economics*, 46: 177–188.

Lalive, Rafael, Arvind Magesan, and Stefan Staubli. 2017. "Raising the full retirement age: Defaults vs incentives."

Lieber, Ethan M.J., and Lee M. Lockwood. 2019. "Targeting with in-kind transfers: Evidence from Medicaid home care." *American Economic Review*, 109(4): 1461–1485.

Liebman, Jeffrey B., Erzo F.P. Luttmer, and David G. Seif. 2009. "Labor supply responses to marginal Social Security benefits: Evidence from discontinuities." *Journal of Public Economics*, 93(11): 1208–1223.

Lundberg, Shelly. 1985. "The added worker effect." *Journal of Labor Economics*, 3(1, Part 1): 11–37.

Manoli, Dayanand S, and Andrea Weber. 2016. "The effects of the Early Retirement Age on retirement decisions."

Mastrobuoni, Giovanni. 2009. "Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities." *Journal of Public Economics*, 93(11): 1224–1233.

Meyer, Bruce D, and Derek Wu. 2018. "The poverty reduction of social security and means-tested transfers." *ILR Review*, 71(5): 1106–1153.

Morris, Todd. 2019. "Large response to delayed eligibility or a pre-existing trend in female participation? Re-examining an Australian pension reform."

OECD. 2015. *In it together: Why less inequality benefits all.* OECD Publishing, Paris.

Oguzoglu, Umut, Cain Polidano, and Ha Vu. 2016. "Impacts from delaying access to retirement benefits on welfare receipt and expenditure: Evidence from a natural experiment."

Panis, Constantijn, Michael Hurd, David Loughran, Julie Zissimopoulos, Steven Haider, and Patricia St Clair. 2002. "The effects of changing social security administration's early entitlement age and the normal retirement age."

Rothe, Christoph. 2012. "Partial distributional policy effects." *Econometrica*, 80(5): 2269–2301.

Rust, John, and Christopher Phelan. 1997. "How Social Security and Medicare affect retirement behavior in a world of incomplete markets." *Econometrica*, 781–831.

Scholz, John Karl, Robert Moffitt, and Benjamin Cowan. 2009. "Trends in income support." *Changing poverty, changing policies*, 203–241.

Selin, Hakan. 2017. "What happens to the husband's retirement decision when the wife's retirement incentives change?" *International Tax and Public Finance*.

Smith, Warwick, and David Hetherington. 2018. "Not so super, for women: Superannuation and women's retirement outcomes." Per Capita, https://bit.ly/2EKex0n (accessed on October 23, 2018).

Snyder, Stephen E, and William N Evans. 2006. "The Effect of Income on Mortality: Evidence from the Social Security Notch." *Review of Economics and Statistics*, 88(3): 482–495.

Song, Jae, and Joyce Manchester. 2007. "Have people delayed claiming retirement benefits-Responses to changes in social security rules." *Social Security Bulletin*, 67(2): 1–23.

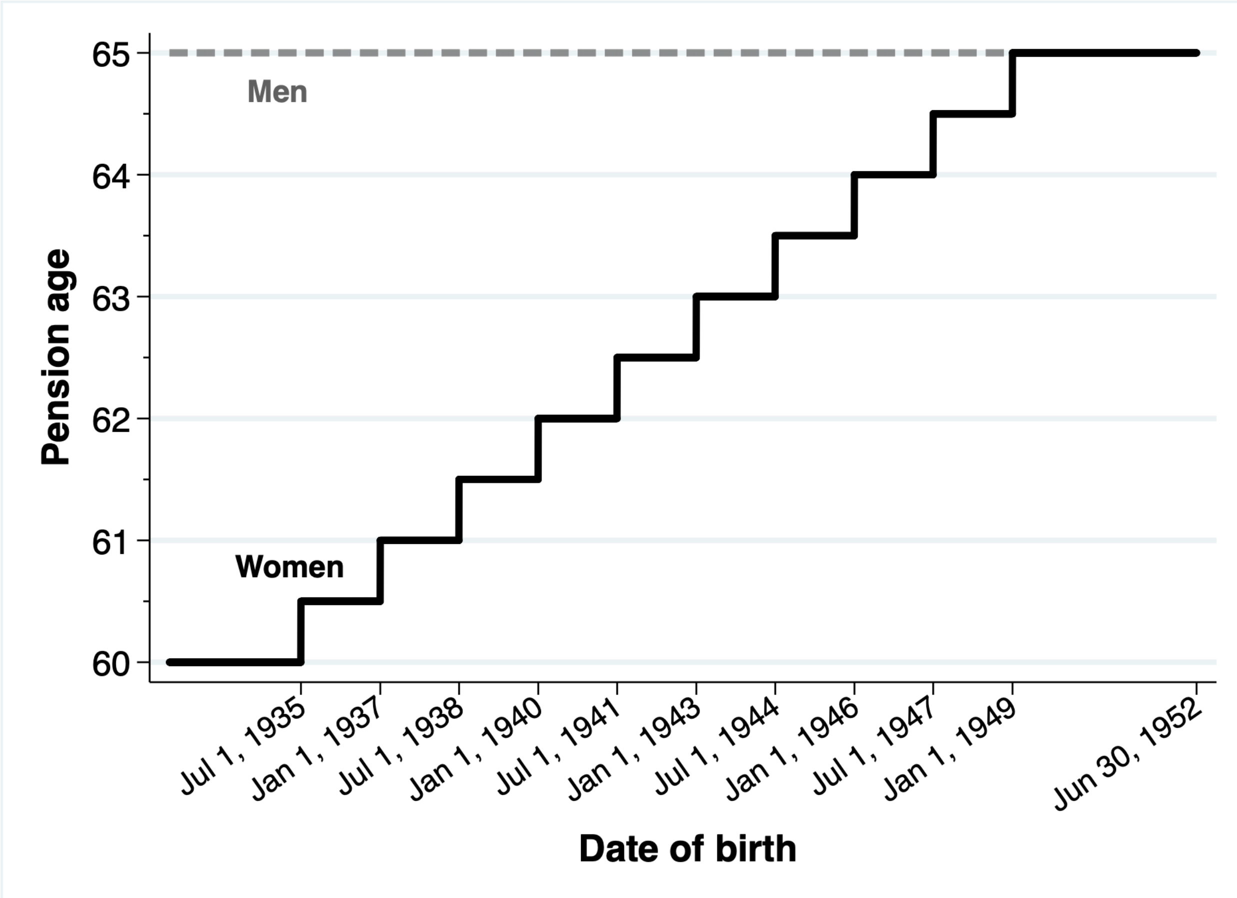
Staubli, Stefan, and Josef Zweimüller. 2013. "Does raising the early retirement age increase employment of older workers?" *Journal of Public Economics*, 108: 17–32.

Vestad, Ola Lotherington. 2013. "Labour supply effects of early retirement provision." *Labour Economics*, 25: 98–109.

Waldron, Hilary. 2007. "Trends in mortality differentials and life expectancy for male social security-covered workers, by socioeconomic status." *Social Security Bulletin*, 67(3): 1–28.

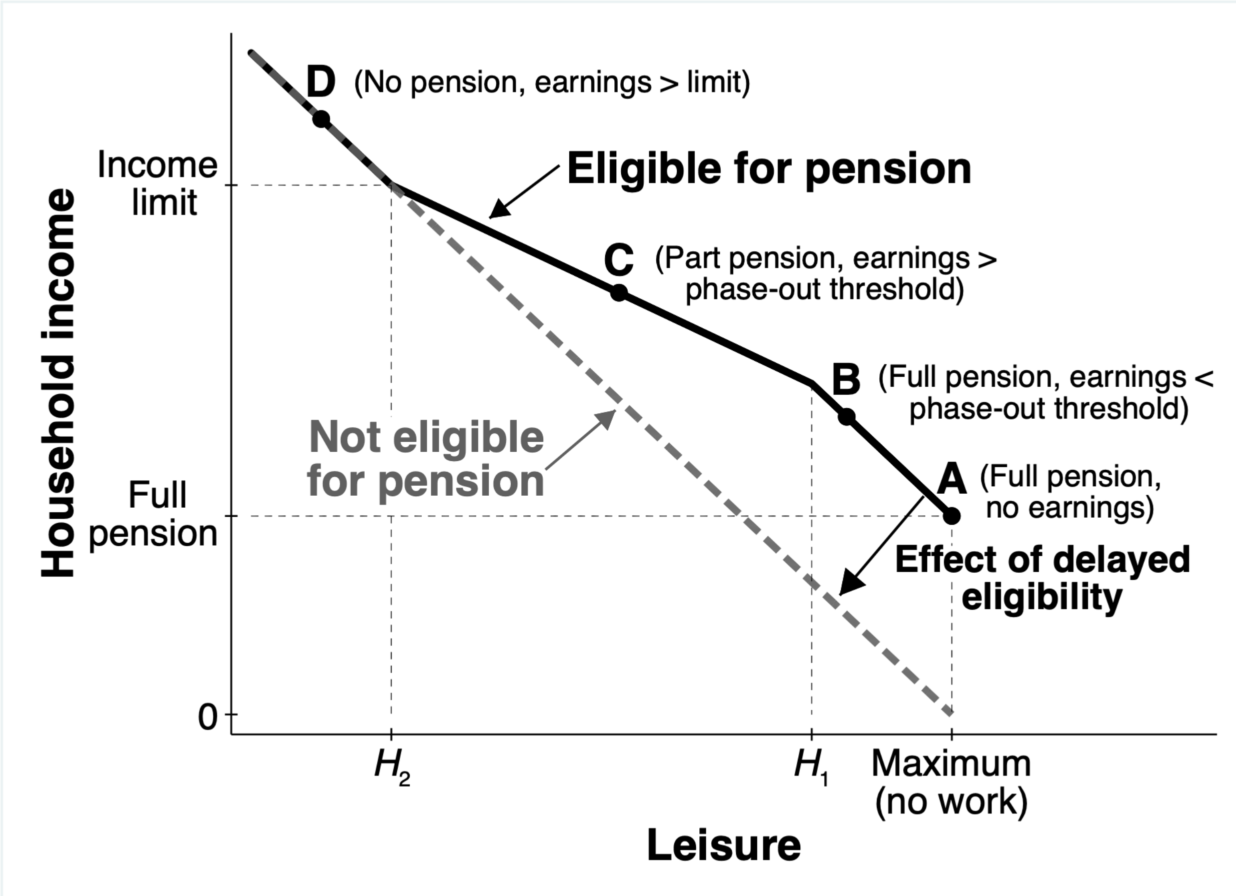
Watson, Nicole, and Mark P. Wooden. 2012. "The HILDA Survey: A case study in the design and development of a successful household panel survey." *Longitudinal and Life Course Studies*, 3(3): 369–381.

Figure 1: Pension age of Australians by date of birth and gender



Notes: This figure shows the Age Pension Age (APA) of Australian men and women based on their date of birth. The phased increases in women's APA from 60 to 65 were due to the 1994 reform. See the text in Section 2.2 for more details on the reform.

Figure 2: Households' budget constraint and the expected effects on earnings and total income



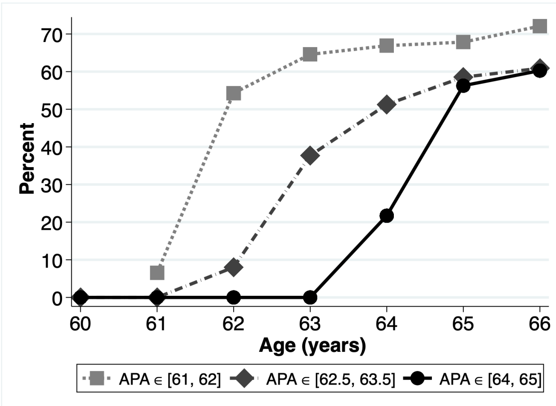
Summary of effects on each type of household

Location when eligible (1)	Effect of delay in eligibility on: Household income (2)	Effective wage/hour (3)	Expected effect on: Earnings (4)	Total income (5)
A	negative	zero	zero or positive	negative
B	negative	zero	positive	negative
C	negative	positive	positive	uncertain
D	zero	zero	zero	zero

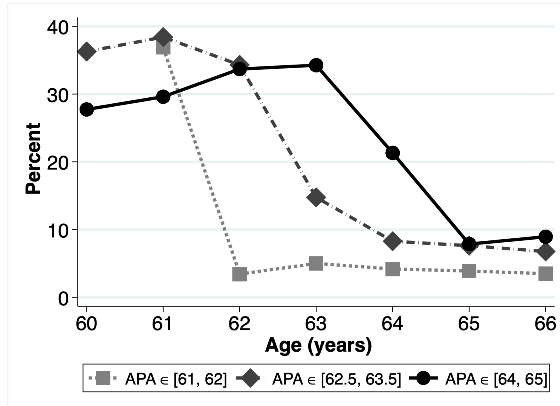
Notes: This figure shows the effect of delaying women from reaching the pension age on their households' budget constraint, in terms of the set of feasible combinations of total income and leisure (assuming that households are eligible for the maximum payment under the assets test and ignoring other government transfers). We can think about this effect as a downward shift in the household budget constraint from the black line to the gray line. The size of the shift varies with household income because of the income test for the Age Pension. The table above summarizes the effects on the four different types of household in the counterfactual scenario (points **A** to **D**), where women face no delay in reaching the pension age. Columns 2 and 3 summarize the direct effects of the delay in eligibility on the income and effective wage rate of households, and columns 4 and 5 show the expected effects on earnings and total income. See the text in Section 2.2 for more details.

Figure 3: Mean outcomes of women by age and pension age

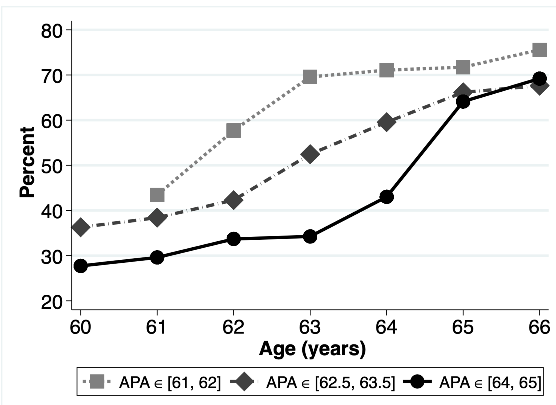
(a) Receives Age Pension



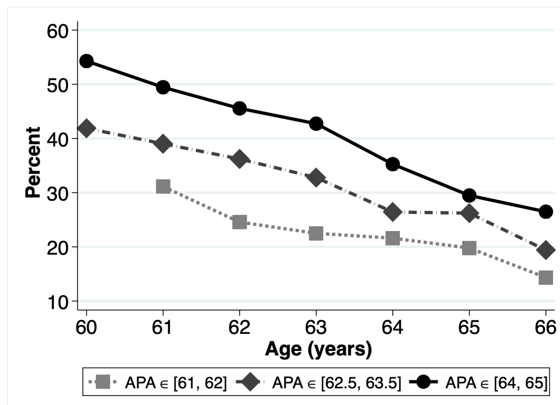
(b) Receives other government transfer



(c) Receives any government transfer



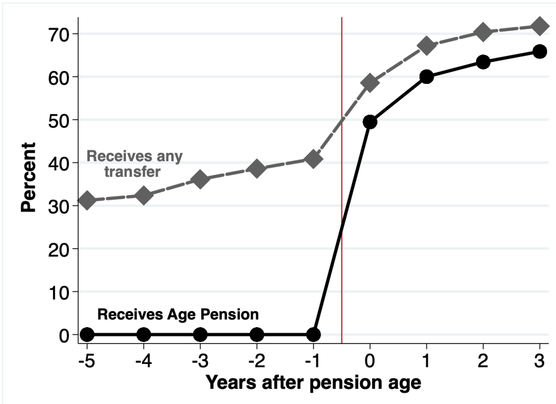
(d) In labor force



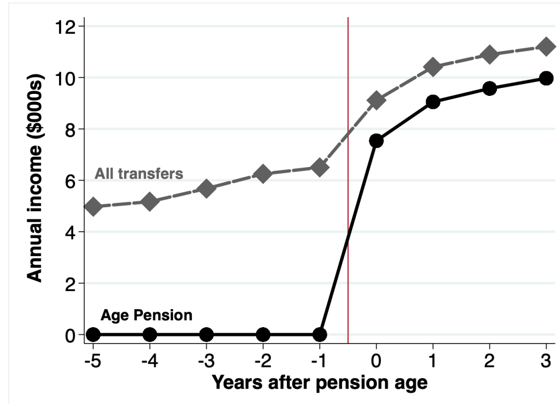
Notes: These figures present the mean outcomes of women by their age in years and Age Pension Age (APA). For clarity, I remove data points that are based on a relatively small number of observations (less than 50% of the average for the relevant group). The sample includes women aged 60–66 from waves 1–14 of the HILDA survey.

Figure 4: Mean outcomes of women by years from pension age

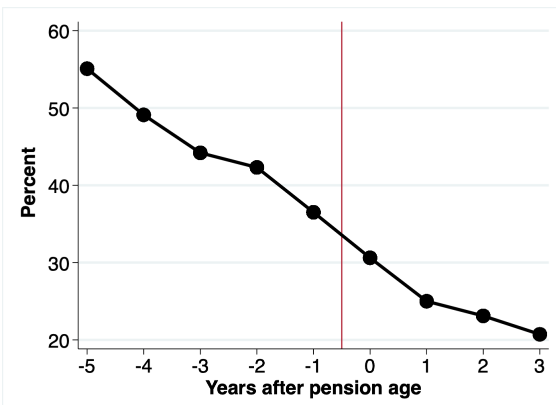
(a) Receipt of government transfers



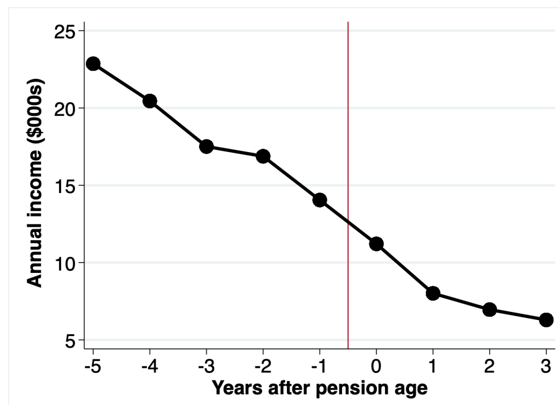
(b) Income from government transfers



(c) In labor force



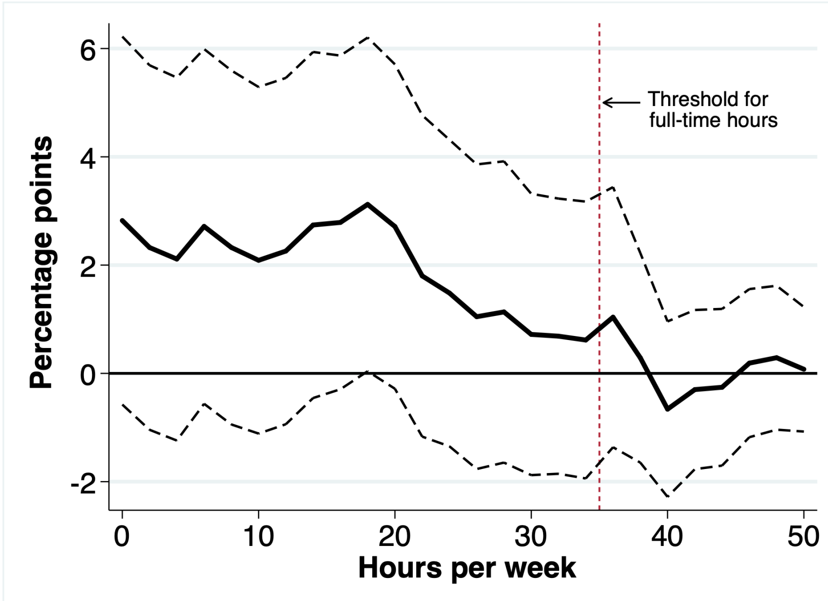
(d) Labor earnings



Notes: These figures present the mean outcomes of women with respect to the number of years between their age and their Age Pension Age (APA), i.e. $\lfloor age - APA \rfloor$, where $\lfloor\rfloor$ is the floor function. Incomes are in 2016 Australian dollars. These figures use an expanded sample containing all women who are between -5 and 3 years above the APA from waves 1–14 of the HILDA survey (rather than just those aged 60–66).

Figure 5: Labor supply responses across the distributions of hours worked and earnings

(a) Hours worked $> x$ per week

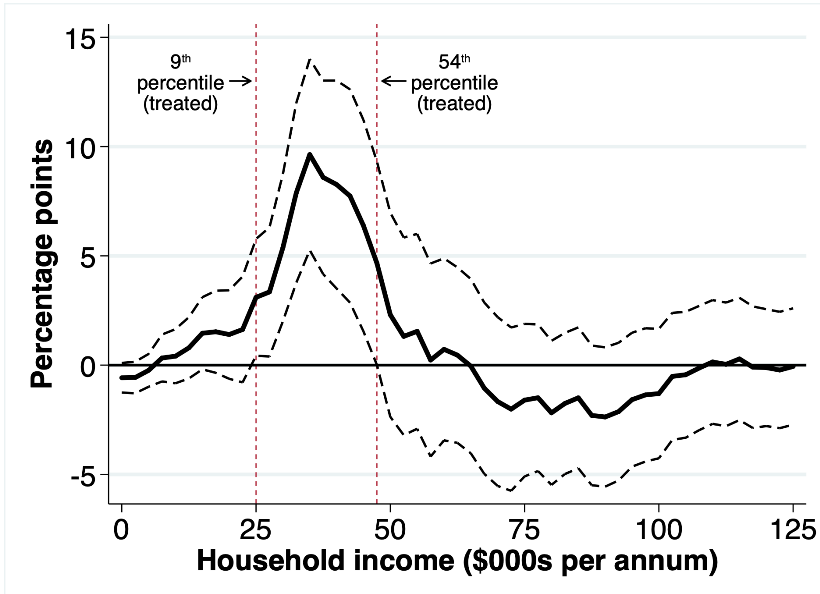


(b) Labor earnings $> \$x$ per annum

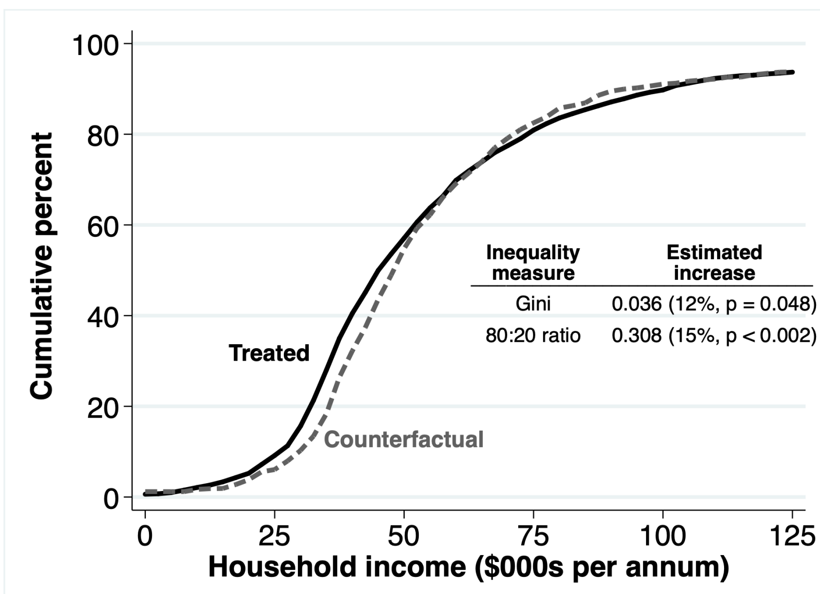
Notes: These figures present distribution regression estimates (estimated by OLS), with 95% confidence intervals, showing the effects of the reform on treated women's earnings (in 2016 Australian dollars) and hours worked at different points in the distribution. In each regression, the dependent variable is an indicator for hours/earnings being above a given level, say x, and x is spaced at two-hour intervals in (a) and $2,500 intervals in (b). The estimates are modest in size but consistent with the theoretical predictions in Figure 2; the estimates provide evidence of an increase in labor supply from part-time and low-earning women but not from full-time and high-earning women. The sample includes women aged 60–66 from waves 1–14 of the HILDA survey. 'Average full-time earnings' comes from Australian Bureau of Statistics' (2018*a*) estimates of the average total earnings of all full-time adult workers in Australia in November 2016. See the text in Section 5.2 for more details.

Figure 6: Overall distributional effects and impact on inequality measures

(a) Distribution regression estimates: Household income < $\$x$ per annum

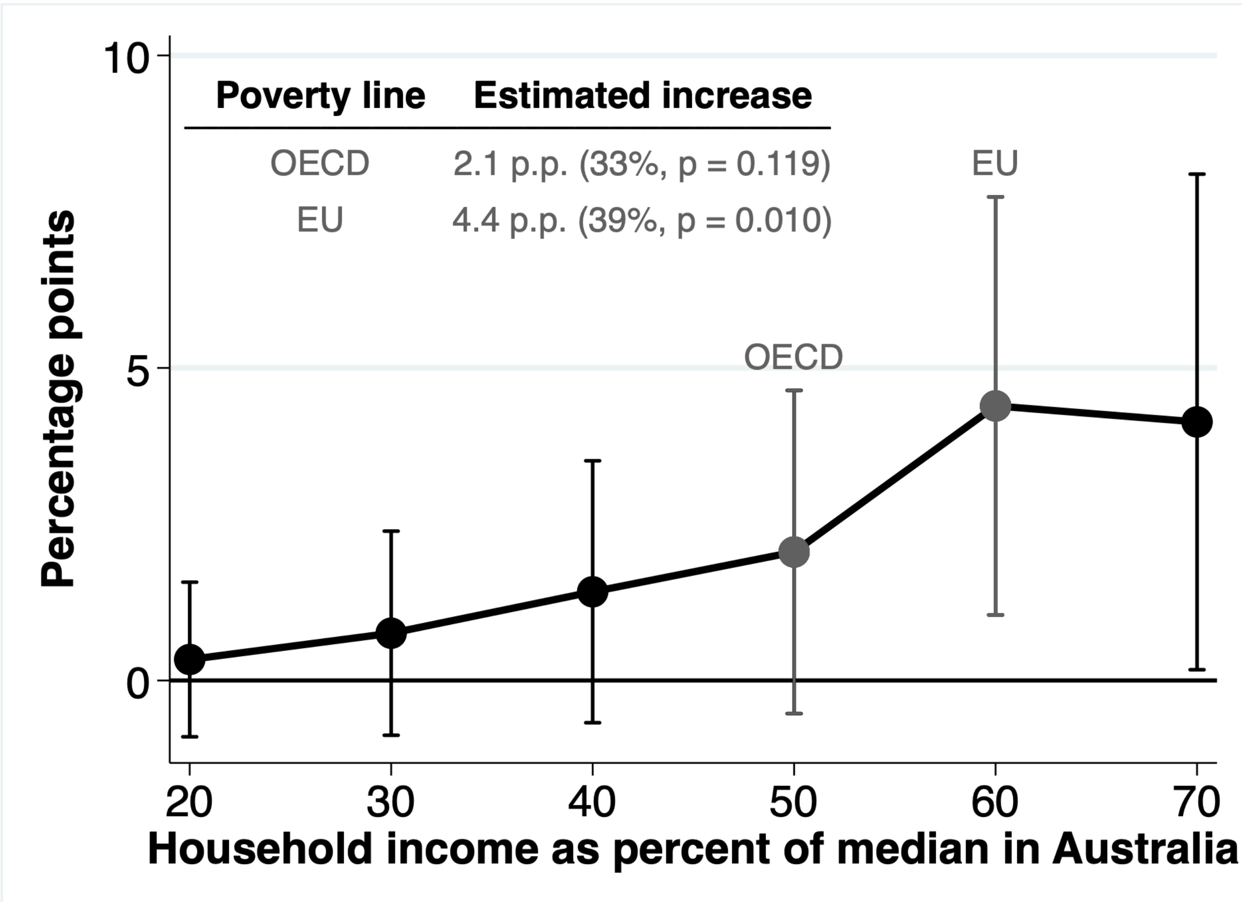


(b) Distribution of household income, estimated counterfactual and estimated effect on inequality



Notes: These figures show the distributional effects of the reform on household incomes (after adjustments for housing costs and the value of in-kind benefits). Incomes are in 2016 Australian Dollars. (a) presents distribution regression estimates (estimated by OLS), with 95% confidence intervals, showing the estimated impact of the reform on the fraction of households with income below a given level, say x, with x spaced at $2,500 intervals. The estimates are consistent with the theoretical predictions in Figure 2; the estimates indicate that the reform had negative effects on the incomes of low-to-middle-income households and little impact on higher income households. The estimates in (a) are used to estimate the counterfactual distribution of household income in (b). The estimated effect on inequality measures among treated women is equal to the difference in such measures between the treated and counterfactual distributions. The p-values shown are for a two-tailed test and are constructed using a pairs cluster bootstrap with 999 replications. The sample includes women aged 60–66 from waves 1–14 of the HILDA survey. See the text in Section 6.2 for more details.

Figure 7: Estimated effects on relative poverty: Household income $< x\%$ of the median level in the Australian population



Notes: This figure presents OLS regression estimates of the causal effect of the reform on relative poverty rates, with 95% confidence intervals, from estimates of equation (1). The dependent variable in each regression is an indicator for household income (after adjustments for housing costs and the value of in-kind benefits) being less than a given percentage, say $x\%$, of the median level in the Australian population, with x spaced at intervals of 10 from 20 to 70. Regardless of the poverty line, the estimates are positive, implying an increase in relative poverty, and the estimates are statistically significant at the 5% level for the 60% and 70% poverty lines. The p-values shown are for a two-tailed test. The sample includes women aged 60–66 from waves 1–14 of the HILDA survey. See the text in Section 6.3 for more details.

Table 1: Characteristics of the sample

	All waves	Wave 1 (2001–02)	Wave 14 (2014–15)
Demographics			
Age (years)	63.4	63.4	63.5
(Std. dev.)	(2.0)	(2.1)	(2.0)
Partnered	66.3%	72.4%	65.2%
Home owner	84.2%	87.3%	84.1%
Completed high school	25.5%	18.8%	30.6%
Completed year 10	74.7%	65.6%	80.6%
Government transfers			
Receiving Age Pension	27.6%	46.8%	15.5%
Average payments if receiving Age Pension	14,378	14,552	15,310
(Std. dev.)	(5,115)	(4,055)	(6,897)
Receiving any transfer	48.2%	60.3%	40.5%
Average payments if receiving any transfer	14,831	15,006	15,969
(Std. dev.)	(6,049)	(5,467)	(7,397)
Labor market			
In labor force	35.5%	23.0%	41.4%
Employed	34.5%	22.6%	39.8%
Earnings if employed	45,102	37,676	45,916
(Std. dev.)	(27,660)	(25,299)	(27,617)
Spousal variables (if partnered)			
Age (years)	65.7	65.9	66.0
(Std. dev.)	(4.9)	(4.8)	(4.7)
Receiving Age Pension	31.6%	35.9%	28.8%
Average payments if receiving Age Pension	12,769	13,384	13,023
(Std. dev.)	(4,867)	(3,324)	(5,220)
Receiving any transfer	45.6%	52.2%	40.7%
Average payments if receiving any transfer	14,742	14,122	15,572
(Std. dev.)	(8,863)	(6,366)	(9,488)
In labor force	42.6%	33.5%	46.0%
Employed	41.6%	32.7%	44.2%
Earnings if employed	63,614	55,817	68,580
(Std. dev.)	(45,877)	(35,088)	(51,851)
Number of observations	7,999	474	761
Number of individuals	1,945	474	761

Notes: This table summarizes the characteristics of the sample and presents the means of the key variables. Incomes are in 2016 Australian dollars. The sample consists of women aged 60–66 years old from waves 1–14 of the HILDA survey. See the text in Section 3.1 for more details on the sample.

Table 2: Estimated effects on women's receipt of government transfers and labor supply

	Transfer receipt			Labor supply		
	Age Pension (1)	Other transfer (2)	Any transfer (3)	In labor force (4)	Employed (5)	Hours per week (6)
Effect of reform	-0.483***	0.300***	-0.182***	0.031*	0.027	0.63
	(0.018)	(0.019)	(0.019)	(0.018)	(0.017)	(0.59)
Mean at pension age	0.478	0.083	0.561	0.311	0.303	7.99
R-squared	0.475	0.178	0.226	0.121	0.116	0.120
Observations	7,999	7,999	7,999	7,999	7,999	7,882

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by individual.

Notes: This table presents OLS regression estimates of the causal effect of the reform, δ, from equation (1). 'Mean at pension age' is the mean of the dependent variable among individuals who have reached the pension age in the last 12 months. The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey.

Table 3: Estimated effects on women's income

	Age Pension (1)	Other transfers (2)	Labor earnings (3)	Transfers + earnings (4)
Effect of reform	-6,957***	4,365***	1,458	-1,177
	(301)	(312)	(892)	(855)
Mean at pension age	6,879	1,244	11,106	19,317
R-squared	0.436	0.181	0.110	0.089
Observations	7,999	7,999	7,913	7,913

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by individual.

Notes: This table presents OLS regression estimates of the causal effect of the reform, δ, from equation (1). 'Mean at pension age' is the mean of the dependent variable among individuals who have reached the pension age in the last 12 months. Incomes are annualized and in 2016 Australian dollars. Columns 3 and 4 exclude women who ever had earnings in the top 1% of workers. The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey.

Table 4: Estimated effects on net government expenditure per affected woman

	Age Pension (1)	Other transfers (2)	In-kind benefits (3)	Income taxes (4)	Net expenditure (5)
Effect of reform	-6,957***	4,365***	-3,341***	1,001***	-$6,934
	(301)	(312)	(264)	(277)	
Mean at pension age	6,879	1,244	23,550	1,202	
R-squared	0.436	0.181	0.663	0.092	
Observations	7,999	7,999	7,999	7,734	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by individual.

Notes: This table presents OLS regression estimates of the causal effect of the reform, δ, from equation (1). 'Mean at pension age' is the mean of the dependent variable among individuals who have reached the pension age in the last 12 months. The estimates are annualized and in 2016 Australian dollars. The dependent variable in column 3 is the imputed value of in-kind benefits to households, which is constructed using the 2003–04 and 2009-10 cross-sections of the Australian Bureau of Statistics' Fiscal Incidence Study (see Section 6.1 for more information). Column 4 excludes women who ever had income tax in the top 1% of the sample. The effect on net government expenditure in column 5 is calculated as the sum of the treatment effects in columns 1–3 minus the treatment effect in column 4. The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey.

Online Appendix for "The Unequal Burden of Retirement Reform: Evidence from Australia"

Todd Morris

Munich Center for the Economics of Aging, Max Planck Institute for Social Law and Social Policy

A Data appendix

A.1 Testing for anticipatory behavior

As my estimates calculate the effects of the reform on women's outcomes net of any anticipatory responses before age 61.5, I examine whether the reform caused any effects on women's income or labor supply at ages 55–61. To test for such effects, I assume that women with a higher Age Pension Age (APA), who had a larger reduction in their expected future pension income, would have had a stronger response if women were forward-looking. Specifically, among women in the sample with an APA of 65.0 and below,[29] I test whether women's APA affected their outcomes at ages 55–61 using the following regressions:

$$y_{it} = \beta x_{it} + \delta_{ant} APA_i + \theta birth\_year_i + FE\_age\_0.5yrs_{it} + \varepsilon_{it} \tag{A1}$$

where the notation and covariates are the same as in equation (1) except $birth\_year_i$ is each woman's year of birth. Including $birth\_year_i$ allows for a linear cohort trend in the unobservable determinants of y_{it}, which is essential in this context due to the strong trends across the relevant cohorts in terms of women's education, number of children, attitudes towards work and labor force participation that would otherwise make it impossible to distinguish anticipatory behavior from unrelated cross-cohort trends in women's outcomes. The key explanatory variable is APA_i, a woman's APA in years. A significant estimate of δ_{ant} would indicate that raising women's APA affected their outcomes at ages 55–61.

Table A3 shows that none of the estimates are statistically significant at the 10% level. Thus, there is no conclusive evidence of anticipatory behavior. However, the power of these tests is limited; the standard errors on these estimates are relatively large, around three-times larger than the standard errors in Tables 2 and 3, because of high multicollinearity between women's APA and year of birth.

[29] I exclude women with an APA of 65.5 or above for two reasons. First, the phased increases in the APA from 65.0 to 67.0 were only announced in 2009, eight years after the start of the sample period. Second, the treatment effect is identified mainly using variation in the APA among women with an APA below 65.5.

A.2 Imputing concession card ownership

This section describes the imputation of concession card ownership for women and their spouses. As I require information on concession card ownership to estimate the total value of in-kind benefits for households over the course of the previous financial year, I estimate the fraction of time respondents would have owned a concession card during this period. I consider two types of concession cards: the Pensioner Concession Card (PCC) and the Commonwealth Seniors Health Card (CSHC). To estimate the fraction of time respondents owned a PCC, I first add up the number of weeks they received a transfer that directly qualified them for a PCC and divide by 52.[30] Then, I add to this the fraction of weeks they received other transfers that would qualify them for a PCC if they were over the age of 60, multiplied by the fraction of time they were over 60.[31] Finally, I restrict the fraction of time respondents owned a PCC to be no larger than 1.

To estimate the fraction of time respondents owned a CSHC, I set this equal to the fraction of time they were above the Age Pension Age (APA) in the preceding financial year if they satisfied the following two conditions: (i) they did not own a PCC and (ii) their equivalised household disposable income was below the 90$^{\text{th}}$ percentile in the sample. Otherwise, I set this fraction equal to 0.

Finally, I estimate the fraction of time respondents owned a concession card in the previous financial year as the sum of the fraction of time they owned a PCC and the fraction of time they owned a CSHC. Overall, this imputation is effective in (i) fitting the level of concession card ownership among women in the sample and (ii) predicting which individuals owned concession cards. In waves 9 and 13, where information on concessions is available, I estimate that 44.0% of women owned a concession card in the previous financial year. This is similar to their reported level of concession card ownership of 46.9%.[32] Moreover, the imputation aligns with women's reports for 88.0% of observations.

A.3 Value of in-kind benefits for households

After imputing the fraction of time respondents owned a concession card in the previous financial year, I calculate the average number of concession cards owned by each household in the sample. This is equal to the fraction of time women owned a concession card in the previous financial year plus the corresponding fraction for their spouse (if they had one). Then, I construct the total value of in-kind benefits for each

[30] These transfers include the Age Pension, Carer Payment, Disability Support Pension, Mature Age Allowance, Parenting Payment Single, Mature Age Partner Allowance, Bereavement Allowance and Wife Pension.

[31] These transfers include Newstart Allowance, Parenting Payment Partnered, Partner Allowance, Widow Allowance, Sickness Allowance and Special Benefit.

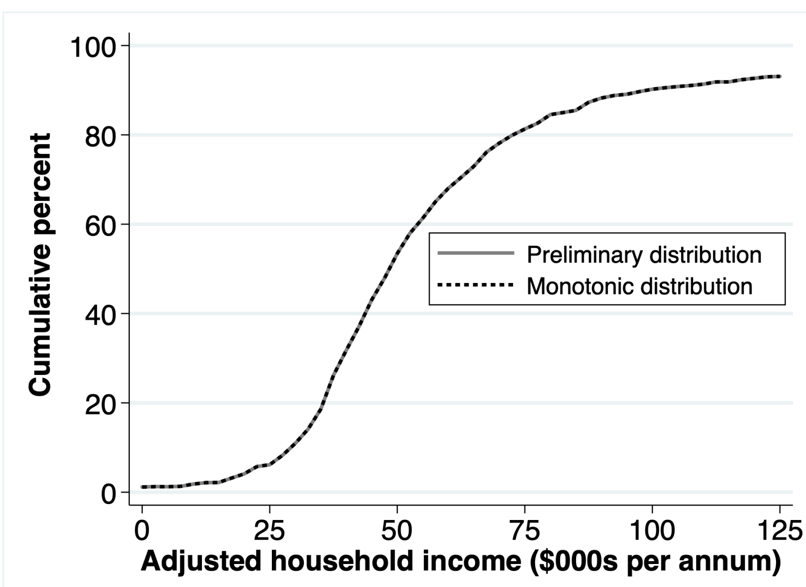
[32] This small difference makes sense because concession card ownership increases with age, and women are younger in the financial year prior to the survey date than when they report concession card ownership (at the survey date).

household in the previous financial year using the average values for the different household types shown in Table A7 (which come from the 2003–04 and 2009–10 cross-sections of the Australian Bureau of Statistics' Fiscal Incidence Study). I use linear interpolation to calculate the total value of in-kind benefits for households whose average number of concession cards is not equal to zero, one or two. For example, the total value of in-kind benefits for a single women who owned 0.5 concession cards on average in the previous financial year is calculated as $0.5 \times \$9,365 + 0.5 \times \$15,559 = \$12,462$.

A.4 Monotonicity restriction on counterfactual distribution

To construct the counterfactual distribution in Figure 6b, I start by subtracting the distribution regression estimates in Figure 6a from the treated distribution at each value of x. Then, I restrict the counterfactual distribution to be non-decreasing in income. To explain how, let me introduce some notation. Let $\mathbf{x} = (x_0, x_1, \ldots, x_{400})$ be the grid of values for household income, spaced at \$2,500 intervals, where $x_0 = \$0$, $x_1 = \$2,500$ and $x_{400} = \$1,000,000$, and let $\mathbf{y} = (y_0, y_1, \ldots, y_{400})$ be the corresponding set of estimates of the counterfactual distribution after subtracting the distribution regression estimates in Figure 6a from the treated distribution. To restrict the counterfactual distribution to be non-decreasing, I start at the top of the income distribution and replace y_{399} with the minimum of y_{399} and y_{400}. Then, I do the same for y_{398} (setting it equal to the minimum of y_{398} and y_{399}) and continue in this manner until I reach y_0. As Figure A1 shows, the impact of this algorithm on the counterfactual distribution is very minor (and visually imperceptible).[33]

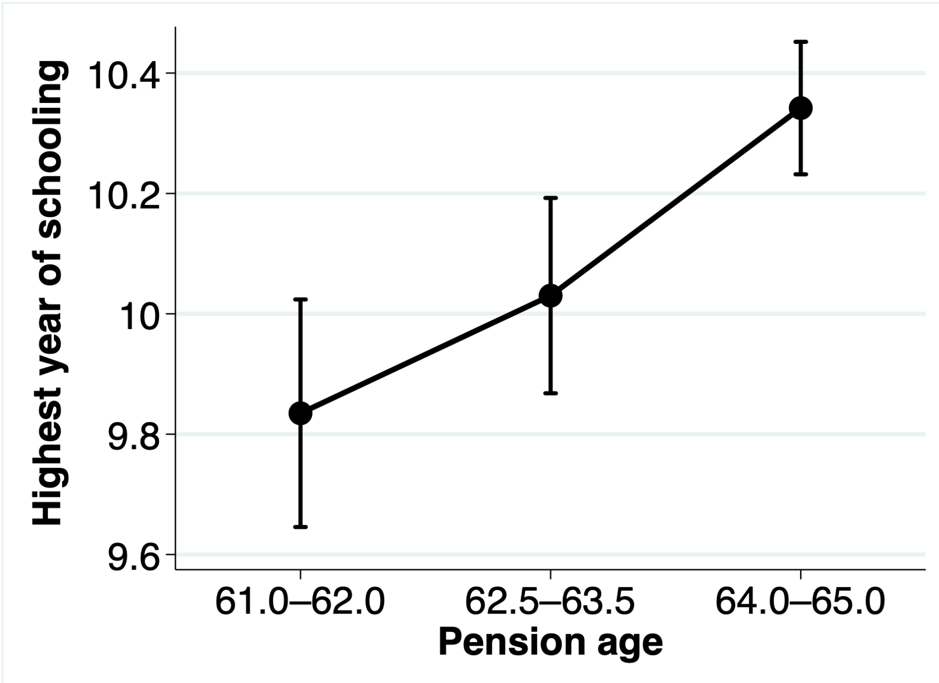
Figure A1: Impact of the monotonicity restriction on the counterfactual distribution



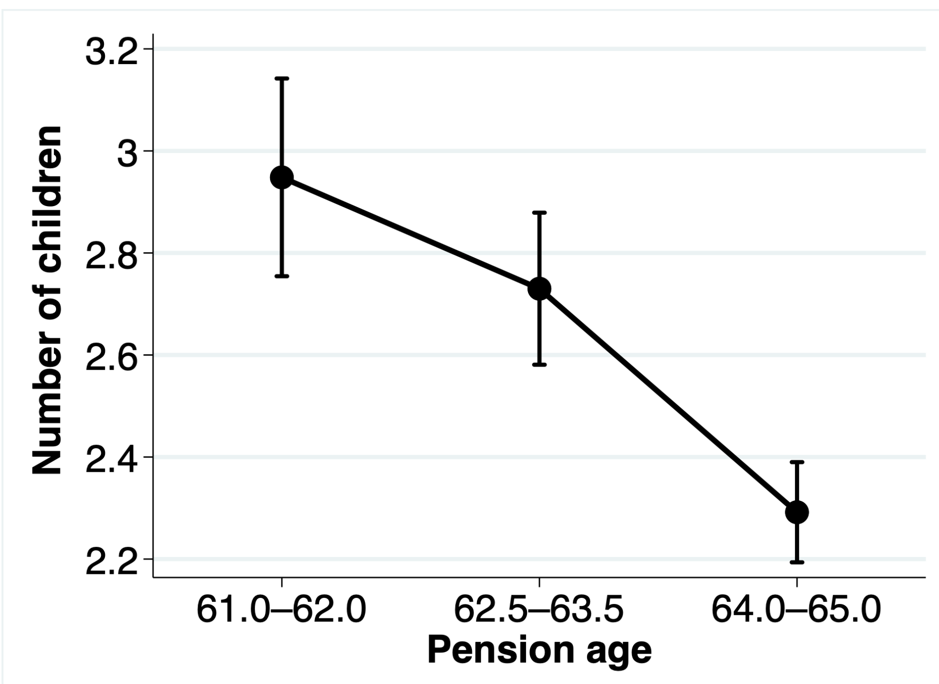
[33] An alternative approach would be to start at the bottom of the distribution and replace y_n with the maximum of y_{n-1} and y_n at each step. I use the former approach because it results in slightly smaller estimates of the impact on inequality measures.

Figure A2: Trends in the characteristics of women by their pension age

(a) Highest year of schooling completed



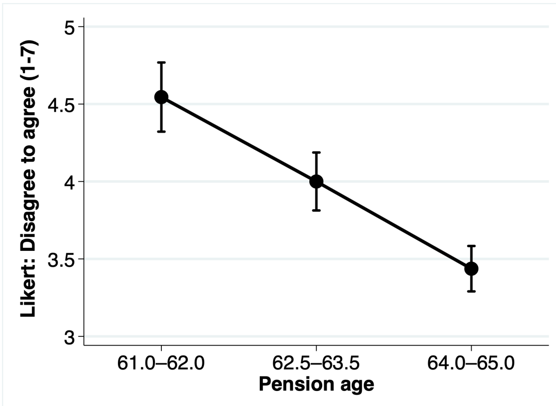
(b) Number of children



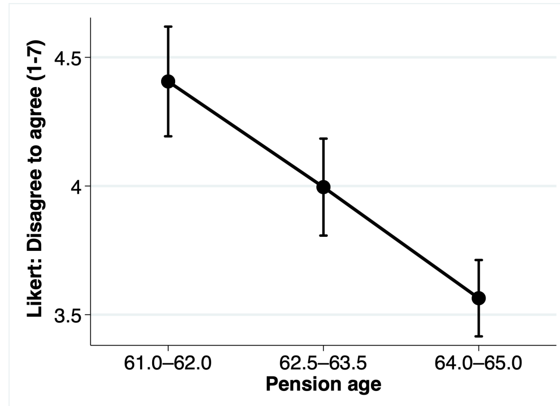
<u>Notes</u>: These figures show how average birth rates and education vary with women's Age Pension Age (APA), with 95% confidence intervals around these averages. These figures show that women with a higher APA are more educated and have fewer children on average. In (a), I exclude 1% of women who responded that they did not complete primary school. The sample consists of women aged 60–66 from waves 1–14 of the HILDA Survey.

Figure A3: Trends in the attitudes of women by their pension age

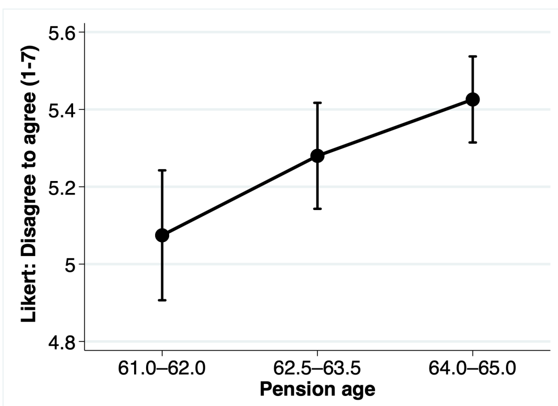
(a) ‘Better for everyone involved if the man works and the woman takes care of home and children’



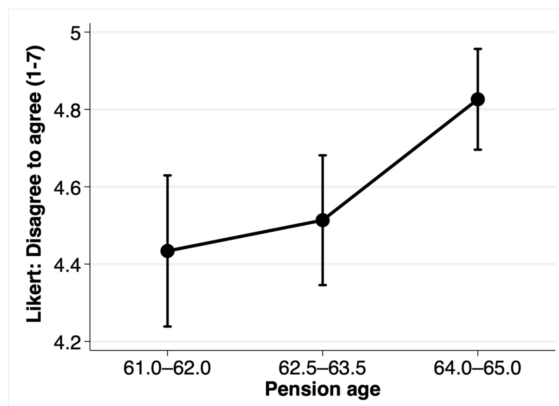
(b) ‘Mothers who don’t really need the money shouldn’t work’



(c) ‘Children do just as well if the mother works and the father cares for the home and children’

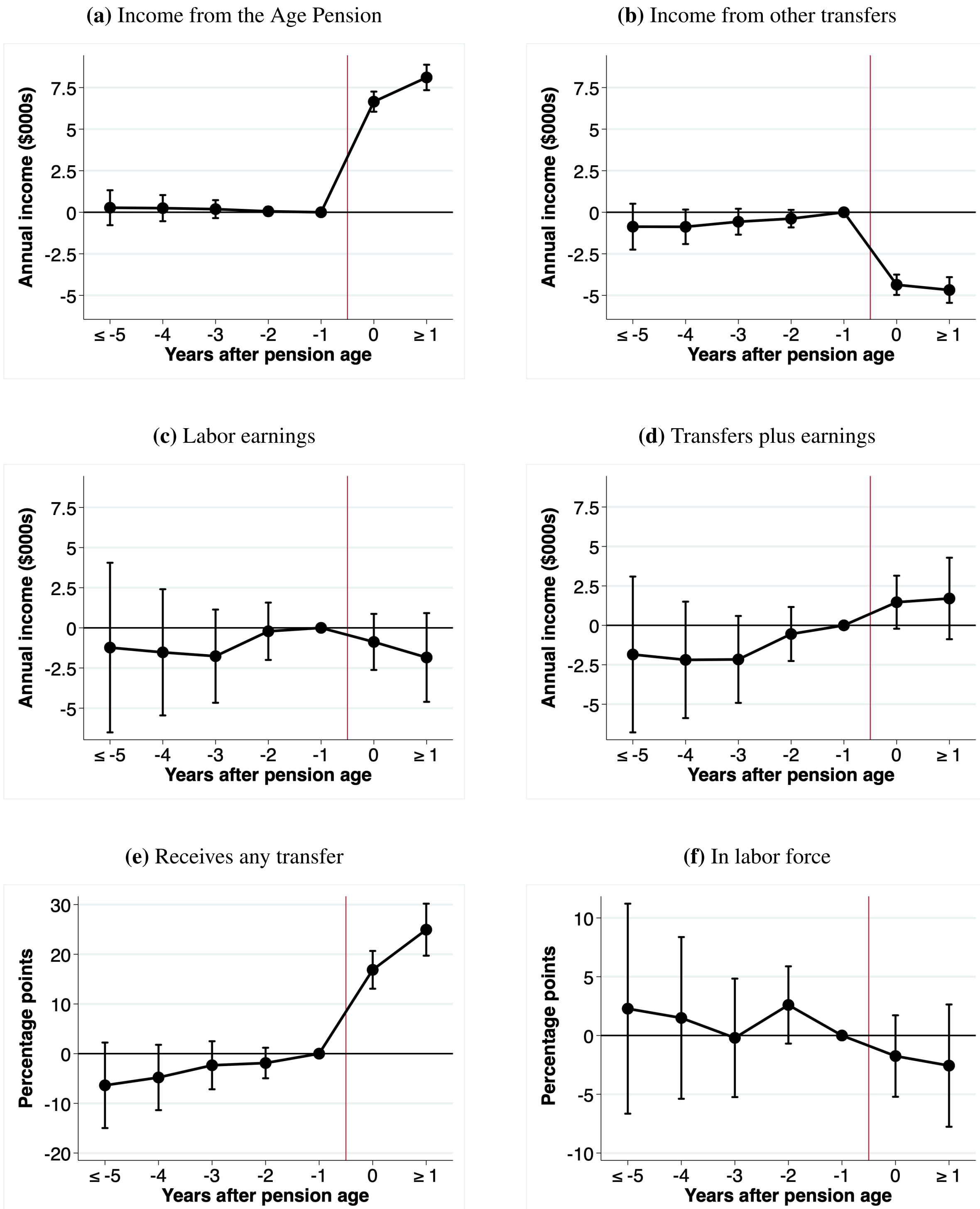


(d) ‘Working mothers can establish equally good relationships with their children’



Notes: These figures show how the average attitudes of women vary with their Age Pension Age (APA), with 95% confidence intervals around these averages. The figures indicate that women with a higher APA had more progressive attitudes about how paid work and child-rearing should be shared within the household between men and women. The statements above are paraphrased for presentation purposes. The full statements corresponding to each figure are: a) “It is better for everyone involved if the man earns the money and the woman takes care of the home and children”; b) as stated in the caption above; c) “Children do just as well if the mother earns the money and the father cares for the home and the children”; and d) “A working mother can establish just as good a relationship with her children as a mother who does not work for pay”. The sample for these graphs includes women aged 60–66 from waves 1, 5, 8 and 11 of the HILDA Survey, the waves that asked these questions to survey participants.

Figure A4: Estimated effects on women's outcomes relative to year before the pension age



(a) Income from the Age Pension

(b) Income from other transfers

(c) Labor earnings

(d) Transfers plus earnings

(e) Receives any transfer

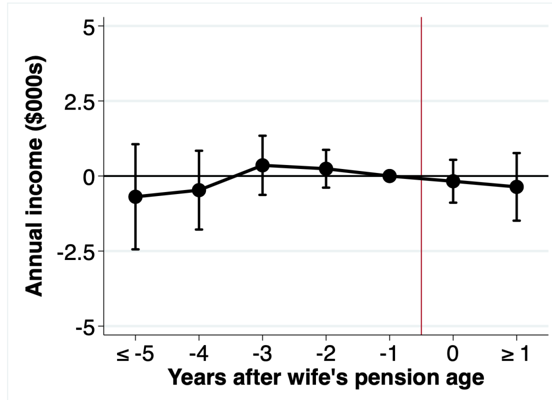
(f) In labor force

Notes: These figures plot the estimated δ coefficients, with 95% confidence intervals, from OLS estimates of equation (2). The estimates show the causal effect of women being a given number of years above or below the Age Pension Age (APA) relative to them being less than or equal to one year below the APA. Incomes are in 2016 Australian dollars. For these regressions, I expand the age sample to 55–69 (but only include women who are ever in the main sample). The sample comes from waves 1–14 of the HILDA survey.

Figure A5: Estimated effects on spousal outcomes relative to year before women's pension age

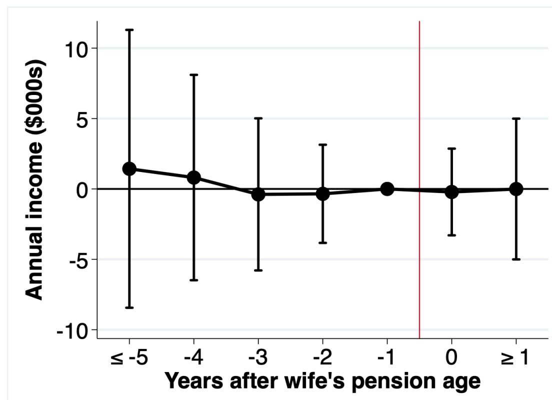
(a) Income from the Age Pension

(b) Income from other transfers



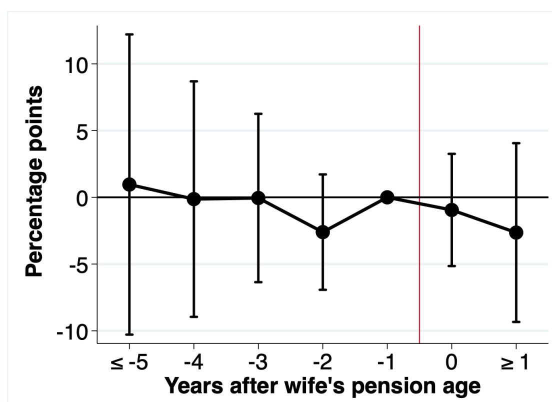
(c) Labor earnings

(d) Transfers plus earnings



(e) Receives any transfer

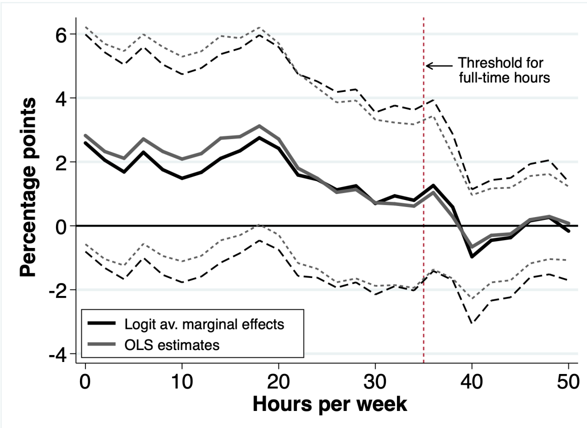
(f) In labor force



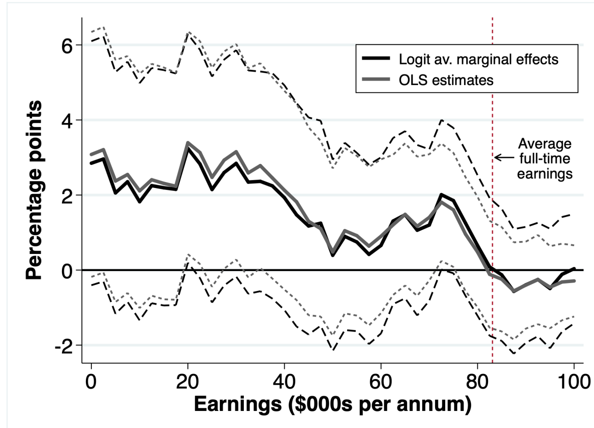
Notes: These figures plot the estimated δ coefficients, with 95% confidence intervals, from OLS estimates of equation (2). The estimates show the causal effect on spousal outcomes of women being a given number of years above or below the Age Pension Age (APA) relative to them being less than or equal to one year below the APA. Incomes are in 2016 Australian dollars. For these regressions, I expand the age sample for the female spouse to 55–69 (but only include women who are ever in the main sample). The sample comes from waves 1–14 of the HILDA survey.

Figure A6: Distribution regression estimates: Logit vs OLS

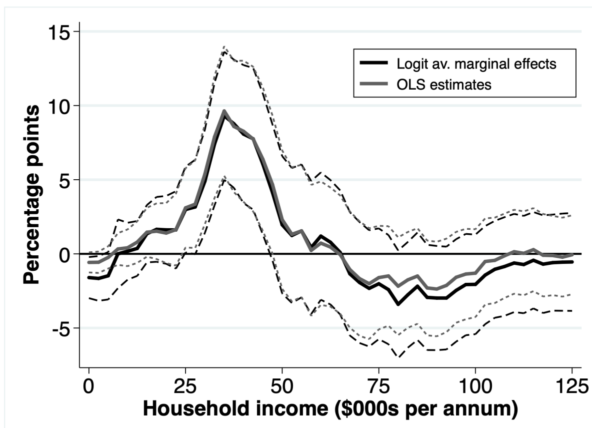
(a) Hours worked > $\$x$ per week



(b) Labor earnings > $\$x$ per annum



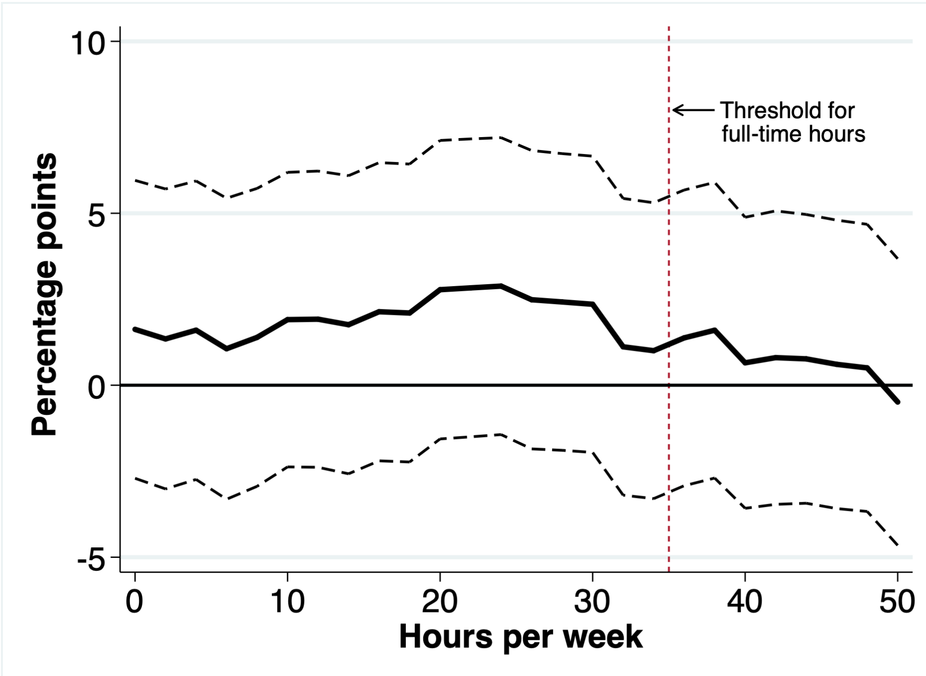
(c) Household income < $\$x$ per annum



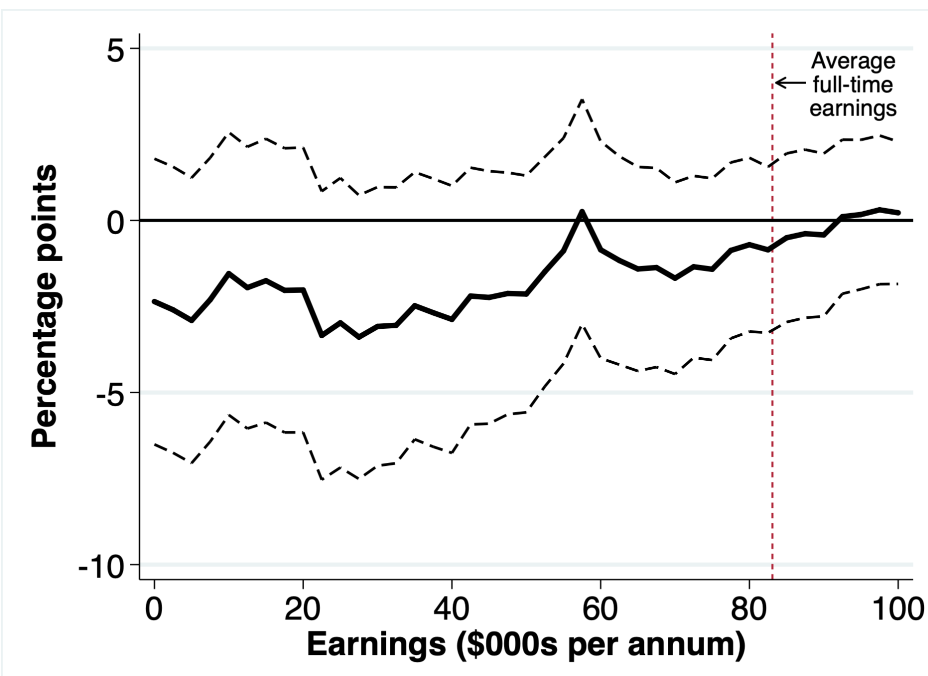
Notes: These figures compare the distribution regression estimates for OLS and logit models. Evidently, the average marginal effects from a logit model (and the associated 95% confidence intervals) are extremely similar to the OLS estimates in Figures 5 and 6.

Figure A7: Spousal labor supply responses across the distributions of hours worked and earnings

(**a**) Hours worked > *x* per week

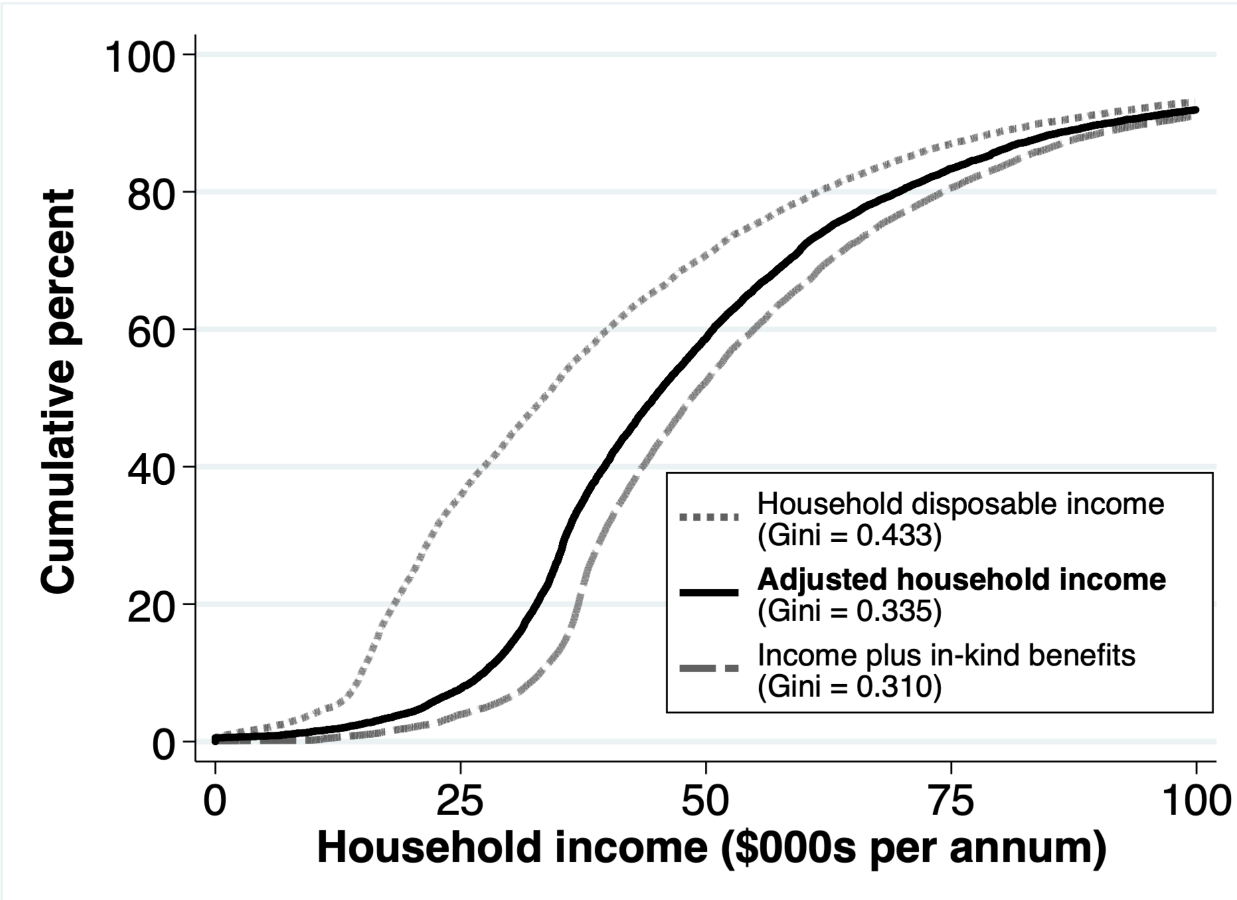


(**b**) Labor earnings > $*x* per annum



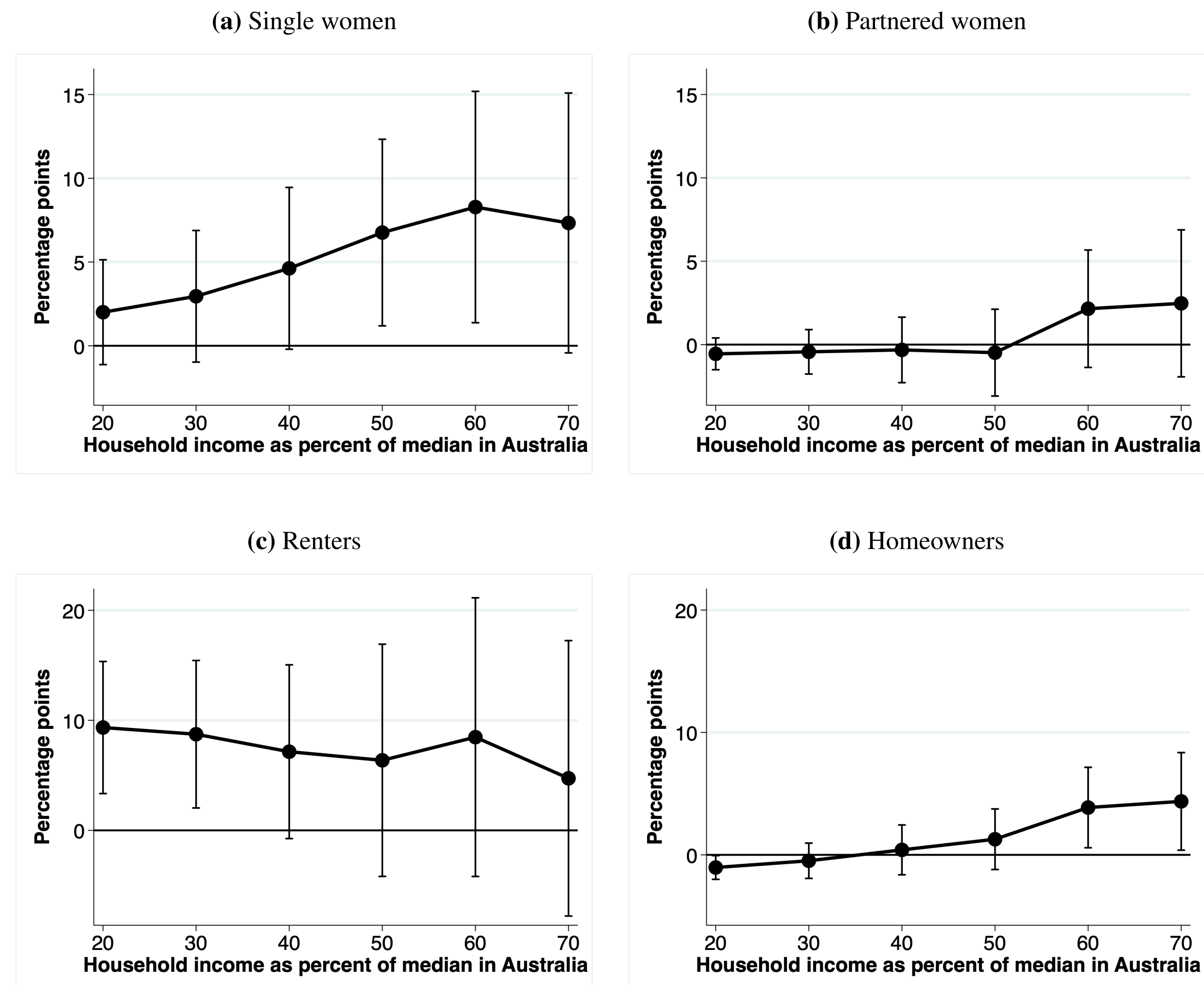
Notes: These figures present distribution regression estimates, with 95% confidence intervals, showing the causal effect of the reform on spouses' hours worked and earnings at different points in the distribution. In each regression, the dependent variable is an indicator for hours/earnings being above a given level, say *x*, and *x* is spaced at two-hour intervals in (a) and $2,500 intervals in (b). The sample consists of the spouses of women aged 60–66 from waves 1–14 of the HILDA survey. See the text in Section 5.2 for more details.

Figure A8: Impact of adjustments for housing costs and in-kind benefits on the income distribution



Notes: This figure shows how accounting for housing costs and in-kind benefits affects the cumulative distribution of household income for women in the sample. All three measures of household income are in 2016 dollars and are adjusted for household size. 'Income plus In-kind Benefits' is equal to household disposable income plus the value of in-kind benefits. 'Adjusted household income' is equal to household disposable income plus the value of in-kind benefits minus housing costs. The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey. See the text in Section 6.1 for more details.

Figure A9: Heterogeneity in the estimated effects on relative poverty: Household income $< x\%$ of the median level in the Australian population



Notes: This figure presents OLS regression estimates of the causal effect of the reform on relative poverty rates for different subgroups, with 95% confidence intervals, from estimates of equation (1). The dependent variable in each regression is an indicator for household income (after adjustments for housing costs and the value of in-kind benefits) being less than a given percentage, say $x\%$, of the median level in the Australian population, with x spaced at intervals of 10 from 20 to 70. The estimates are statistically different for single and partnered women at the 5% level for the 50% poverty line and at the 10% level for the 40% poverty line. The estimates are statistically different for renters and homeowners at the 1% level for the 20% and 30% poverty lines. The sample includes women aged 60–66 from waves 1–14 of the HILDA survey. See the text in Section 6.3 for more details.

Table A1: Number of observations in sample by women's pension age and age

Date of birth	Pension age	Age (in half-years) 60.0–61.0	61.5	62.0	62.5	63.0	63.5	64.0	64.5	65.0–66.5
Control cohorts										
Before July 1935	60.0									50
July 1935 – December 1936	60.5								17	175
January 1937 – June 1938	61.0					11	43	45	47	171
July 1938 – December 1939	61.5		13	34	45	51	43	57	45	179
Treated cohorts										
January 1940 – June 1941	62.0	111	**60**	47	49	44	48	39	54	191
July 1941 – December 1942	62.5	158	**50**	**50**	48	52	46	49	52	195
January 1943 – June 1944	63.0	174	**50**	**53**	**48**	58	43	58	47	212
July 1944 – December 1945	63.5	179	**56**	**56**	**57**	**54**	52	59	49	261
January 1946 – June 1947	64.0	203	**66**	**68**	**61**	**69**	**62**	84	79	333
July 1947 – December 1948	64.5	191	**60**	**68**	**68**	**77**	**81**	**80**	74	303
January 1949 – June 1952	65.0	525	**185**	**197**	**179**	**166**	**123**	**107**	**79**	72
Partially treated cohorts										
July 1952 – December 1953	65.5	257	**51**	**22**						
January 1954 – June 1955	66.0	74								

Notes: Each cell in this table presents the number of observations in the sample by women's age (in half years) and Age Pension Age (APA). Shaded cells indicate women who are below the APA. Women are treated when they are below the APA *and* aged 61.5–64.5 (cells in bold text). The sample consists of women aged 60–66 from waves 1–14 of the HILDA Survey.

Table A2: Robustness of main estimates to additional controls for time- and health-specific factors

	Annualised income				Other outcomes	
	Age Pension (1)	Other transfers (2)	Earnings (3)	Transfers + earnings (4)	Any transfer (5)	In labor force (6)
Baseline estimates						
Effect of reform	-6,957***	4,365***	1,458	-1,177	-0.182***	0.031*
	(301)	(312)	(892)	(855)	(0.019)	(0.018)
Observations	7,999	7,999	7,913	7,913	7,999	7,999
Including survey-wave dummies						
Effect of reform	-6,918***	4,434***	1,052	-1,472*	-0.178***	0.024
	(305)	(306)	(884)	(846)	(0.019)	(0.018)
Observations	7,999	7,999	7,913	7,913	7,999	7,999
Including controls for physical and mental health at survey date						
Effect of reform	-6,959***	4,349***	1,514*	-1,144	-0.184***	0.031*
	(301)	(309)	(893)	(859)	(0.019)	(0.018)
Observations	7,999	7,999	7,913	7,913	7,999	7,999

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by individual.

Notes: This table shows the robustness of the estimates in Tables 2 and 3 to the inclusion of additional controls for time- and health-specific factors. Incomes are in 2016 Australian dollars. Columns 3 and 4 exclude women who ever had earnings in the top 1% of workers. The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey.

Table A3: Testing for anticipatory changes in women's income and labor supply at ages 55–61

	Annualised income			Other outcomes	
	All transfers (1)	Earnings (2)	Transfers + earnings (3)	Any transfer (4)	In labor force (5)
Effect of one-year higher pension age	376	5,018	5,376	0.035	0.018
	(819)	(3,894)	(3,580)	(0.052)	(0.061)
Observations	5,036	4,896	4,896	5,036	5,036

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by individual.

Notes: This table presents OLS regression estimates of δ_{ant} from equation (A1). These regressions test whether there were any anticipatory changes in women's outcomes at ages 55–61. See Appendix A.1 for more details on these regressions. Incomes are in 2016 Australian dollars. Columns 2 and 3 exclude women who ever had earnings in the top 1% of workers. The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey.

Table A4: Characteristics of each subgroup: All waves

	Marital status		Home-ownership		Household wealth		
	Single women (1)	Partnered women (2)	Renters (3)	Home owners (4)	Bottom third (5)	Middle third (6)	Top third (7)
Demographics							
Age (years)	63.5	63.3	63.3	63.4	63.4	63.4	63.3
(Std. dev.)	(2.0)	(2.0)	(2.0)	(2.0)	(2.0)	(2.0)	(2.0)
Partnered	0%	100%	32.0%	72.2%	50.6%	73.4%	76.3%
Completed high school	24.5%	26.1%	17.0%	27.1%	15.2%	22.6%	38.0%
Completed year 10	73.2%	75.5%	58.2%	77.6%	59.9%	76.0%	87.5%
Home owner	69.5%	91.7%	0%	100%	61.2%	95.9%	97.1%
Household wealth in wave 2 ($000s)	480	804	140	774	110	453	1,522
(Std. dev.)	(617)	(1,288)	(363)	(1,023)	(92)	(120)	(1,631)
Government transfers							
Receiving Age Pension	31.2%	25.8%	37.2%	26.0%	40.9%	31.8%	13.8%
Payments if receiving Age Pension	17,543	12,435	17,298	13,676	16,176	13,076	11,421
(Std. dev.)	(4,669)	(4,351)	(4,283)	(5,049)	(4,099)	(4,874)	(5,729)
Receiving any transfer	58.8%	42.8%	76.2%	43.6%	75.4%	49.4%	19.9%
Payments if receiving any transfer	17,869	12,710	17,727	13,992	16,328	13,354	12,161
(Std. dev.)	(6,184)	(4,949)	(6,269)	(5,728)	(5,276)	(5,618)	(6,028)
Labor market							
In labor force	38.1%	34.1%	25.9%	37.0%	27.8%	36.6%	40.1%
Earnings if employed	48,505	42,964	40,920	45,538	40,298	44,585	48,820
(Std. dev.)	(27,412)	(27,610)	(25,417)	(27,869)	(25,007)	(27,209)	(29,726)
Number of observations	2,694	5,305	1,091	6,739	2,217	2,214	2,217

Notes: This table presents the means of the key variables for each subgroup and summarizes their characteristics. Income and wealth information is in 2016 Australian dollars. Wealth information is not available in wave 2 for 17% of observations. Household wealth is divided by 1.5 for partnered women, the equivalence scale implied by the Age Pension payment rules for singles and couples. The sample consists of women aged 60–66 years old from waves 1–14 of the HILDA survey. See the text in Section 3.1 for more details on the sample.

Table A5: Heterogeneity in the estimates by women's marital status and estimated effects on their spouses

	Annualised income				Other outcomes			
	Age Pension (1)	Other transfers (2)	Labor earnings (3)	Transfers + earnings (4)	Any transfer (5)	In labor force (6)	Employed (7)	Hours per week (8)
Panel A: Estimates by marital status								
Single women								
Effect of reform	-10,106***	6,443***	3,850**	104	-0.165***	0.069**	0.066**	1.83
	(610)	(687)	(1,849)	(1,703)	(0.031)	(0.032)	(0.031)	(1.12)
Mean at pension age	9,860	1,546	14,651	26,214	0.618	0.338	0.321	10.12
Partnered women								
Effect of reform	-5,337***	3,272***	405	-1,687*	-0.193***	0.013	0.010	0.09
	(314)	(307)	(951)	(943)	(0.024)	(0.021)	(0.021)	(0.69)
Mean at pension age	5,297	1,084	9,233	15,673	0.531	0.297	0.293	6.84
p-value on equality of treatment effects	<0.001	<0.001	0.099	0.358	0.492	0.147	0.129	0.186
Panel B: Effects on spouses								
Effect of reform	-70	381	-1,797	-1,467	-0.007	0.003	0.002	0.59
	(298)	(375)	(1,716)	(1,655)	(0.022)	(0.022)	(0.022)	(0.97)
Mean at pension age	4,195	2,617	16,230	23,092	0.484	0.404	0.397	12.76

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by female individual.

Notes: This table presents OLS regression estimates of the causal effect of the reform, δ, from equation (1) for different subgroups based on women's marital status. Single women comprise 33% of the full sample; partnered women comprise 67% (see Table A4 for the characteristics of each group). 'Mean at pension age' is the mean of the dependent variable among individuals who have reached the pension age (or whose wife reached the pension age in Panel B) in the last 12 months. Incomes are annualized and in 2016 Australian dollars. Columns 3 and 4 exclude individuals who ever had earnings in the top 1% of workers. The sample in Panel A consists of women aged 60–66 from waves 1–14 of the HILDA survey. The sample in Panel B consists of their spouses (for those who are partnered).

Table A6: Heterogeneity in the estimated effects on women by financial circumstances

	Annualised income				Other outcomes			
	Age Pension (1)	Other transfers (2)	Labor earnings (3)	Transfers + earnings (4)	Any transfer (5)	In labor force (6)	Employed (7)	Hours per week (8)
Panel A: Estimates by home-ownership status								
Renters								
Causal effect of reform	-12,661***	9,271***	3,962*	614	-0.122***	0.089*	0.087**	2.12
	(890)	(1,149)	(2,129)	(1,819)	(0.047)	(0.048)	(0.044)	(1.58)
Mean at pension age	12,642	2,317	4,599	19,691	0.825	0.175	0.158	4.70
Home owners								
Causal effect of reform	-5,853***	3,602***	1,108	-1,192	-0.181***	0.017	0.014	0.42
	(309)	(301)	(1,000)	(969)	(0.021)	(0.020)	(0.019)	(0.64)
Mean at pension age	5,875	1,101	12,068	19,124	0.516	0.331	0.324	8.32
p-value on equality of treatment effects	<0.001	<0.001	0.222	0.376	0.253	0.164	0.126	0.314
Panel B: Estimates by household wealth								
Bottom third								
Causal effect of reform	-11,688***	8,694***	4,120***	1,126	-0.163***	0.073**	0.072**	1.86*
	(564)	(713)	(1,363)	(1,203)	(0.030)	(0.031)	(0.029)	(1.03)
Mean at pension age	11,137	2,394	8,154	21,685	0.826	0.263	0.242	6.63
Middle third								
Causal effect of reform	-6,004***	3,877***	-1,861	-4,058**	-0.181***	0.010	-0.001	-1.15
	(550)	(506)	(1,840)	(1,744)	(0.039)	(0.034)	(0.033)	(1.09)
Mean at pension age	6,614	1,028	13,476	21,258	0.577	0.302	0.302	8.75
Top third								
Causal effect of reform	-3,276***	1,128***	2,572	381	-0.203***	0.004	0.008	0.48
	(430)	(323)	(1,740)	(1,742)	(0.032)	(0.032)	(0.031)	(1.02)
Mean at pension age	2,867	249	11,393	14,573	0.289	0.366	0.358	8.63
p-value on equality of treatment effects	<0.001	<0.001	0.032	0.045	0.674	0.225	0.190	0.133

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by individual.

Notes: This table presents OLS regression estimates of the causal effect of the reform, δ, from equation (1) for different subgroups. Renters comprise 14% of the full sample; homeowners comprise 84% (see Table A4 for the characteristics of each group). 'Mean at pension age' is the mean of the dependent variable among individuals who have reached the pension age in the last 12 months. Incomes are annualized and in 2016 Australian dollars. Columns 3 and 4 exclude women who ever had earnings in the top 1% of workers. Information on household wealth comes from wave 2, the first wave with information on wealth, and is missing for 17% of observations. I adjust for household size by dividing household wealth by 1.5 for couples. The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey.

Table A7: Average total value of in-kind benefits per annum to households by household composition and concession card ownership, adjusted for household size

	Number of concession card owners in household		
	0	1	2
Single women	\$9,365	\$15,559	
Couples	\$8,500	\$12,738	\$16,114

Notes: This table presents the average total value of in-kind benefits to households, in 2016 dollars, based on household composition and concession card ownership. I adjust for household size by dividing the value of in-kind benefits by 1.5 for couples. The estimates come from the 2003–04 and 2009–10 cross-sections of the Australian Bureau of Statistics' Fiscal Incidence Study (with the sample restricted to households containing women aged 60–64).

Table A8: Estimated effects on women's ownership of concession cards

	Any card (1)	Seniors card (2)
Effect of reform	-0.446***	-0.254***
	(0.021)	(0.017)
Mean at pension age	0.776	0.226
R-squared	0.380	0.181
Observations	7,999	7,999

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by individual.

Notes: This table presents OLS regression estimates of the causal effect of the reform, δ, from equation (1). 'Mean at pension age' is the mean of the dependent variable among individuals who have reached the pension age in the last 12 months. Information on concession card ownership is imputed based on women's age, receipt of qualifying transfers and household income (see Appendix A.2 for the details). The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey.

Table A9: Sensitivity of inequality and poverty estimates to value of in-kind benefits

	In-kind benefits valued at x% of levels in Fiscal Incidence Study					
	125% (1)	**100% (Baseline) (2)**	75% (3)	50% (4)	25% (5)	0% (6)
Panel A: Estimated effects on inequality measures						
Gini	+13.0%	**+11.5%**	+9.5%	+8.1%	+6.9%	+4.8%
	[$p = 0.036$]	[$p = 0.048$]	[$p = 0.084$]	[$p = 0.134$]	[$p = 0.182$]	[$p = 0.288$]
80:20 ratio	+14.3%	**+15.4%**	+14.7%	+17.4%	+18.7%	+21.0%
	[$p < 0.002$]	[$p < 0.002$]	[$p = 0.002$]	[$p = 0.004$]	[$p = 0.006$]	[$p = 0.012$]
Panel B: Estimated effects on relative poverty rates						
50% of median	0.019	**0.021**	0.027*	0.041**	0.029	0.040*
	(0.012)	(0.013)	(0.015)	(0.016)	(0.020)	(0.022)
60% of median	0.046***	**0.044****	0.036*	0.031	0.047**	0.018
	(0.015)	(0.017)	(0.019)	(0.021)	(0.022)	(0.022)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The p-values in Panel A are constructed using a pairs cluster bootstrap with 999 replications. The standard errors in Panel B (in parentheses) are clustered by female individual.

Notes: This table examines the sensitivity of the estimated effects on inequality measures and relative poverty rates to different assumptions about the value of in-kind benefits. Column 2 presents the baseline estimates, where in-kind benefits are valued according to the levels in the Fiscal Incidence Study (FIS). Column 1 presents the estimates when in-kind benefits are valued *higher*, at 125% of the levels in the FIS, while columns 3–6 present the estimates when in-kind benefits are valued *lower*, at 75%, 50%, 25% and 0% of the levels in the FIS. Relative poverty lines are adjusted in line with the assumed value of in-kind benefits. The sample includes women aged 60–66 from waves 1–14 of the HILDA survey. See the text in Section 6.4 for more details.