MAX-PLANCK-INSTITUT FÜR SOZIALRECHT UND SOZIALPOLITIK
MAX PLANCK INSTITUTE FOR SOCIAL LAW AND SOCIAL POLICY

mea *Munich Center for the Economics of Aging*

The Unequal Burden of Retirement Reform: Evidence from Australia

Todd Morris

10-2019

MEA DISCUSSION PAPERS

mea – Amalienstr. 33_D-80799 Munich_Phone +49 89 38602-355_Fax +49 89 38602-390_**www.mpisoc.mpg.de**

The Unequal Burden of Retirement Reform: Evidence from Australia

Todd Morris

Updated version: June 29, 2020

Abstract:

As governments try to contain rising expenditure on retirement pensions by increasing eligibility ages, there are concerns that such reforms disproportionately affect poorer households. Using detailed longitudinal data, I examine this trade-off in the context of an Australian reform that increased women's pension-eligibility age from 60 to 65. While this reform led to significant reductions in net government expenditure, the negative effects on household incomes were concentrated among poorer households. These unequal impacts meant that, among affected cohorts, the reform increased relative poverty rates by 20 to 37 percent and inequality measures by 11 to 36 percent.

Zusammenfassung:

Mit staatlichen Eindämmungsversuchen steigender Rentenausgaben durch eine Anhebung des Renteneintrittsalters gehen Bedenken einher, dass solche Reformen ärmere Haushalte unverhältnismäßig betreffen. Anhand detaillierter Längsschnittdaten untersuche ich diesen Zielkonflikt im Zusammenhang mit einer australischen Reform, bei der das Renteneintrittsalter für Frauen von 60 auf 65 Jahre angehoben wurde. Während die Reform zu einer signifikanten Senkung der staatlichen Nettoausgaben geführt hat, konzentrieren sich die negativen Auswirkungen auf Einkommen ärmerer Haushalte. Diese ungleichen Auswirkungen der Reform haben dazu geführt, dass sich die relative Armutsquote um 20 bis 37 Prozent und Ungleichheitsmaße um 11 bis 36 Prozent in den betroffenen Kohorten erhöht hat.

Keywords:

Retirement age, distributional effects, poverty, inequality

JEL Classification:

H55, I38, J26

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

First version: July 5, 2019

This version: June 29, 2020

Abstract

As governments try to contain rising expenditure on retirement pensions by increasing eligibility ages, there are concerns that such reforms disproportionately affect poorer households. Using detailed longitudinal data, I examine this trade-off in the context of an Australian reform that increased women's pension-eligibility age from 60 to 65. While this reform led to significant reductions in net government expenditure, the negative effects on household incomes were concentrated among poorer households. These unequal impacts meant that, among affected cohorts, the reform increased relative poverty rates by 20 to 37 percent and inequality measures by 11 to 36 percent.

*Acknowledgements: For helpful comments and suggestions, I thank John de New, Bill Evans, Greg Gilpin, Tim Moore, Stefan Staubli, Cain Polidano, Roger Wilkins and Jenny Williams. Useful feedback was also provided by participants at many 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:** Max Planck Institute for Social Law and Social Policy, Amalienstrasse 33, Munich 80799, Germany. Email: toddstuartmorris@gmail.com.
Data availability: This paper uses a restricted release of the HILDA survey that is available to researchers in Australia. I agree to provide assistance to other researchers wishing to obtaining access. Other replication materials, such as code, will be posted in an Online Appendix. **Disclosure statement:** I do not have anything to disclose.

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, 2020) and increasing the employment rates of older individuals (Staubli and Zweimüller, 2013; Atalay and Barrett, 2015; Cribb, Emmerson and Tetlow, 2016; Manoli and Weber, 2016; Geyer and Welteke, 2019).

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 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

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

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 the reform on households' labor supply and income. I start my analysis by estimating the average impacts 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, 2020), 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 transfers. In contrast, the impact 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. These groups also 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 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 across the entire 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 preferred measure of household income is household disposable income net of 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 and no 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 11–36% among affected households (households containing women in the affected cohorts). This increase is substantial; the same increase could be achieved by transferring around AU$5,000 per annum (US$3,500) 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 around 3 percentage points or 20–37%, with single women and renters explaining most of this increase.

As a final step, I estimate the net fiscal impact of the reform, considering (i) changes in government expenditure on the retirement pension, other transfers and in-kind benefits and (ii) changes in government revenue from income tax. I estimate that each one-year increase in the pension age saved around 750 million dollars per annum, which is equivalent to 1.7% of current government expenditure on the retirement pension.

Overall, these findings highlight a significant trade-off for policymakers: while raising pension-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 twice as large (1.6 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 over three-times larger (39% vs. 11% based on the Gini coefficient) and the increase in relative poverty rates seven-times larger (22 vs. 3 percentage points). While these calculations ignore behavioral responses, they demonstrate the importance of other government transfers as a buffer for low-income households.

This paper relates to a growing 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). These studies consistently find that Social Security is the most important welfare program in reducing poverty. However, these studies are purely descriptive and there have been few attempts to estimate the distributional effects of changes in retirement policy. One early exception is Engelhardt and Gruber (2004). Using cohort variation from the Social Security notch, they estimate that a 1% increase in Social Security income decreases old-age poverty by a similar percentage. More recently, two studies exploit pension reforms in Europe to study the impact of delayed pension eligibility on household incomes (Geyer et al., 2020; Cribb and Emmerson, 2019). Geyer et al. (2020) study a German reform that increased the Early Retirement Age (ERA) of women from 60 to 63 and find no effect on average household incomes, both overall and for various subgroups. Cribb and Emmerson (2019) study a similar reform in the U.K. and find negative effects on household incomes and an increase in relative poverty rates by 6.4 percentage points (43%).

In this paper, I build on these studies in several ways. First, I estimate the impact of a large Australian reform on the entire distribution of household income and quantify the effects on inequality measures. My estimates account for a comprehensive set of responses affecting household incomes and eligibility for in-kind benefits. Moreover, like Bitler, Gelbach and Hoynes (2006), my estimates allow for income changes within — as well as between — subgroups, which is important as most of the increase in inequality is explained by within-group changes. Second, I find material increases in poverty and inequality measures despite stronger substitution to other government transfers than similar reforms in other countries (Duggan, Singleton and Song, 2007; Staubli and Zweimüller, 2013; Geyer and Welteke, 2019) and an overall welfare system with explicit redistributional goals. Thus, the estimated effects on poverty and inequality measures may not be an upper bound for the effects in countries with less targeted pensions, especially countries with weaker safety nets for people of working age. Third, my results highlight the importance of other government transfers in attenuating the regressive impacts of increases in pension-eligibility ages. While many studies highlight the fiscal effects of substitution between government programs (e.g., Staubli, 2011; Staubli and Zweimüller, 2013; Borghans, Gielen and Luttmer, 2014; Inderbitzin, Staubli and Zweimüller, 2016; Autor et al., 2019), there has been little consideration of how this behavior may limit the regressive impacts of cuts to social insurance programs.

My paper also relates to a broader 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 studies initially focused on increases in the Normal Retirement Age in the U.S., a policy that is distinct from the Australian reform because it reduces the generosity of retirement benefits but does not affect the minimum claiming age.[2] The Australian reform is more like an increase in the ERA, a policy that is occurring widely. Several recent papers examine changes in the ERA and find large employment responses (Börsch-Supan and Schnabel, 1998; Vestad, 2013; Staubli and Zweimüller, 2013; Manoli and Weber, 2016; Cribb, Emmerson and Tetlow, 2016; Geyer and Welteke, 2019), while others document spillovers on other government programs (Staubli and Zweimüller, 2013; Geyer et al., 2020). For example, Staubli and Zweimüller (2013) use a similar identification strategy to study gradual increases in the ERA in Austria. They estimate large increases in employment rates (10–11 percentage points) and similar increases in the receipt of unemployment insurance. Relative to this literature, I consider a more comprehensive set of responses, allowing for earnings and benefit-substitution responses from both women and their spouses, and heterogeneity in the labor supply response based on earnings and household wealth. Overall, 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 existing studies (Staubli and Zweimüller, 2013; Cribb, Emmerson and Tetlow, 2016; Geyer et al., 2020). These differences are consistent with the more targeted nature of Australia's welfare system, which meant that the reform had stronger impacts on poorer households, who had a relatively weak attachment to the labor force and a strong reliance on other transfers.

[2]For example, see Duggan, Singleton and Song (2007); Song and Manchester (2007); Mastrobuoni (2009); Blau and Goodstein (2010); and Behaghel and Blau (2012). Other studies examine similar reforms in Europe (Hanel and Riphahn, 2012; Lalive, Magesan and Staubli, 2017; Engels, Geyer and Haan, 2017) or exploit cohort variation in Social Security income (Krueger and Pischke, 1992; Snyder and Evans, 2006; Gelber, Isen and Song, 2016).

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 55 for all cohorts in my sample. Although private sources of retirement funding are becoming more important, most Australians still partially or fully fund their retirement via the Age Pension; in 2013, the Age Pension was received by 73% of women above the APA and 68% of men (Oguzoglu, Polidano and Vu, 2020). Age Pensioners receive payments every two weeks (fortnightly). Payment rates do not depend on the age at which individuals start claiming, which creates a strong incentive to start claiming at the APA. Eligibility is not contingent on employment history but is subject to a means test.

The means test also determines 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 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 per person 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).[3] 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.

For households who are not eligible for the maximum payment, payments gradually decrease with income or assets above the relevant phase-out threshold until they reach zero at the income or assets limit. In practice, the income test is relatively strict and binds for about twice as many pensioners as the assets test (Chomik et al., 2018). The income test is based on household income per fortnight and assesses most forms of income, including earnings, superannuation income and investment income. In 2016, each dollar of income above the income threshold reduced payments by 50 cents (25 cents each for couples). The assets test excludes the family home — the largest component of wealth for elderly Australians — but assesses most other assets. In 2016, each $1,000 of assets above the asset threshold reduced payments by $1.50.[4]

While the Age Pension is the main source of public retirement funding, around one in three women receive a regular transfer from the Australian Government before they reach the APA (Oguzoglu, Polidano and Vu, 2020). These transfers have stricter eligibility conditions than the Age Pension but are similar in other respects: they are non-contributory and have similar means tests and payment rates. For example, the two most common transfers, the Disability Support Pension and Carer Payment, have the same means test and payment rates, while other transfers, such as the transfer for the unemployed, are less generous. Generally, recipients of other transfers are moved onto the Age Pension at the APA, which results in an increase in income or no change.

Older Australians also receive valuable in-kind benefits through large discounts on healthcare and other expenses. Eligibility for these discounts is based on having a concession card (which can be carried like a credit card). There are two main concession cards: the Pensioner Concession Card

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

[4] In 2016, the phase-out thresholds for the income test were $162 per fortnight for singles and $292 for couples, and the respective income limits were $1,918.20 and $2,936.80. Most full-time workers would exceed these limits, with average full-time earnings equal to $3,146 per fortnight. For the assets test, the phase-out threshold for homeowners was $209,000 for singles and $296,500 for couples, and the respective limits were $788,250 and $1,170,000. For renters, the respective phase-out thresholds and limits are around $150,000–200,000 higher.

(hereafter called the "pensioner card") and the Commonwealth Seniors Health Card (the "seniors card"). Eligibility for both cards increases sharply at the APA. The pensioner card is available to older Australians who receive a government transfer; Age Pensioners qualify automatically, as do recipients of most other transfers if they are over 60. The seniors card is available to a wealthier group of older Australians who (i) have reached the APA, (ii) are not receiving a government transfer, and (iii) have household income below a maximum level (which is similar to the income limit for the Age Pension). Both cards considerably reduce out-of-pocket medical expenses and provide other discounts.[5] Overall, these benefits are substantial; according to Harmer's (2008) estimates, the total value of in-kind benefits for Age Pension households is equivalent to 82% of their pension income, with around three-quarters of this value coming from cheaper healthcare.

All things considered, the Age Pension is an especially important source of retirement funding for women. Conditional on reaching the APA, women are more likely to receive the pension than men, and they generally receive it for longer due to their greater longevity. Women also have fewer assets to self-fund their retirement, especially in superannuation. According to Clare (2015), the mean superannuation balance of men aged 60–64 in 2013–14 was double that of women ($292,510 vs. $138,154), and men's median balance was almost four-times larger ($100,000 vs. $28,000).

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, rising by six months every two years from July 1995 to July 2013. Initially proposed in 1993, the reform aimed to foster employment among older women, reduce government expenditure, and ultimately harmonize the APA of men and women. As Figure 1 shows, the impact of the reform varied by birth cohort; while the reform had no effect

[5] For example, both cards offer large discounts on doctor visits and prescription medicines. 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. Furthermore, there are discounts on public transport fares, council and water rates, electricity costs, and vehicle registration charges (Harmer, 2008).

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 — by an extra six months for each eighteen-month cohort — until the APA reached 65 for women born after December 1948.[6]

For the rest of this section, I discuss the expected effects of the reform on the labor supply and total income of affected households. In line with the empirical analysis, I focus on the effect of women being delayed from reaching the APA and abstract from possible effects at ages where women's pension eligibility was not affected. This approach is consistent with most studies in the literature examining changes in eligibility ages (e.g., Staubli and Zweimüller, 2013; Cribb, Emmerson and Tetlow, 2016; Oguzoglu, Polidano and Vu, 2020).[7]

Expected effects of the reform. Basic labor supply theory predicts heterogeneous impacts of the reform on households' labor supply and income. To explain these predictions, Figure 2 depicts the impact on the household budget constraint, in terms of the set of feasible combinations of total income and leisure. Figure 2 shows the respective budget constraints of households with women who are above the APA (in black) and below the APA (in gray).[8] We can think about the impact of the reform as a shift in the budget constraint from the black line to the gray line.

The shift in the budget constraint varies with earnings because of the income test. At zero or low levels of hours worked, where households earn less than the phase-out threshold, the budget constraint shifts down by the amount of the maximum pension payment. For a slightly higher level of hours, 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 only eligible for a reduced pension. In addition, the slope of the budget constraint doubles in this region because the income test reduces

[6]Under a 2009 reform, the APA is increasing further towards 67. However, my estimates do not capture any effects of this reform as pension eligibility was not affected until 2017, two years after the end of the sample period.

[7]For completeness, I also examine whether the reform caused anticipatory effects at earlier ages. The estimates (discussed in Section 4) do not show statistically significant evidence of such effects.

[8]For simplicity, I assume that households are eligible for the maximum Age Pension payment under the assets test and ignore the presence of in-kind benefits and other government transfers.

pension payments by 50 cents for each dollar of earnings above the phase-out threshold. Finally, for households earning 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. Otherwise, I consider a similar static model of labor supply to Bitler, Gelbach and Hoynes (2006): I assume (i) that households can freely choose hours of work at the given wage, which is constant, and (ii) that leisure and consumption are both normal goods.

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

Several insights further inform the predictions. First, due to the assets test and the higher payment rate for singles, the impact on the budget constraint is larger for single and poorer households, and there is no impact for the wealthiest households. As such, we would expect the effects on labor supply and total income to be larger among single women and poorer households. Second, as many women receive other transfers prior to the APA, the labor supply and distributional impacts are likely to be attenuated by benefit substitution. Third, the reform is likely to delay many women from becoming eligible for the generous in-kind benefits associated with concession card ownership. While these benefits are less targeted than the Age Pension, they still disproportionately benefit lower income households (Harmer, 2008). Thus, in estimating the overall distributional impacts of the reform, I consider changes in both household income and in-kind benefits.

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 the key variables. 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.[9] Then, I examine the effects on

[9] 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

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.[10] 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 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 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. The sample consists of women who are 63.4 years old on average. 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 half receive a transfer from the Australian Government

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).

[10]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. This restriction excludes about 1% of the sample. The estimated effect on women's earnings is 18% smaller without this restriction.

(48.2%), with 27.6% receiving the Age Pension, while only 35.5% 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, spouses have a higher labor force participation rate (42.6%) and, conditional on employment, higher earnings ($63,614 per annum).

Table 1 also shows how the characteristics of the sample have changed over the sample period, presenting the means of these variables in the first and last waves (1 and 14). 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 other cohort factors are present.

3.2 Graphical evidence: Trends around the Age Pension Age

This section presents graphical evidence of the effects of the reform. Figure 3 shows trends in women's labor force participation and transfer receipt by age 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 65% for women in the first cohort, 40% for women in the second cohort, and 0% for women in the third cohort. However, these differences mostly 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 many women. While each cohort has similar rates of transfer receipt at age 65, there are large differences at 62–64. For example, women with an APA of 64.0–65.0 are around 35 percentage points less likely to receive a transfer at age 63 than women with an APA of 61.0–62.0, but the gap is around 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. Rather, Figure 3d shows persistently higher labor force participation rates for women with a higher APA. While these differences may partially reflect responses to the reform, there is strong evidence that other cohort factors explain most of this trend. In a companion paper (Morris, 2019), I document a strong upward trend in female labor force participation rates across the relevant cohorts prior to the reform. This trend is also consistent with trends in education, birth rates and social norms among women in the sample; Figures A2 and A3 show that women with a higher APA have higher levels of education, fewer children and more progressive attitudes about the division of paid work and child-rearing.

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. Figure 4a shows that women's Age Pension receipt increases by around 50 percentage points in the year they reach the APA, but the overall increase in transfer receipt is 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 transfer income increases by around \$2,500. There remains only weak evidence of women decreasing their labor supply at the APA; while labor force participation and earnings decrease 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 the reform 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.[11] 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, 2020). 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 by estimating 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

[11] 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.

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.[12]

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 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 macroeconomic conditions at each age. Over the sample period, Australia's economy was very stable, with low and stable unemployment rates (4.2–6.7%) and no recession. Moreover, the Australian 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 at-

[12] I estimate all regressions by OLS. I cluster standard errors by 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.

tenuate estimates of δ as these responses are likely to be absorbed by the cohort fixed effects. In Appendix A.1, I adopt the approach of Duggan, Singleton and Song (2007) to test whether women responded to changes in the net present value of their future Age Pension income at ages 55–61. The estimates are small and statistically insignificant at conventional levels. These null results are consistent with well-identified estimates of anticipatory responses to similar reforms in other countries. Using a regression discontinuity design, Geyer and Welteke (2019) examine a sharp increase in the pension-eligibility age from 60 to 63 for German women born after December 1951. They find a precisely estimated zero effect on employment prior to age 60 despite (i) large effects at ages 60–61 (14.4 percentage points) and (ii) the fact that the reform was well anticipated.

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 discuss the average impacts on women and verify the robustness of the empirical strategy. I also discuss heterogeneity in the effects based on household characteristics and present the estimated effects on spouses. In Section 5.2, I examine how the labor supply response varied across the distributions of earnings and hours worked.

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 (Atalay and Barrett, 2015; Oguzoglu, Polidano and Vu, 2020), the estimates show that many women offset their delayed Age Pension receipt by extending their receipt of other transfers. While the reform decreased Age Pension receipt among women by an estimated 48.3 percentage points ($p < 0.01$), it increased their receipt of other transfers 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 ($p < 0.01$).

Table 2 also presents the estimated effects on women's labor supply (in columns 4–6). Overall, the estimates indicate that the reform caused modest increases in female labor supply. The estimated increase in women's labor force participation is 3.1 percentage points ($p = 0.085$), with most of this effect explained by a 2.7-percentage-point increase in employment ($p = 0.121$). The estimated effect on hours worked is an increase of just 0.6 hours per week ($p \geq 0.1$).

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. In a companion paper (Morris, 2019), I show that this difference stems from different empirical strategies. Atalay and Barrett (2015) use a differences-in-differences approach in which male cohorts form the comparison group for changes in female participation rates. However, Morris (2019) shows that this strategy is likely to overestimate the impact on female participation rates; Morris (2019) documents opposing pre-reform trends for the relevant cohorts of men and women, with a strong positive trend in female labor force participation rates and a weak negative trend in male participation rates. As I directly control for differences in participation rates across female cohorts, my estimates are considerably smaller.[13]

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 ($p \geq 0.1$).

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

[13] The smaller effects estimated here cannot be explained by differences in the sample. Atalay and Barrett (2015) use repeated cross-sections of the nationally representative Income and Housing Costs surveys from 1994 to 2010. Morris (2019) successfully replicates this sample and estimates similar regressions to equation (1), finding a similarly modest increase in female labor force participation of 4.1 percentage points ($p = 0.109$).

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, I estimate the following regressions:

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

where these regressions modify equation (1) by replacing the key regressor, $\mathbf{1}(age_{it} < APA_i)$, with six indicators that 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 who are between zero and 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 between zero and 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 $j < -1$.[14]

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) sharp increases in women's total income from transfers and transfer receipt (both $p < 0.01$); (ii) a sharp decrease in women's income from transfers other than the Age Pension ($p < 0.01$), as many women move from other transfers onto the Age Pension; (iii) relatively small and statistically

[14] To enhance precision, I use an expanded age sample of 55–69 (but only include women in the main sample).

insignificant decreases in women's earnings and labor force participation (both $p \geq 0.1$); and (iv) a modest increase in women's combined income from transfers and earnings ($p = 0.087$).

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 economic 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 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 show larger income losses from the Age Pension 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 over twice as much Age Pension income as homeowners and (ii) women in the poorest third of households lost nearly twice as much as women in the middle third and over three-times as much as women in the wealthiest third.[15] Second, the estimated increase in income from other transfers is larger among single and poorer groups of women in both absolute terms and as a proportion of lost Age Pension income. 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 the differences between groups are generally not statistically significant. For example, the estimated increase in labor force participation is 6.6 percentage points for single

[15] 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.

women ($p = 0.030$), 8.9 percentage points for renters ($p = 0.065$) and 7.3 percentage points for women in the poorest third of households ($p = 0.019$), while the estimates for all other groups are close to zero and statistically insignificant at the 10% level.

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.[16] 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., 2020).

However, this finding contrasts with the results of Atalay, Barrett and Siminski (2019), who study the same reform and estimate an increase in the participation rates of spouses of 6.8 percentage points (which is outside the 99% confidence interval of my estimate). This difference appears to stem from the fact that my regressions include fixed effects for the birth cohorts of 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 ($p = 0.014$, not shown).[17] In this context, it is important to include such controls. There is a strong trend in female labor force participation rates across the relevant cohorts prior to the reform, with much higher participation rates for later cohorts (Morris, 2019), and this trend is especially large for partnered women. This trend could directly increase the participation rates of husbands due to preferences for joint retirement. Such preferences have been emphasized as an explanation for the tendency of couples to retire close together regardless of age differences (e.g., Gustman and Steinmeier, 2000, 2004; Coile, 2004).

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

[17] The estimates here (and generally) are similar with fixed effects for survey wave rather than women's birth cohort; with wave fixed effects, the estimated effect is 0.9 percentage points ($p \geq 0.1$, not shown).

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, we expect heterogeneity in the response across the distributions of hours worked and earnings. Therefore, 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 a sequence of binary regressions, where in each regression the dependent variable is an indicator for the relevant variable — in this case hours or earnings — being above/below a given threshold. This is analogous to the common practice of estimating quantile regressions at a range of quantiles.[18]

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 and their 95% confidence intervals. Overall, the estimates are modest in size but consistent with the theoretical predictions. The estimates show the strongest responses from part-time and lower earning women, and there is little evidence of any response from full-time workers or women with earnings above the income limits for the pension. Specifically, 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 estimates for earnings. For earnings-thresholds up to

[18] In this context, I favor distribution regression over quantile regression for two reasons (though quantile regression estimates show similar effects). First, distribution regression outperforms quantile regression in the presence of mass points (Chernozhukov, Fernández-Val and Melly, 2013), which is relevant as most women in the sample do not work. Second, distribution regression allows the response to vary with 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.

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

$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]

Overall, the results in this section indicate that the Australian reform had a smaller impact on labor supply than similar reforms in other countries (e.g., Staubli and Zweimüller, 2013; Vestad, 2013; Cribb, Emmerson and Tetlow, 2016; Manoli and Weber, 2016; Geyer et al., 2020). These relatively modest impacts are somewhat surprising, given (i) the strong incentive to start claiming the Age Pension at the APA and (ii) the disincentive to continue working beyond the APA provided by the income test. However, the modest impacts can also be rationalized, given (i) the low labor force participation rates among women near the APA, (ii) the very strong substitution to other government transfers, and (iii) the concentrated impact on poorer households, who had a particularly weak attachment to the labor market and a particularly high reliance on other government transfers.

6 Distributional effects on household incomes

In this section, I examine the overall distributional effects of the reform. In Section 6.1, I explain the key variable used for analysis: household disposable income net of 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 I quantify the impact on inequality measures. In Section 6.3, I estimate the effects on relative poverty rates. 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

My preferred measure of household income is household disposable income net of housing costs and the value of in-kind benefits. Considering in-kind benefits is important as many Australians become eligible for higher benefits at the APA, and these benefits are disproportionately received by low-income households (Harmer, 2008). Considering housing costs is important for the mea-

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

surement of relative poverty. Relative poverty lines are typically set equal to 50% or 60% of the median income in the entire population. However, older Australians generally have relatively low housing costs 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, so accounting for this variation results in a more sensible comparison of these groups.

HILDA contains comprehensive information on 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 it 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 other costs, such as council rates, that are only incurred by homeowners.[21]

Unfortunately, HILDA lacks information on in-kind benefits. I address this issue by using additional data to construct an imputed value of in-kind benefits for each household in the HILDA sample. Specifically, I use the 2003–04 and 2009–10 cross-sections of the Australian Bureau of Statistics' Fiscal Incidence Study, a representative survey that estimates the value of in-kind benefits for individual households.[22] As explained in Appendix A.2, I restrict the sample to households containing women of the relevant age and calculate the average value of in-kind benefits for different types of households. Specifically, I divide households into groups based on their composition (single/couple) and concession card ownership, which increases sharply for women at the APA and is positively related to the value of in-kind benefits.[23] Then, I set the value of in-kind benefits for each household in HILDA equal to the average of households in the same group.

[21]Specifically, I increase housing costs for homeowners by $2,474: the mean annual cost of housing (net of mortgage repayments) for homeowners in the 2015–16 Australian Bureau of Statistics' Income and Housing Costs Survey.

[22]The Fiscal Incidence Study is conducted every six years and estimates the value of in-kind benefits to households from 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, (ii) the reported expenditure of the household on a comprehensive list of items asked in the accompanying Household Expenditure Survey, and (iii) the various concessions available to that household.

[23]In Appendix A.3, I estimate that the reform decreased concession card ownership among women by 44.6 percentage points ($p < 0.01$). The average value of in-kind benefits is 66% higher for single women who own a concession card than those who do not (see Table A7), 50% higher for couples who have one concession card owner, and 90% higher for couples who have two concession card owners.

After this imputation, I construct the key variable, adjusted household income, as household disposable income *plus* the value of in-kind benefits *less* housing costs. I adjust for household size by dividing by 1.5 for couples, the equivalence scale implied by Age Pension payment levels.

Figure A8 shows how in-kind benefits and housing costs affect the income distribution among households in the sample. Evidently, both adjustments have stronger impacts towards the bottom of the distribution but in opposing directions. Due to the stronger impact of in-kind benefits, the net effect is a reduction in inequality measures. For example, the Gini coefficient — a standard inequality measure ranging from zero to one — is 0.433 for household disposable income and 0.356 for adjusted household income. Overall, these measures indicate that there is a higher level of inequality among households in the sample than in the Australian population overall, in which the Gini coefficient for household disposable income is 25% lower (0.323).

6.2 Effect on inequality measures

In this section, I estimate the impact of the reform on inequality measures. I start with a brief overview of my four-step approach. *Step 1*: I use distribution regressions to estimate the impact on adjusted household incomes across the entire distribution. *Step 2*: I use the estimates from *Step 1* to estimate the counterfactual distribution of household income for treated women. *Step 3*: I compare the observed and counterfactual Lorenz curves for treated women and calculate the difference in inequality measures. *Step 4*: I bootstrap *Steps 1–3* to construct p-values for the estimates.

To implement the distribution regressions (*Step 1*), I modify equation (1). In each regression, the dependent variable is an indicator for adjusted household income being *less* than \$$x$, with x set at \$2,500 intervals up to one million dollars. As the dependent variable contains information for the financial year prior to the survey date (rather than at the survey date), I replace 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.[24]

[24] I make similar adjustments to other right-hand-side variables: I specify the age fixed effects with respect to women's age at the start of the previous financial year (rather than her age at the survey date), and I control for the average state-level unemployment rate in the previous financial year (rather than in the survey month).

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 (as expected).

The estimates indicate that the reform had a strong negative impact on the distribution for household incomes below \$45,000 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 \$42,500 (corresponding to the 7$^{\text{th}}$ and 43$^{\text{rd}}$ percentiles among treated women), indicating that the reform increased the proportion of households with incomes below these relatively modest cut-offs. In contrast, the estimates are close to zero and statistically insignificant at higher cut-offs. Overall, these effects align with the theoretical predictions, in that there is a 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 constructing the empirical distribution function for treated women at each value of x in the income grid (the "treated distribution"). Here, 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.[25] 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 construct the respective Lorenz curves for the treated and counterfactual distributions and compare standard measures of inequality. Figure 7 presents the Lorenz curves. Evidently, the Lorenz curve for the treated distribution is consistently below the counterfactual. In fact, if we exclude incomes in the bottom 2% of each distribution, the counterfactual distribution has Lorenz dominance over the treated distribution (that is, its Lorenz curve is always closer to the 45-degree line of complete equality). This implies

[25] I also restrict the counterfactual distribution to be non-decreasing in income (see Appendix A.5). The impact of this restriction is minor.

that nearly all inequality measures will be higher for the treated distribution. Figure 7 shows that this is the case for standard measures like the Gini coefficient, which is 11% higher. Other inequality measures show larger effects as they are more sensitive to changes towards the bottom of the income distribution. For example, the Theil index (Theil's T) is 22% higher, the mean log deviation index is 36% higher, the ratio of the income share of the top income quintile to that of the bottom 40% (the "80/40 ratio") is 18% higher, and the 80/20 ratio is 20% higher. All of these increases are economically significant. An 11% increase in the Gini coefficient is equivalent to an annual transfer of $7,000 from households in the bottom quintile of the sample to households in the top quintile, while a 20% increase in the 80/20 ratio is equivalent to a transfer of $3,800. These transfers amount to 19–35% of the average income of households in the bottom quintile.

To construct p-values (*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 inequality measure. 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 0.050 for the Gini coefficient, 0.172 for the Theil index, 0.026 for the mean log deviation index, 0.028 for the 80/40 ratio and 0.036 for the 80/20 ratio. Hence, four of the five estimates are either significant at the 5% level or bordering significance, and all are significant at the 10% level under a one-sided test.

Added value of the approach. The distribution regression approach is crucial in detecting the regressive impact of the reform because it allows for within-group income changes in addition to between-group changes. In Panel A of Table A9, I present the average impacts on adjusted household income for the full sample and various subgroups. For the full sample, the estimated effect is a decrease of $2,025 per annum ($p \geq 0.1$), and while the estimates show larger and at times statistically significant effects for lower income groups (single and poorer households), the differences between groups are at best significant at the 10% level. These estimates struggle to detect the regressive impacts because they only capture between-group income changes, which are less important in this context. In Panel B, I decompose the estimated increase in inequality measures into changes within and between groups, drawing on the fact that the Theil and mean log deviation indices can be decomposed into within- and between-group components. For both measures, over 80% of the increase in inequality is explained by within-group changes. This shows the importance of going beyond mean estimates when the impacts and responses are likely to vary across the income distribution, the point emphasised by Bitler, Gelbach and Hoynes (2006).

6.3 Effect on relative poverty rates

In this section, I quantify the impact on relative poverty rates. For developed countries like 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 fraction of the median household income in the entire population. Namely, as explained in Appendix A.4, I set the poverty line in each year based on the median adjusted household income in Australia.

I estimate similar regressions to the distribution regressions described above in Section 6.2. Here, though, the dependent variables are indicators for adjusted household income being less than 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 80 (at intervals of 10). Figure 8 presents the estimates (with 95% confidence intervals), showing two patterns of note. First, 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 50% 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 meaningful increases in relative poverty rates in percentage terms (11–56%).

Figure 8 also 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 3.0 percentage points (37%, $p = 0.041$). At the 60% poverty line, the estimated increase in poverty is similar — 3.1 percentage points (20%) — but marginally insignificant at the 10% level ($p = 0.107$). Therefore, under conventional definitions of relative poverty, the estimated increase in poverty rates is around 3 percentage points or 20–37%.

These effects are concentrated among single women and renters (see Figure A9). The estimates are larger for these groups at nearly every poverty line. For example, under the OECD definition of relative poverty, the estimated increase in poverty is 6.9 percentage points for single women ($p = 0.023$) and 10.8 percentage points for renters ($p = 0.054$). The corresponding estimates for partnered women and homeowners are 0.6 and 1.8 percentage points (both $p \geq 0.1$). Concerningly, there is also a large increase in the proportion of renters with household incomes below 20% of the median income (10.9 percentage points, $p < 0.01$) and 30% of the median income (8.5 percentage points, $p = 0.013$). These effects reflect the higher housing costs of renters and suggest that the reform pushed some renters deeper into poverty.

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 in-kind benefits, since the Fiscal Incidence Study (FIS) may overestimate or underestimate their value. On the one hand, the FIS 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 FIS may overestimate the value of in-kind benefits. On the other hand, the FIS ignores the value of in-kind benefits as a form of insurance against health shocks, which may be large for older households.

In Table A10, I present the estimated effects on poverty and inequality measures for different values of in-kind benefits. Column 2 presents the baseline estimates, in which in-kind benefits are valued in the same way as the FIS (i.e., \$1 spent by governments is worth \$1 to households). Column 1 presents the estimates for higher values of in-kind benefits (125% of expenditure), while columns 3–6 present the estimates for lower values (0–75% of expenditure). Despite variation across columns in effect size and statistical significance, 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 is 5–13% with the Gini coefficient, 9–25% with Theil's T, 14–36% with the mean log deviation index, 14–18% with the 80/40 ratio, and 17–22% with the 80/20 ratio. The estimated increase in relative poverty rates is 2.0–4.5 percentage points at the 50% poverty line and 1.8–6.0 percentage points at the 60% poverty line.

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 by the size of the affected population. Using equation (1), I estimate the average impact 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 imputed values constructed in Section 6.1); and (iv) the total amount of income tax paid by women.[26]

Table 4 presents the estimates. The estimates in columns 1 and 2 replicate Table 3, showing that treated women lost \$6,957 per annum in income from the Age Pension but gained \$4,365 in income from other transfers. Column 3 presents the estimated effect on the amount of in-kind benefits provided to households; the estimate indicates that treated households received \$2,918 less on average in in-kind benefits per annum ($p < 0.01$). In column 4, I present the estimated effect on income tax. 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 increase in earnings. However, as there are tax concessions for people above the APA, women would have paid more tax even without any change in earnings.[28]

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

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

[28]The main tax concession is the Seniors and Pensioners Tax Offset (SAPTO). The SAPTO increases the tax-free

Finally, in column 5, I calculate the average net fiscal impact on treated women. This estimate, of -\$6,511 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 750 million dollars per annum. This saving is equivalent to 1.7% of current government expenditure on cash transfers to Age Pensioners (or 0.2% of total expenditure).

8 Conclusion

I examine the effects of a large Australian reform that increased women's eligibility age for the public retirement pension from 60 to 65. Overall, I find that this reform led to significant reductions in government expenditure and modest increases in female labor supply. However, the impacts and responses were concentrated among single women and poorer households. Moreover, the reform reduced the incomes of low-to-middle-income households but had little impact on households in the top half of the income distribution. These unequal impacts meant that the reform was significantly regressive; my baseline estimates indicate that the reform increased relative poverty rates by around 3 percentage points (20–37%), and inequality measures by 11–36%. These findings highlight a clear trade-off: 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 if women were unable to claim other transfers prior to the pension age. However,

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.

as low-income households disproportionately relied on such transfers, the estimated increase in inequality and poverty measures would have been around three- and seven-times larger respectively. While these estimates ignore behavioral responses, they suggest that governments should exhibit caution when considering simultaneous policies that tighten eligibility for other transfers. While such responses may help governments maximize the fiscal savings of higher pension-eligibility ages, these policies are likely to magnify the effects on poverty and inequality.

The results have clear international relevance. In response to widespread population aging, policymakers are raising pension-eligibility ages in many countries, including the U.K., Germany, France, Italy and Spain. When considering the international 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 progressive — that is, the replacement rate is higher for people with lower earnings. Hence, similar reforms may be 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 exist in other countries, these effects are especially large in the Australian context due to the explicit redistributional aims of its welfare system. Thus, the estimated effects on poverty and inequality measures may not be 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 these reforms, while fiscally effective, can significantly diminish the social safety net for older households, especially women who are single and renting. This highlights a potential role for complementary policies like targeted housing assistance that strengthen the safety net for vulnerable 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. 2019. "Pension incentives and the joint retirement of couples: Evidence from two natural experiments." *Journal of Population Economics*, 32(3): 735–767.

Autor, David, Andreas Kostøl, Magne Mogstad, and Bradley Setzler. 2019. "Disability benefits, consumption insurance, and household labor supply." *American Economic Review*, 109(7): 2613–54.

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.

Borghans, Lex, Anne C. Gielen, and Erzo F.P. Luttmer. 2014. "Social support substitution and the earnings rebound: Evidence from a regression discontinuity in disability insurance reform." *American Economic Journal: Economic Policy*, 6(4): 34–70.

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." Michigan Retirement Research Center Research Paper No. 2014-314.

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.

Chomik, Rafal, Simon Graham, Hazel Bateman, John Piggott, and Sophie Yan. 2018. "Retirement income in Australia: Part II Public support." The ARC Centre of Excellence in Population Ageing Research (CEPAR).

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.” NBER Working Paper No. 10466.

Engels, Barbara, Johannes Geyer, and Peter Haan. 2017. “Pension incentives and early retirement.” *Labour Economics*, 47: 216–231.

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. 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.” mimeo.

Geyer, Johannes, and Clara Welteke. 2019. “Closing Routes to Retirement for Women: How Do They Respond?” *Journal of Human Resources*.

Geyer, Johannes, Peter Haan, Anna Hammerschmid, and Michael Peters. 2020. “Labor market and distributional effects of an increase in the retirement age.” *Labour Economics*.

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).

Inderbitzin, Lukas, Stefan Staubli, and Josef Zweimüller. 2016. “Extended unemployment benefits and early retirement: Program complementarity and program substitution.” *American Economic Journal: Economic Policy*, 8(1): 253–288.

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.” NBER Working Paper orrc17-12.

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." NBER Working Paper No. 22561.

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." SSRN Working Paper No. 3382046.

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

Oguzoglu, Umut, Cain Polidano, and Ha Vu. 2020. "Impacts from delaying access to retirement benefits on welfare receipt and expenditure: Evidence from a natural experiment." *Economic Record*, 96(312): 65–86.

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." RAND Report.

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." In *Changing poverty, changing policies*, 203–241. New York: Russell Sage Foundation.

Selin, Håkan. 2017. "What happens to the husband's retirement decision when the wife's retirement incentives change?" *International Tax and Public Finance*, 24(3): 432–458.

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. 2011. "The impact of stricter criteria for disability insurance on labor force participation." *Journal of Public Economics*, 95(9-10): 1223–1235.

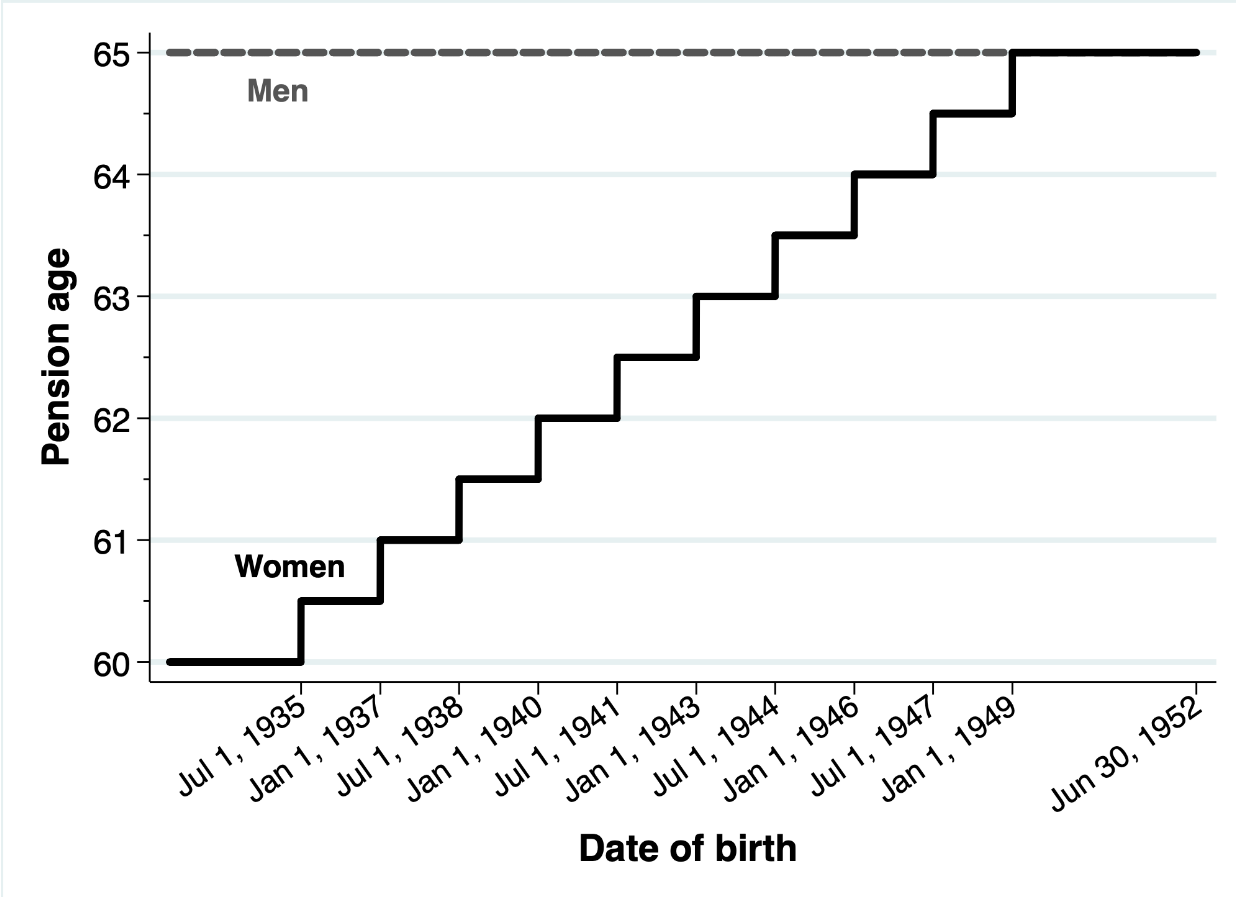
Staubli, Stefan, and Josef Zweimüller. 2013. "Does raising the early retirement age increase employment of older workers?" *Journal of Public Economics*, 108: 17–32.

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.

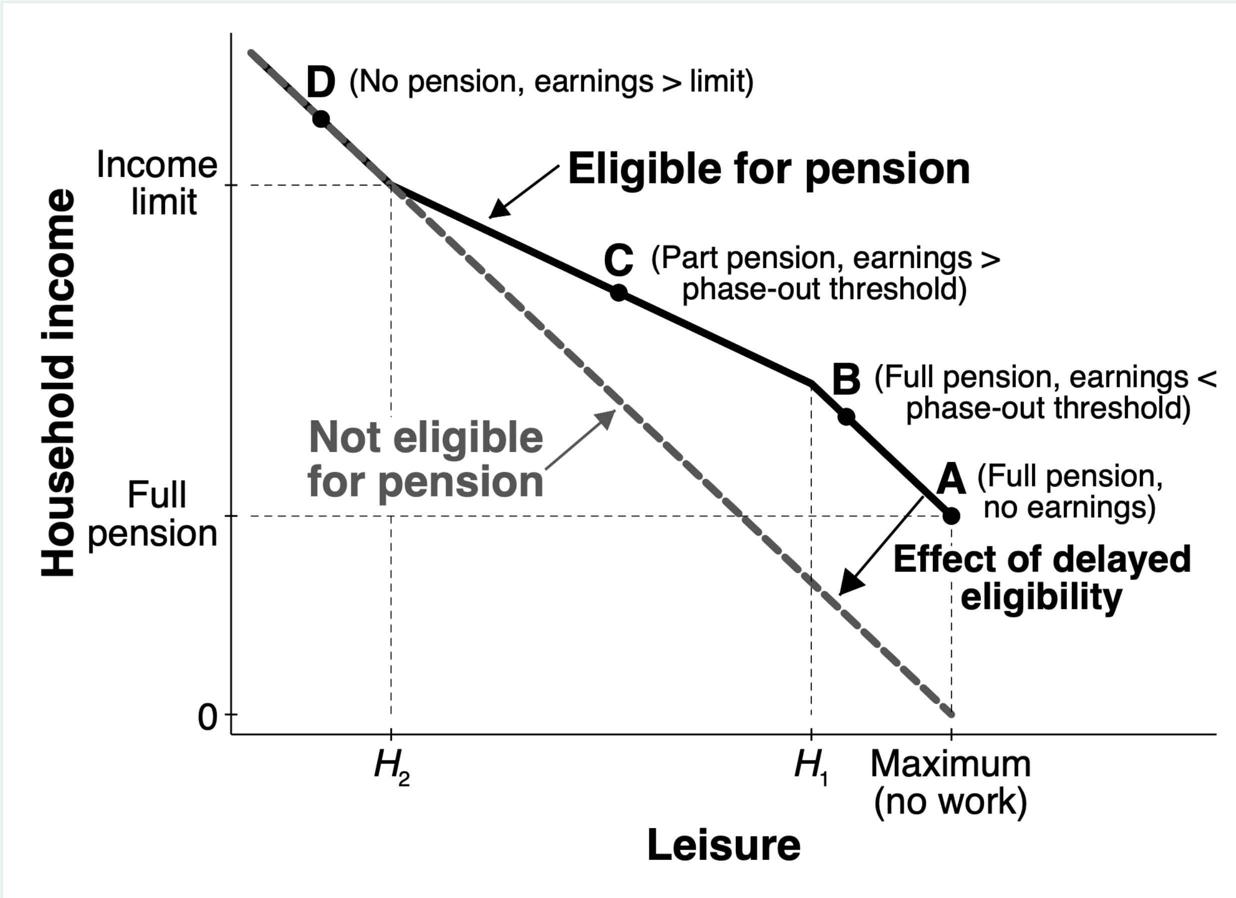
Watson, Nicole, and Mark P. Wooden. 2012. "The HILDA Survey: A case study in the design and development of a successful household panel survey." *Longitudinal and Life Course Studies*, 3(3): 369–381.

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



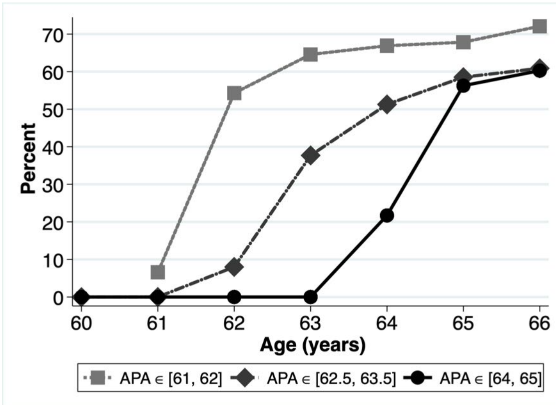
Summary of effects on each type of household

	Direct effect on		Expected effect on	
Location when eligible (1)	Household income (2)	Effective wage (3)	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

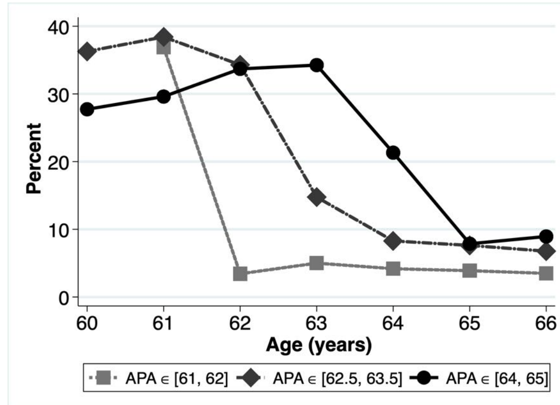
Notes: This figure shows the effect of delaying women from reaching the pension age on their households' budget constraint, in terms of the set of feasible combinations of total income and leisure (assuming that households are eligible for the maximum payment under the assets test and ignoring other 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

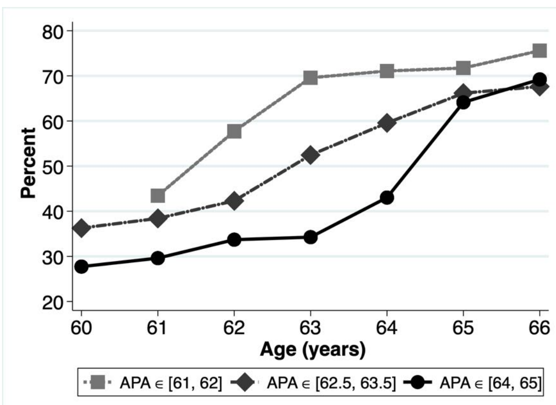
(a) Receives Age Pension



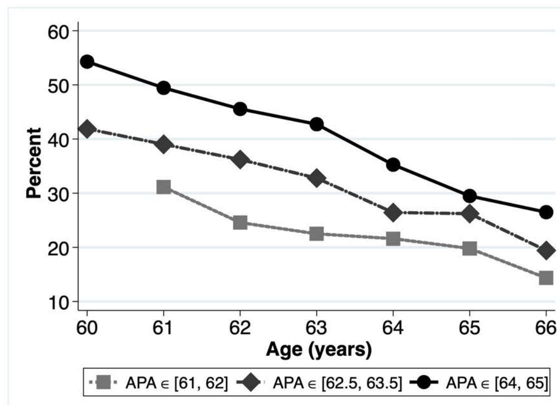
(b) Receives other government transfer



(c) Receives any government transfer



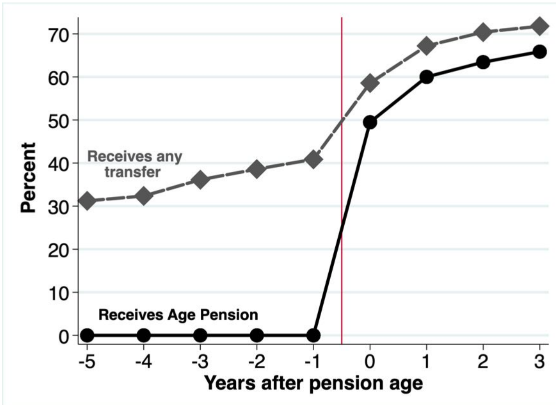
(d) In labor force



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.

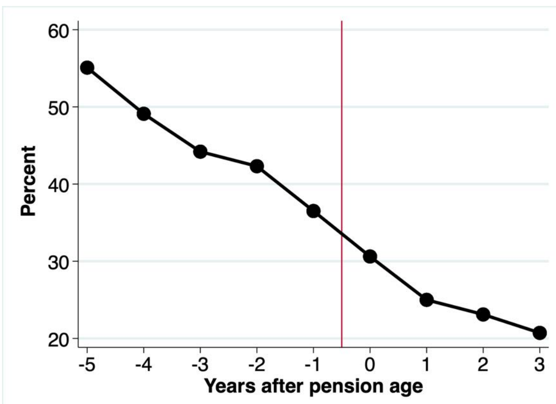
Figure 4: Mean outcomes of women by years from pension age

(a) Receipt of government transfers

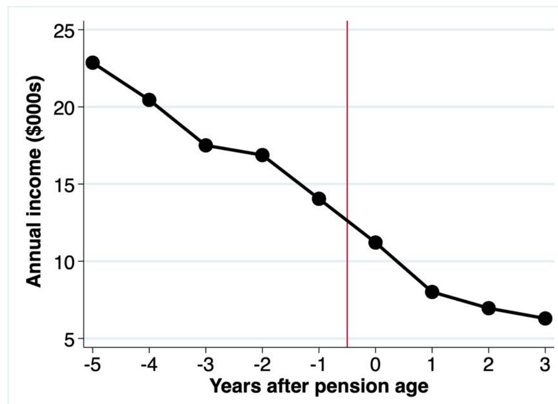


(b) Income from government transfers

(c) In labor force



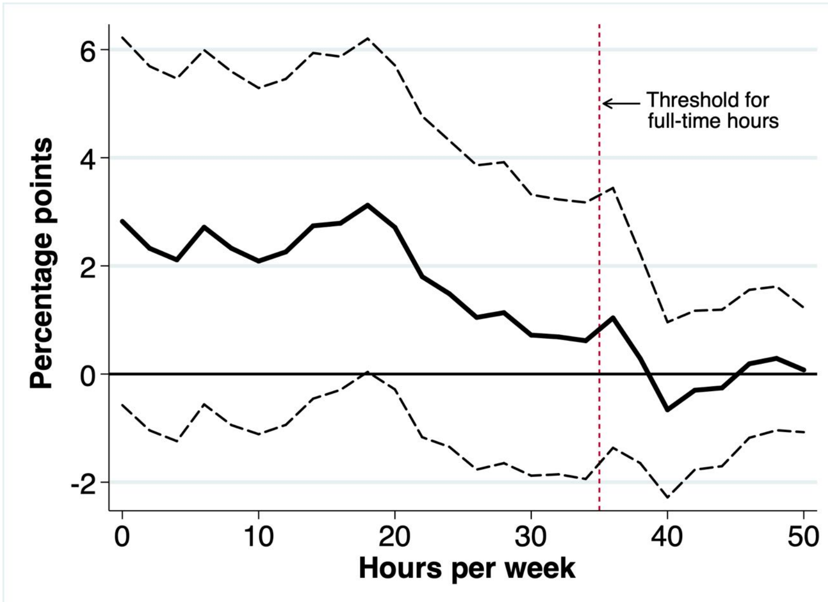
(d) Labor earnings



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

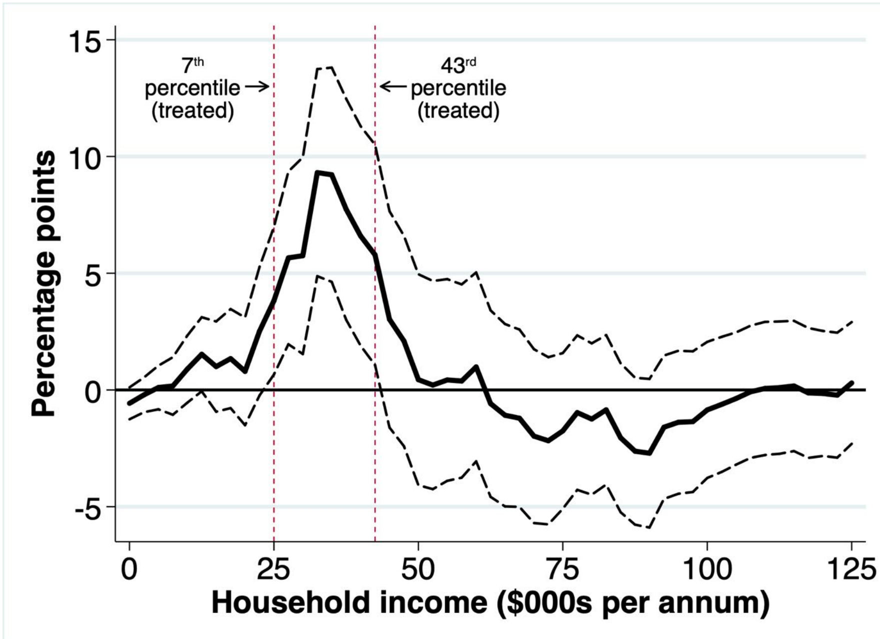


(b) Labor earnings > $*x* per annum

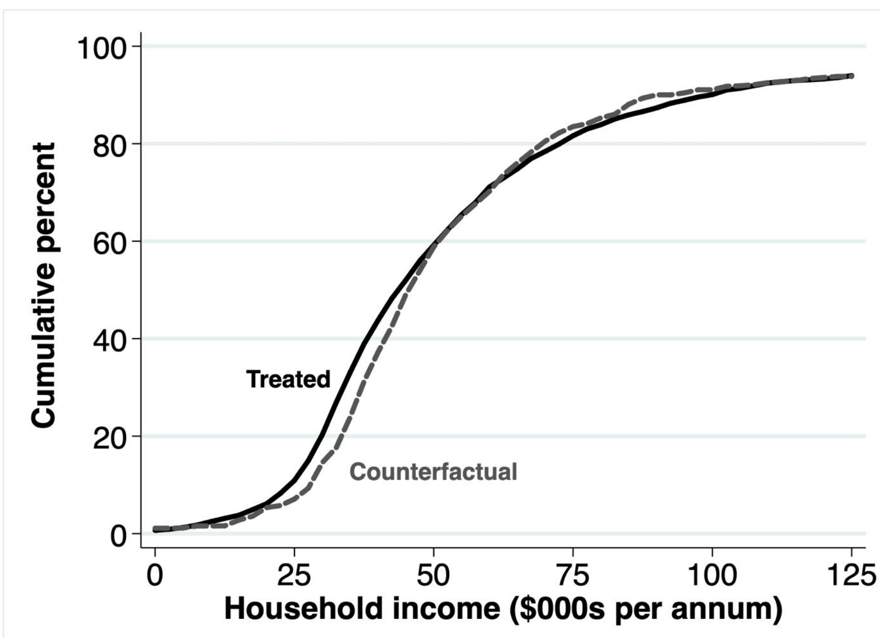
Notes: These figures present distribution regression estimates (estimated by OLS), with 95% confidence intervals, showing the effects of the reform on treated women's earnings (in 2016 Australian dollars) and hours worked at different points in the distribution. In each regression, the dependent variable 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. In (b), 'income limit' corresponds to the Age Pension income limit in 2016 for singles, and 'average full-time earnings' is based on 2016 estimates from the Australian Bureau of Statistics. The sample includes women aged 60–66 from waves 1–14 of the HILDA survey. See the text in Section 5.2 for more details.

Figure 6: Overall distributional effects on household incomes

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

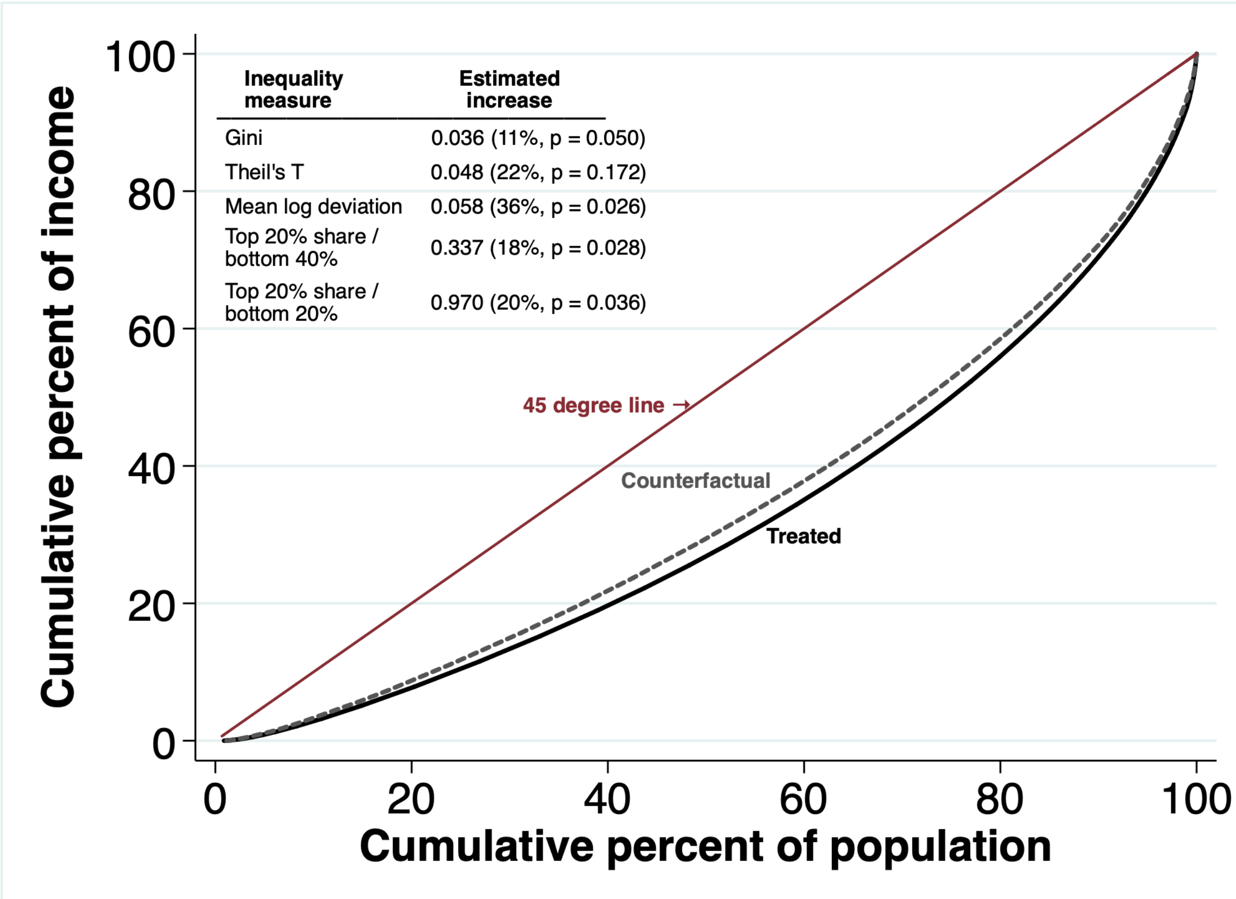


(b) Distribution of household income: Observed (treated women) and estimated counterfactual



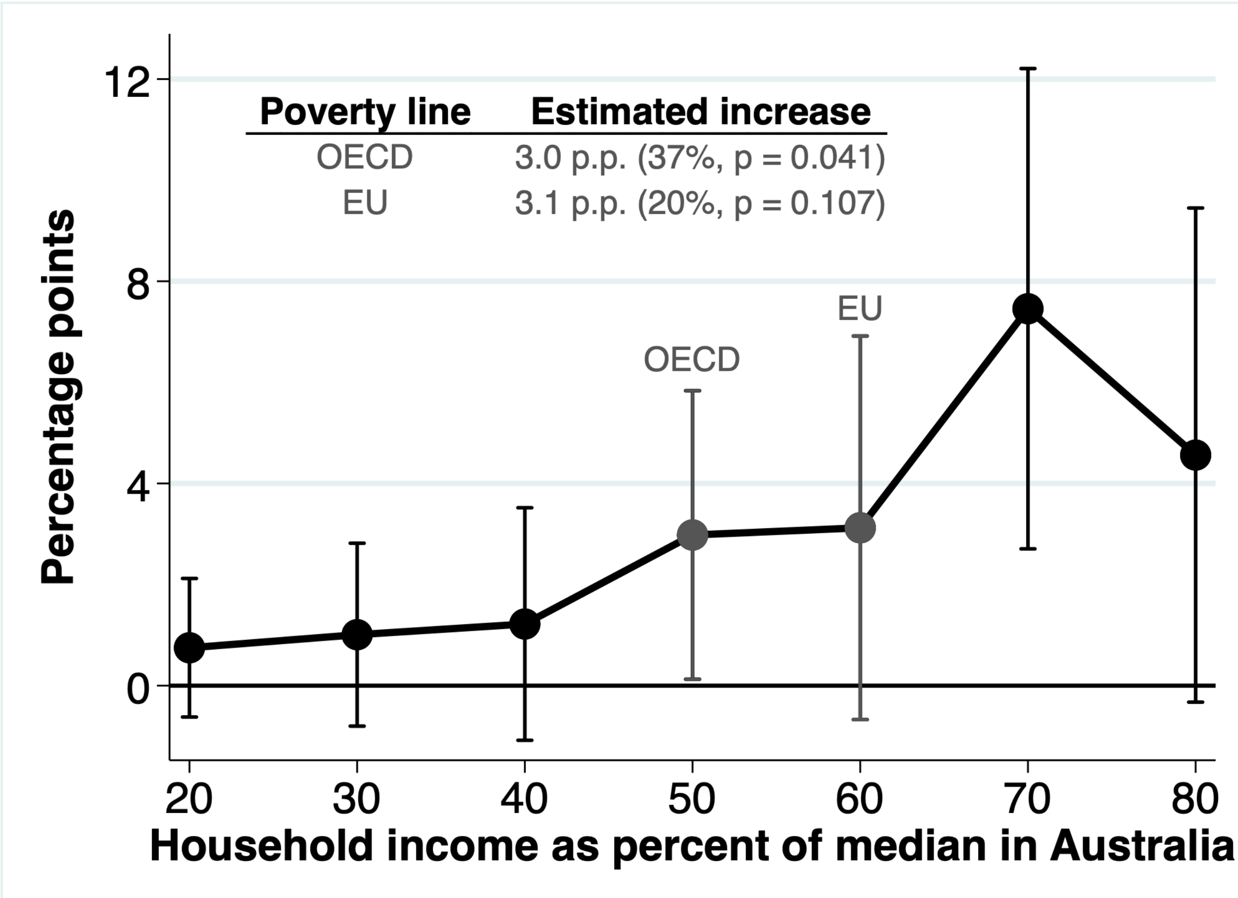
Notes: These figures show the distributional effects of the reform on household incomes (net of housing costs and in-kind benefits). Incomes are in 2016 Australian Dollars. (a) presents distribution regression estimates (estimated by OLS), with 95% confidence intervals, showing the estimated impact of 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). See the text in Section 6.2 for more details.

Figure 7: Observed and counterfactual Lorenz curves, and difference in inequality measures



Notes: This figure shows the Lorenz curves based on the observed distribution of household income (net of housing costs and in-kind benefits) for treated women and the estimated counterfactual. The Lorenz curves are constructed from the distributions shown in Figure 6b (using incomes up to $1,000,000 per annum). 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 8: Estimated effects on relative poverty: Household income less than a given percentage of the median 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 (net of housing costs and 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 80. 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)	(2.0)	(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)	(3,323)	(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***	-2,918***	1,001***	-$6,511
	(301)	(312)	(211)	(277)	
Mean at pension age	6,879	1,244	19,475	1,202	
R-squared	0.436	0.181	0.545	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.

For online publication

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

Contents

List of Figures

List of Tables

A Data appendix

A.1 Testing for anticipatory behavior

In this section, 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 follow the approach of Duggan, Singleton and Song (2007) and test whether women responded to changes in the net present value (NPV) of their future Age Pension income. Specifically, among women in the sample with an Age Pension Age (APA) of 65.0 and below,[A1] I construct the change in the NPV of future Age Pension income for women at each age. I allow the change in NPV to differ by women's APA, age and marital status (given the higher payment rates for single women). I use the following formula to construct the change in NPV for individual i at single year of age $a \in \{55,\dots,61\}$:

$$\Delta NPV_{ia} = \sum_{k=a}^{64} \left[\mathbf{1}(k \geq APA_i) + 0.5 \times \mathbf{1}(k+0.5 = APA_i) - \mathbf{1}(k \geq 60)\right] S_{k,a}(1+r)^{a-k} y_{MS,i} \qquad \text{(A1)}$$

where the term in square brackets is the relative change in pension income for individual i at age k compared to an APA of 60 (equal to 0, -0.5 or -1); $S_{k,a}$ is the conditional probability of a woman living to age k given that she is alive at age a; r is the interest rate used to discount future income; and $y_{MS,i}$ is the average annual pension income (in \$000's) of women above the APA of the same marital status as woman i.[A2]

After calculating ΔNPV_{ia} for each individual at each age, I estimate the following regressions, which replicate equation (3) in Duggan, Singleton and Song (2007):

$$y_{iat} = \beta x_{iat} + \gamma \Delta NPV_{ia} + FE\_age\_yrs_a + FE\_wave_t + \varepsilon_{iat} \qquad \text{(A2)}$$

[A1] 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.

[A2] To construct $S_{k,a}$, I use age-specific death rates for Australian women in 2014 from the Australian Bureau of Statistics. To construct $y_{MS,i}$, I take the mean income from the Age Pension for single and married women in the sample who are above the APA (among recipients and non-recipients).

where the notation and covariates are the same as in equation (1) except $FE\_age\_yrs_a$ are fixed effects for age in single years and FE_wave_t are fixed effects for each survey wave to control for time-specific factors affecting y_{iat}.

The key explanatory variable is ΔNPV_{ia}, the change in the net present value of woman i's future Age Pension income at age a divided by $1,000. Table A3 presents the estimates of γ under different assumptions. In the baseline estimates, I use an interest rate of 4% to discount future income and assume that women remain in their current marital status until age 65. All of the estimates are small and statistically insignificant. For example, the point estimate for earnings indicates that women earn an additional $42 per annum for each $1,000 increase in the NPV of Age Pension income. Similarly, the point estimate in column (5) indicates an increase in labor force participation of 0.001 percentage points for each $1,000 increase in the NPV of Age Pension income. To put these estimates in context, the average change in NPV for an increase in the APA of 0.5 years is -$3,600. Thus, the estimated anticipatory response to a marginal increase in the APA by 0.5 years is a decrease in earnings of $151 per annum ($\sigma = \$1,289$) and a decrease in LFP of 0.0036 percentage points ($\sigma = 1.89$ percentage points). Overall, these estimates are very small, highly statistically insignificant and similar in precision to the baseline estimates in Tables 2 and 3.

Table A3 also shows the robustness of these estimates along two dimensions. First, I vary the interest rate used in equation (A1) to discount future income and re-estimate equation (A2). The estimates and standard errors remain similar to the baseline estimates with higher and lower interest rates ($r = 6\%$ and $r = 2\%$). Second, I calculate the change in NPV under different assumptions about future marital status. For example, individuals who are ever single at ages 55–61 may anticipate this and expect to remain single until age 65. Similarly, individuals who are ever partnered at ages 55–61 may expect to remain partnered until 65, either by maintaining their current relationship or by re-partnering. Under both of these extreme assumptions, the estimates remain statistically insignificant.

A.2 Imputing the value of in-kind benefits to households in HILDA

I construct an imputed value of in-kind benefits for each household in the HILDA sample using the 2003–04 and 2009–10 cross-sections of the Australian Bureau of Statistics' Fiscal Incidence Study. Specifically, I pool the 2003–04 and 2009–10 surveys and restrict the sample to households containing women aged 60–64.[A3] Then, among these households, I calculate the average total value of in-kind benefits for the following types of households (in 2016 dollars): (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.

Table A7 shows the average value of in-kind benefits for each group. As expected, this value is considerably higher for concession card owners. To allocate households 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.3). Specifically, I impute the fraction of time respondents owned a concession card in the previous financial year. Then, I calculate the average number of concession cards owned by each household in the sample over the previous financial year. This is equal to the fraction of time women owned a concession card in the previous financial year plus the fraction for their spouse. Finally, I construct the total value of in-kind benefits for each household in HILDA using the values in Table A7, with linear interpolation 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 $0.5 \times \$9,365 + 0.5 \times \$15,559 = \$12,462$.

[A3] Age is only measured in five-year intervals, so it is not possible to restrict the sample to women aged 60–66 (to match the HILDA sample).

A.3 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 (either a pensioner card or seniors card) during this period. To estimate the fraction of time respondents owned a pensioner card, I first add up the number of weeks they received a transfer that directly qualified them for a pensioner card and divide by 52.[A4] Then, I add to this the fraction of weeks they received other transfers that would qualify them for a pensioner card if they were over the age of 60, multiplied by the fraction of time they were over 60.[A5] Finally, I restrict the fraction of time respondents owned a pensioner card to be no larger than 1.

To estimate the fraction of time respondents owned a seniors card, 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 pensioner card and (ii) their equivalised household disposable income was below the 90$^{\text{th}}$ percentile in the sample. Otherwise, I set this 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 pensioner card and the fraction of time they owned a seniors card. 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 of HILDA, in which information on concession cards 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%.[A6] Moreover,

[A4] 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.

[A5] These transfers include Newstart Allowance, Parenting Payment Partnered, Partner Allowance, Widow Allowance, Sickness Allowance and Special Benefit.

[A6] This small difference makes sense because concession card ownership increases with age, and women are the imputation aligns with women's reports for 88.0% of observations.

Table A8 shows the estimated effect of the reform on concession card ownership among women in the sample.[A7] The estimates show a decrease in concession card ownership of 44.6 percentage points ($p < 0.01$), indicating that the reform has considerably delayed concession card ownership among women. Over half of this effect stems from reduced access to the seniors card, with the rest stemming from reduced access to the pensioner card.

A.4 Calculating the median adjusted household income for the population

In order to define measures of relative poverty, I first calculate the median adjusted household income (household disposable income net of housing costs and the value of in-kind benefits) in the Australian population in the relevant year. I do so 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.[A8] 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. This assumption is supported by the data, with very similar ratios in the 2003–04 and 2015–16 surveys (1.139 and 1.157 respectively).

---

younger in the financial year prior to the survey date than when they report concession card ownership (at the survey date).

[A7] As the dependent variable for these regressions is based on information for the previous financial year rather than at the survey date, I make the same modifications to equation (1) as described in the second paragraph of Section 6.2: (i) 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; (ii) I specify the age fixed effects with respect to women's age at the start of the previous financial year (rather than her age at the survey date); and (iii) I control for the average state-level unemployment rate in the previous financial year (rather than in the survey month).

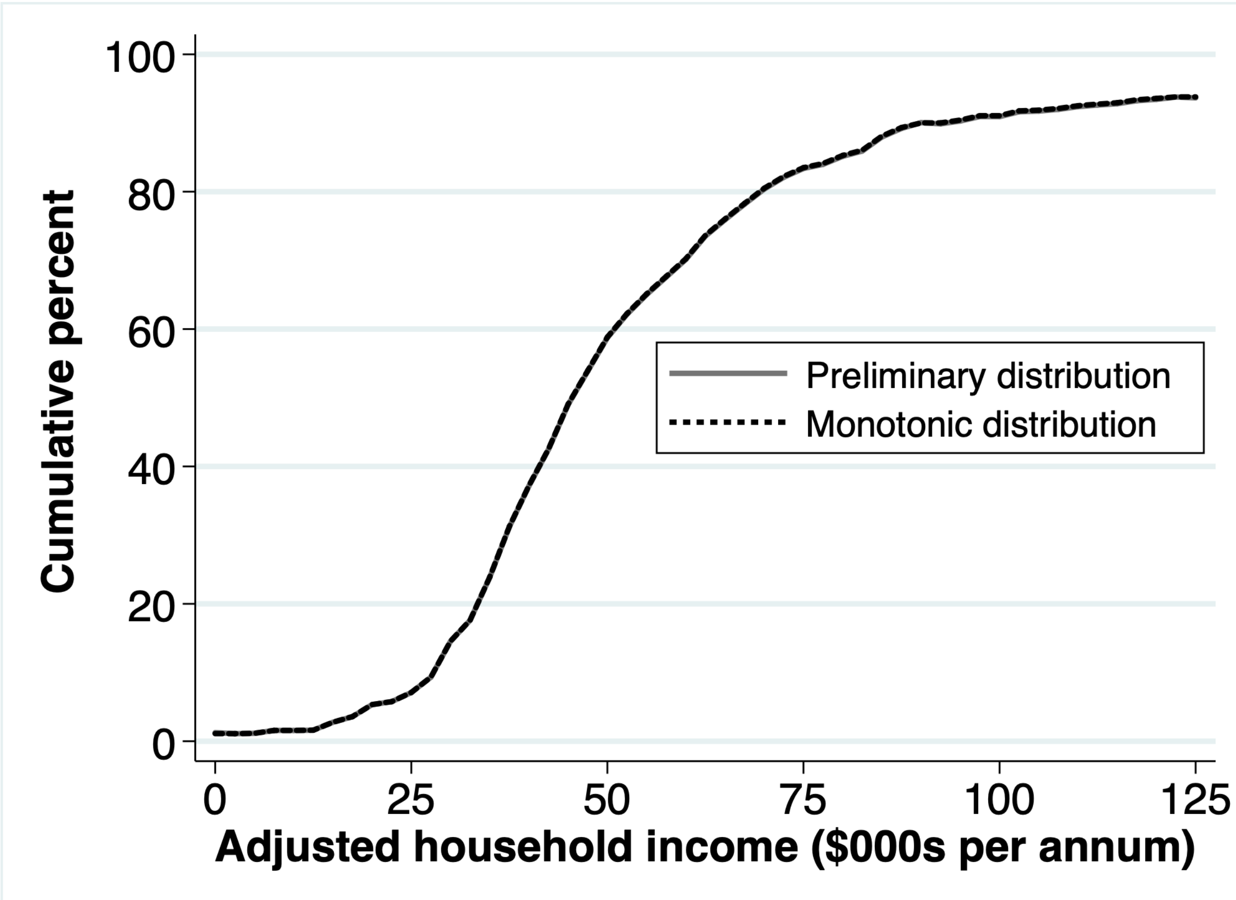
[A8] As these surveys are conducted every second year, I use linear interpolation to estimate the median in intermediate years.

A.5 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).[A9]

[A9] 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 more conservative estimates of the impact on inequality measures.

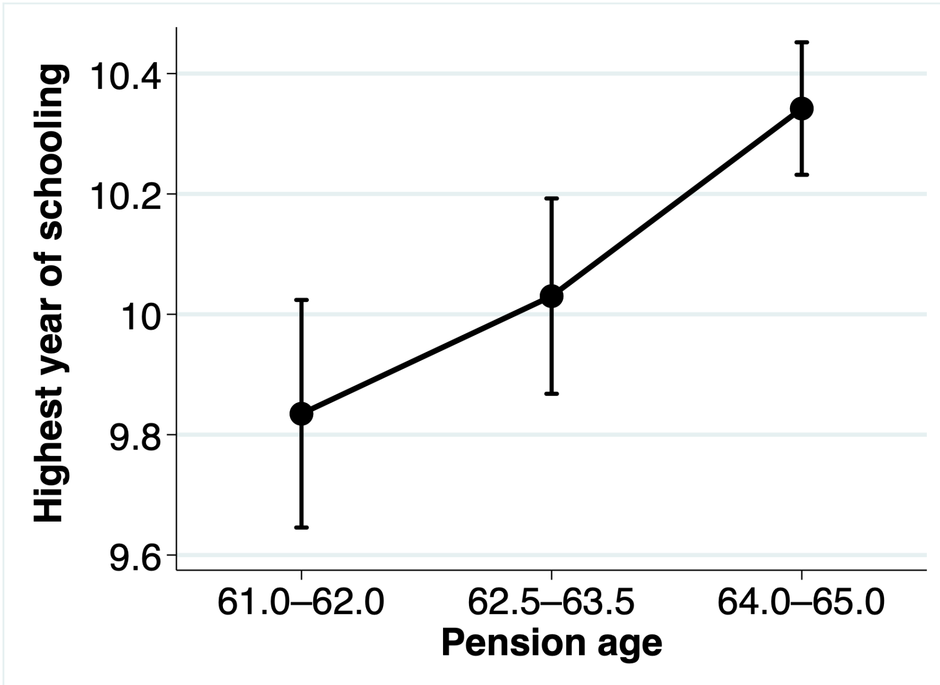
Figure A1: Impact of the monotonicity restriction on the counterfactual distribution



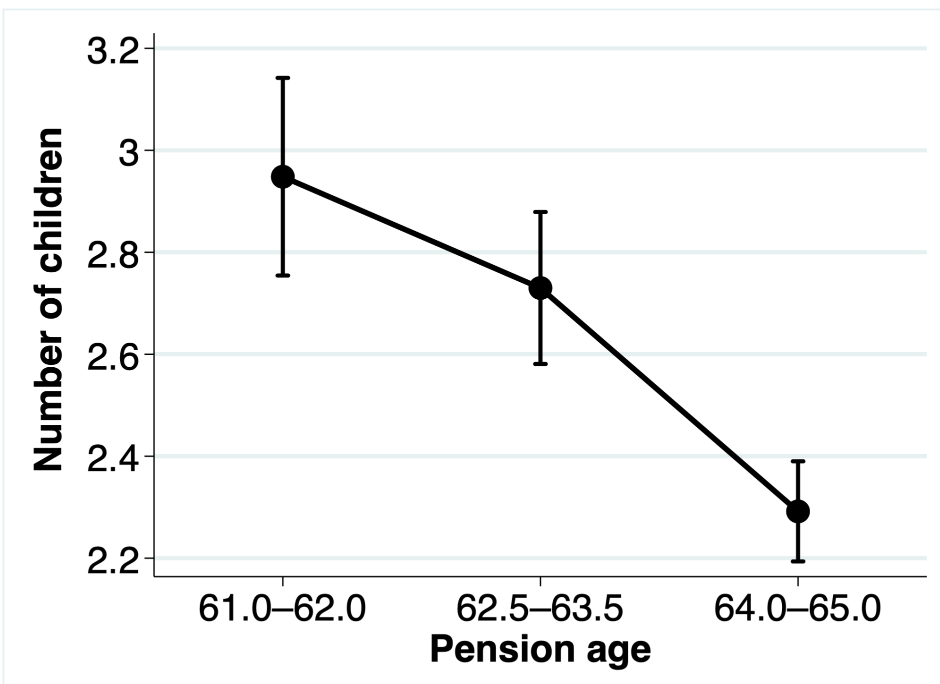
Notes: This figure shows the impact of the monotonicity restriction on the counterfactual distribution of household income (net of housing costs and in-kind benefits). Evidently, the impact is extremely minor (and visually imperceptible). See Appendix A.5 for more details on this restriction. The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey.

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

(a) Highest completed year of schooling



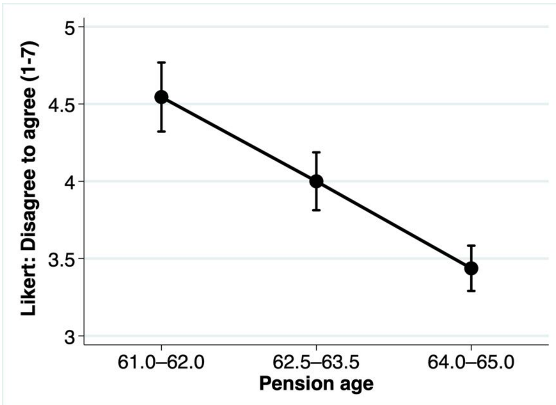
(b) Number of children



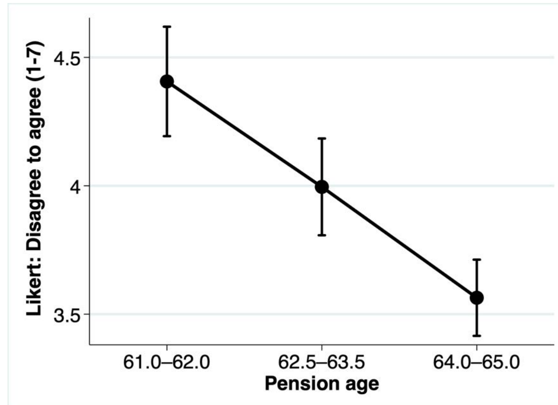
Notes: These figures show how average birth rates and education vary with women's Age Pension Age (APA), with 95% confidence intervals around these averages. These figures show that women with a higher APA are more educated and have fewer children on average. In (a), I exclude 1% of women who responded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

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

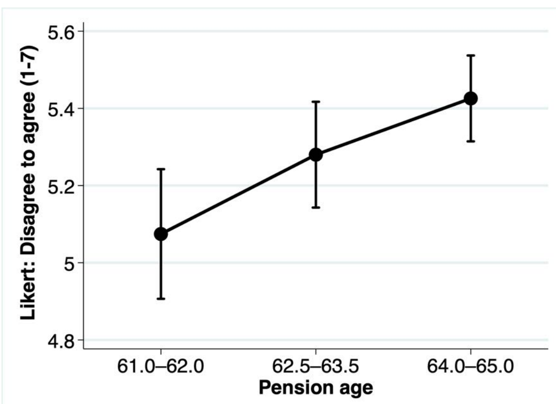
(a) 'Better for everyone involved if the man works and the woman takes care of home and children'



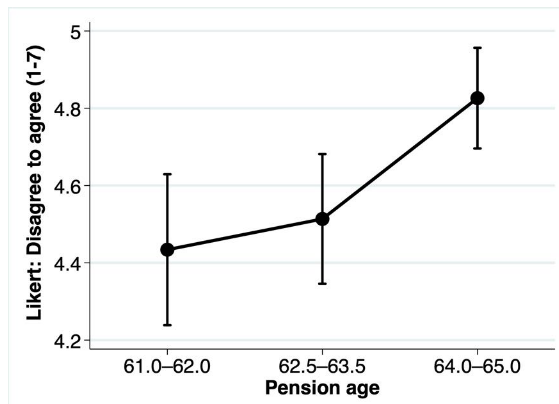
(b) 'Mothers who don't really need the money shouldn't work'



(c) 'Children do just as well if the mother works and the father cares of the home and children'



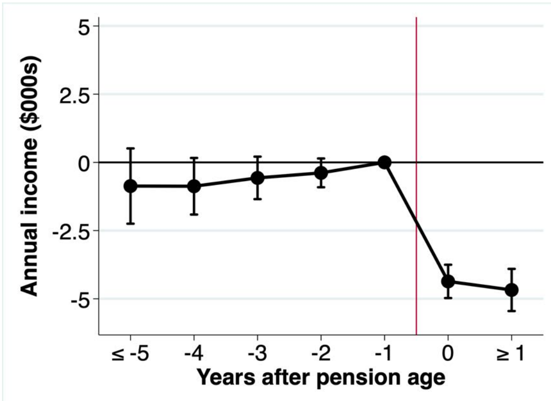
(d) 'Working mothers can establish equally good relationships with their children'



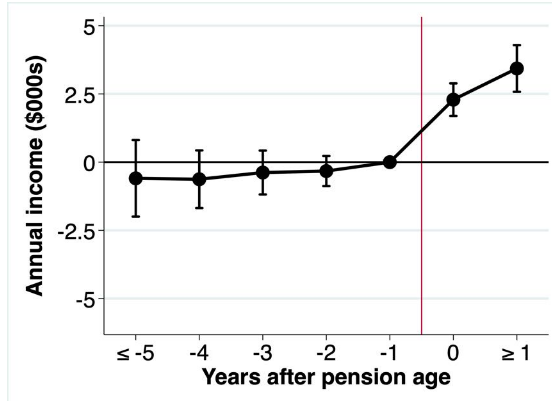
Notes: These figures show how the average attitudes of women vary with their Age Pension Age (APA), with 95% confidence intervals around these averages. The figures indicate that women with a higher APA had more progressive attitudes about how paid work and child-rearing should be shared within the household 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

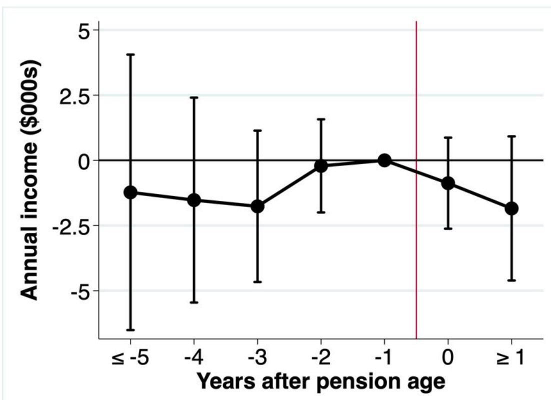
(a) Income from other transfers



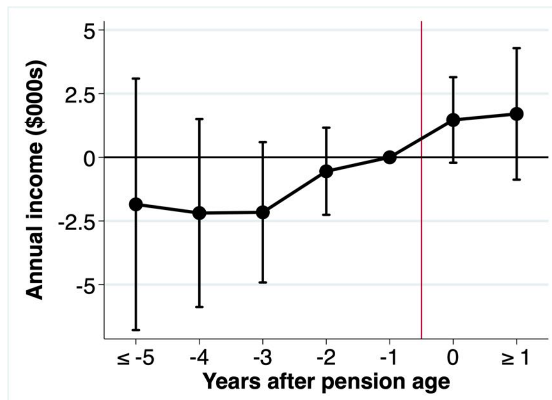
(b) Total income from transfers



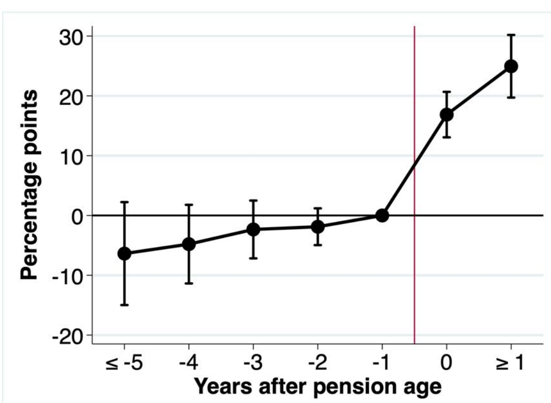
(c) Labor earnings



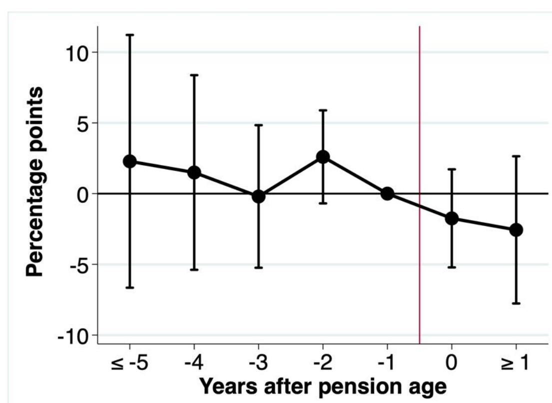
(d) Transfers plus earnings



(e) Receives any transfer



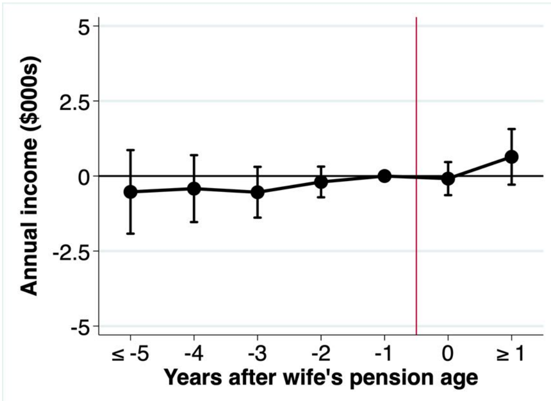
(f) In labor force



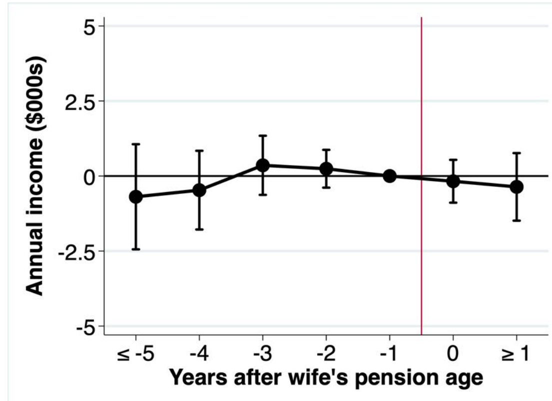
Notes: These figures plot the estimated δ coefficients, with 95% confidence intervals, from OLS estimates of equation (2). The estimates show the causal effect of women being a given number of years above or below the Age Pension Age (APA) relative to them being between zero and one year below the APA. Incomes are in 2016 Australian dollars. For these regressions, I expand the age sample to 55–69 (but only include women if they 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

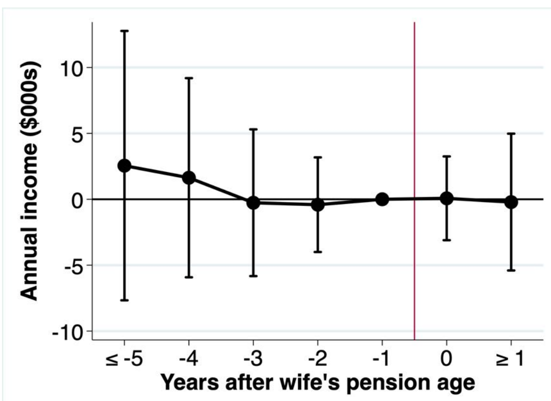
(a) Income from the Age Pension



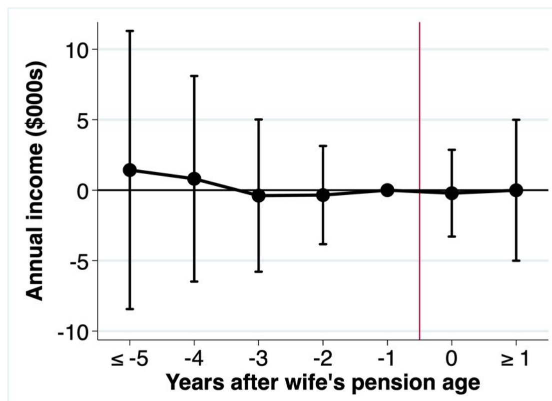
(b) Income from other transfers



(c) Labor earnings

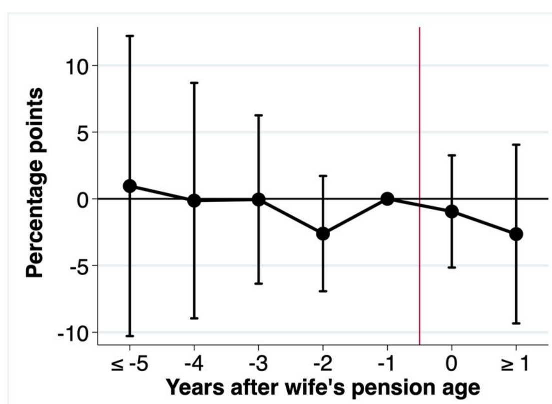


(d) Transfers plus earnings



(e) Receives any transfer

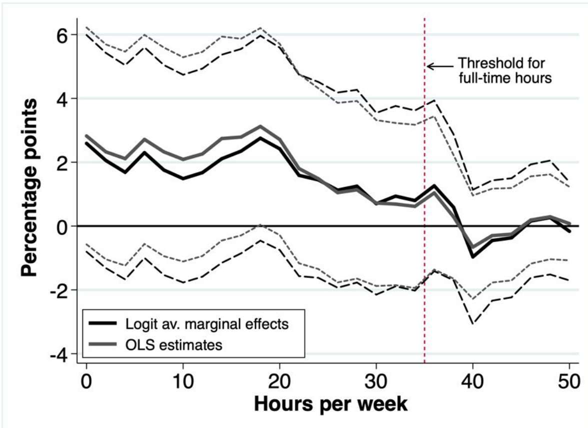
(f) In labor force



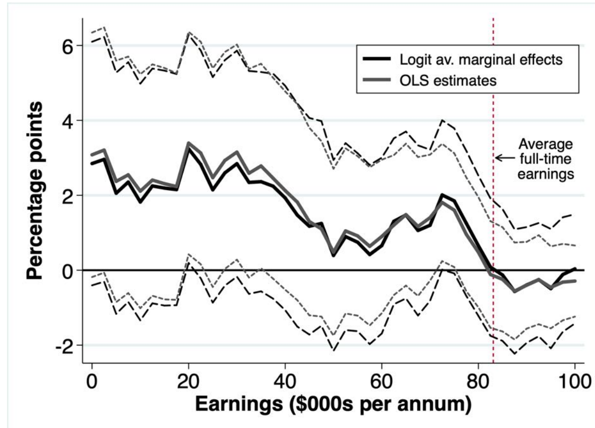
Notes: These figures plot the estimated δ coefficients, with 95% confidence intervals, from OLS estimates of equation (2). The estimates show the causal effect on spousal outcomes of women being a given number of years above or below the Age Pension Age (APA) relative to them being between zero and one 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

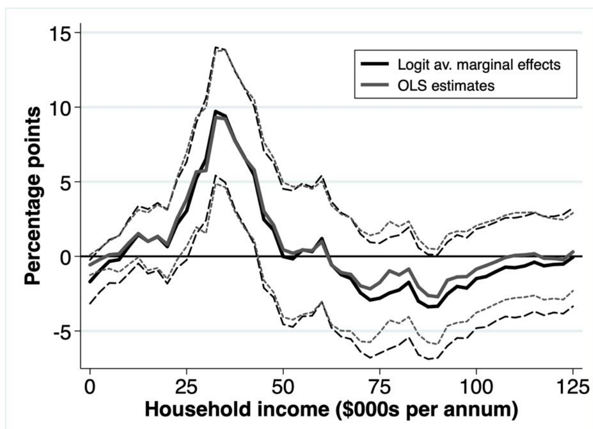
(**a**) Hours worked > *x* per week



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



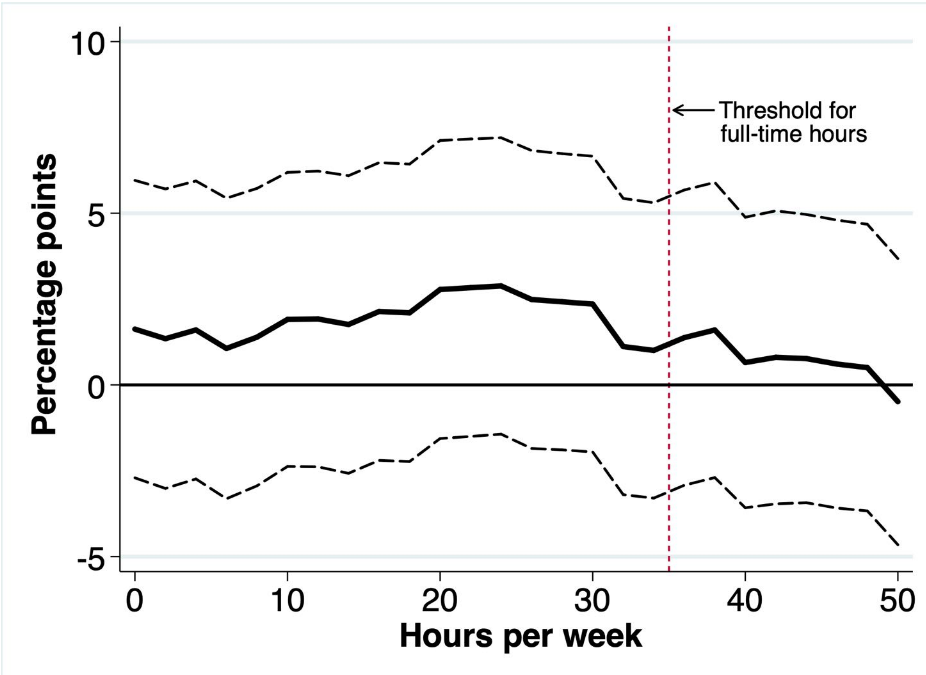
(**c**) Household income < \$*x* per annum



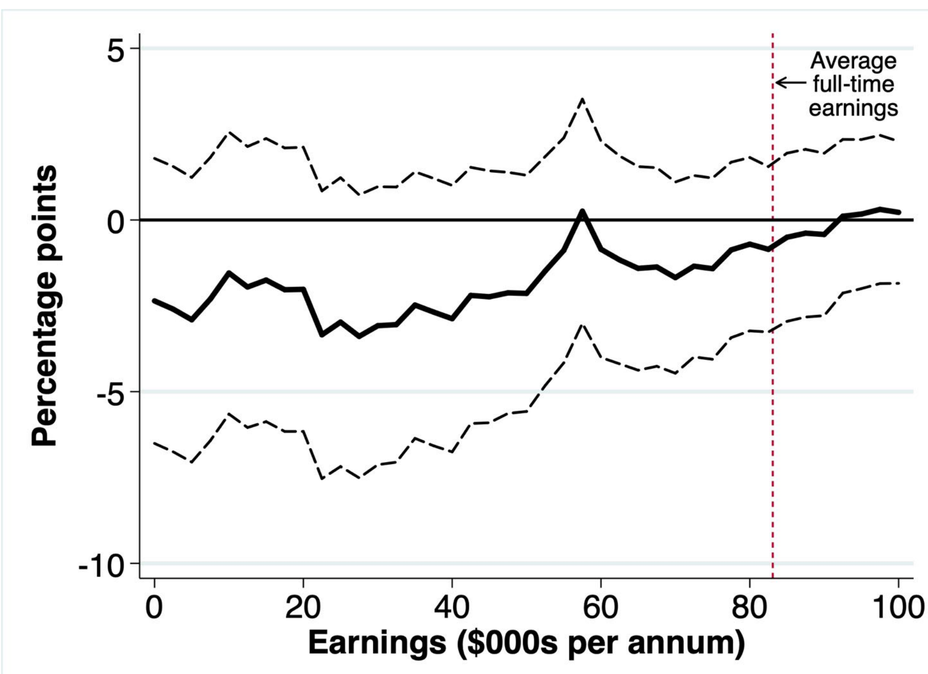
Notes: These figures compare the distribution regression estimates for OLS and logit models. Evidently, the average marginal effects from a logit model (and the associated 95% confidence intervals) are extremely similar to the OLS estimates in Figures 5 and 6.

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

(a) Hours worked > *x* per week

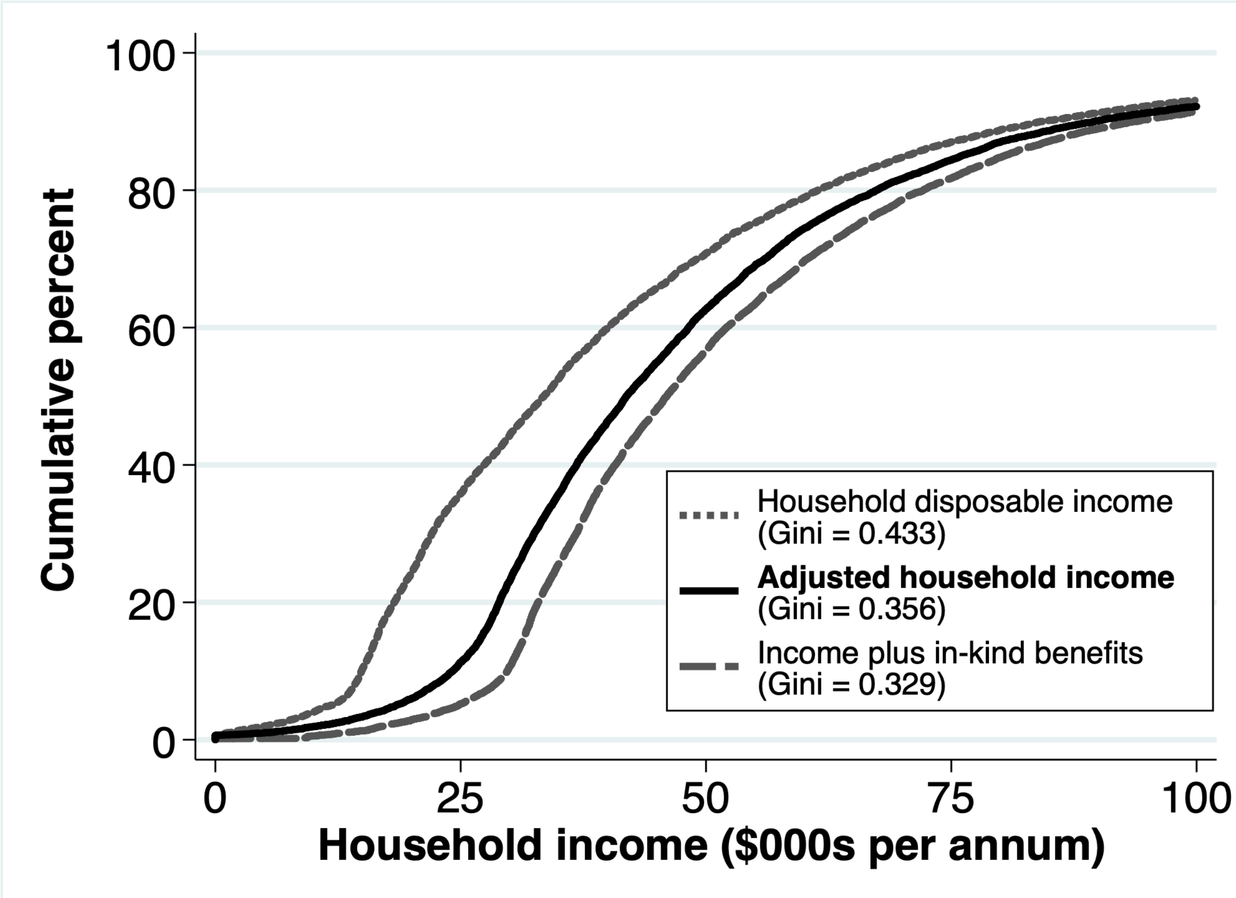


(b) Labor earnings > $*x* per annum



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.

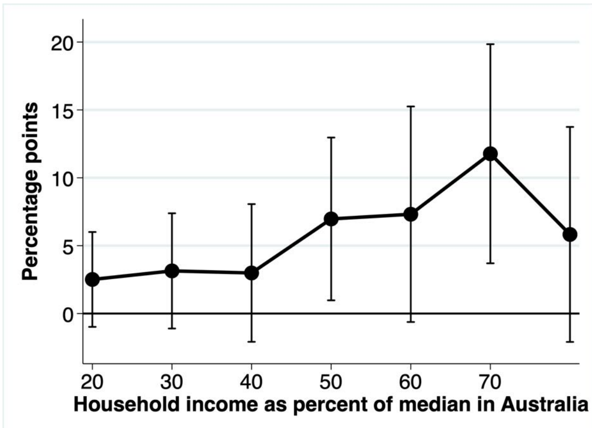
Figure A8: Impact of adjustments for housing costs and in-kind benefits on the income distribution



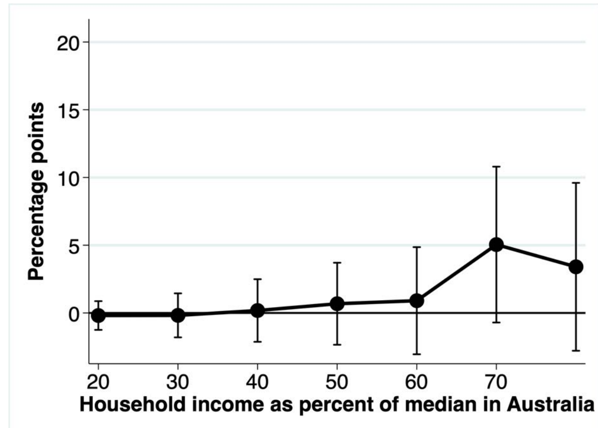
Notes: This figure shows how accounting for housing costs and in-kind benefits affects the cumulative distribution of household income for women in the sample. All three measures of household income are in 2016 dollars and are adjusted for household size. 'Income plus In-kind Benefits' is equal to household 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 less than a given percentage of the median in the Australian population

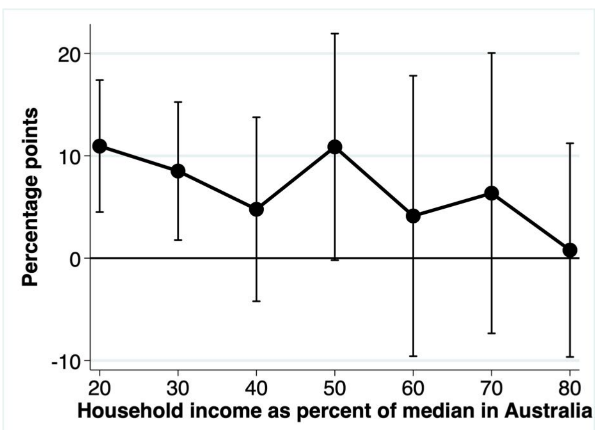
(a) Single women



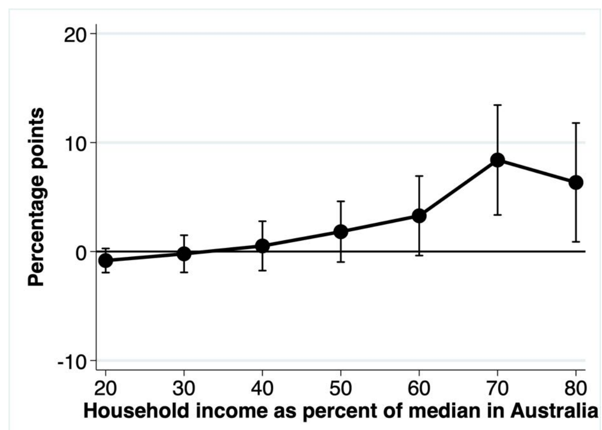
(b) Partnered women



(c) Renters



(d) Homeowners



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 80. See the text in Section 6.3 for more details.

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

		Age (in half-years)								
Date of birth	Pension age	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						
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)
Baseline: Same marital status until age 65, $r = 4\%$					
Δ NPV of Age Pension ($000s)	-49	42	-6	-0.002	0.00001
	(79)	(358)	(326)	(0.005)	(0.00525)
Observations	5,036	4,896	4,896	5,036	5,036
Baseline with $r = 6\%$					
Δ NPV of Age Pension ($000s)	-62	38	-22	-0.002	-0.0001
	(80)	(361)	(330)	(0.005)	(0.0053)
Observations	5,036	4,896	4,896	5,036	5,036
Baseline with $r = 2\%$					
Δ NPV of Age Pension ($000s)	-34	46	10	-0.001	0.0002
	(77)	(345)	(313)	(0.005)	(0.0050)
Observations	5,036	4,896	4,896	5,036	5,036
Single until age 65 if ever single at ages 55–61 ($r = 4\%$)					
Δ NPV of Age Pension ($000s)	-65	-203	-269	-0.004	0.0006
	(47)	(262)	(241)	(0.003)	(0.0034)
Observations	5,036	4,896	4,896	5,036	5,036
Partnered until age 65 if ever partnered at ages 55–61 ($r = 4\%$)					
Δ NPV of Age Pension ($000s)	57	-214	-142	0.002	-0.0044
	(64)	(229)	(197)	(0.003)	(0.0037)
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 γ from equation (A2). 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.

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%
Home owner	69.5%	91.7%	0%	100%	61.2%	95.9%	97.1%
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%
Household wealth in wave 2 ($000s)	479	803	140	774	110	453	1,522
(Std. dev.)	(617)	(1,287)	(363)	(1,022)	(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%
Employed	36.6%	33.4%	24.2%	36.1%	26.3%	36.1%	39.2%
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 APA	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***	407	-1,687*	-0.193***	0.013	0.010	0.09
	(314)	(307)	(951)	(943)	(0.024)	(0.021)	(0.021)	(0.69)
Mean at APA	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 APA	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								
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 APA	12,642	2,317	4,599	19,691	0.825	0.175	0.158	4.70
Homeowners								
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 APA	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								
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 APA	11,137	2,394	8,154	21,685	0.826	0.263	0.242	6.63
Middle third								
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 APA	6,614	1,028	13,476	21,258	0.577	0.302	0.302	8.75
Top third								
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 APA	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.3 for the details). The sample consists of women aged 60–66 from waves 1–14 of the HILDA survey.

Table A9: Estimated effects on household income across groups and decomposition of increase in inequality

		Marital status		Home-ownership		Household wealth		
	Full sample (1)	Single women (2)	Partnered women (3)	Renters (4)	Owners (5)	Bottom third (6)	Middle third (7)	Top third (8)
Panel A: Mean effects on household incomes								
Effect of reform	-1,740	-3,559	-451	-4,133	-1,809	-3,103*	-5,546**	4,158
	(1,542)	(2,222)	(2,020)	(3,479)	(1,714)	(1,635)	(2,529)	(3,657)
Mean at pension age	48,958	43,970	51,585	31,043	51,948	36,961	49,231	60,687
p-value (H_0: equal effects)		0.304		0.542			0.089	
Observations in group	7,771	2,634	5,137	1,084	6,526	2,198	2,183	2,068
Panel B: Decomposition of effects on income inequality								
Increase in Theil's T	22%	22%		16%			42%	
Decomposition								
Within-group contribution		94%		87%			80%	
Between-group contribution		6%		13%			20%	
Increase in mean log deviation	36%	36%		32%			61%	
Decomposition								
Within-group contribution		94%		88%			82%	
Between-group contribution		6%		12%			18%	
Observations	7,771	7,771		7,610			6,449	

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

Notes: Panel A presents OLS regression estimates of the causal effect of the reform, δ, from equation (1) on household incomes (net of housing costs and in-kind benefits). 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 annualised in 2016 dollars and adjusted for household size. I exclude households who ever had income in the top 1% of the sample to reduce the sensitivity of the estimates in Panel A to income changes at the top of the distribution. Panel B shows the impact on income inequality, as measured by the Theil and mean log deviation indices, using the distribution regression approach outlined in Section 6.2. I decompose this impact into changes within and between groups, by comparing the within- and between-group components of inequality in the treated and counterfactual distributions. Information on household wealth comes from wave 2, the first wave with information on wealth, and I adjust for household size by dividing household wealth by 1.5 for couples. This information is missing for 17% of observations. This subsample results in larger estimated effects on inequality measures.

Table A10: 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%	**+11%**	+9%	+8%	+6%	+5%
	[$p = 0.048$]	[$p = 0.050$]	[$p = 0.114$]	[$p = 0.126$]	[$p = 0.186$]	[$p = 0.328$]
Theil's T	+25%	**+22%**	+16%	+14%	+11%	+9%
	[$p = 0.164$]	[$p = 0.172$]	[$p = 0.284$]	[$p = 0.318$]	[$p = 0.388$]	[$p = 0.536$]
Mean log deviation	+35%	**+36%**	+26%	+17%	+11%	+14%
	[$p = 0.028$]	[$p = 0.026$]	[$p = 0.062$]	[$p = 0.194$]	[$p = 0.354$]	[$p = 0.248$]
80/20 ratio	+22%	**+20%**	+18%	+19%	+18%	+17%
	[$p = 0.024$]	[$p = 0.036$]	[$p = 0.076$]	[$p = 0.078$]	[$p = 0.130$]	[$p = 0.220$]
80/40 ratio	+18%	**+18%**	+16%	+17%	+15%	+14%
	[$p = 0.028$]	[$p = 0.028$]	[$p = 0.062$]	[$p = 0.060$]	[$p = 0.096$]	[$p = 0.168$]
Panel B: Estimated effects on relative poverty rates						
50% of median	0.020	**0.030****	0.045***	0.035*	0.033	0.040*
	(0.014)	**(0.015)**	(0.016)	(0.019)	(0.021)	(0.022)
60% of median	0.060***	**0.031**	0.039*	0.060***	0.033	0.018
	(0.018)	**(0.019)**	(0.021)	(0.022)	(0.022)	(0.022)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The p-values in Panel A are for a two-tailed test and 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.