Web Appendix for "Re-examining female labor supply responses to the 1994 Australian pension reform"

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 Other specifications in Atalay and Barrett (2015)

Atalay and Barrett's (2015) other specifications are modifications of their baseline regressions that allow the treatment effect to vary based on women's Age Pension Age (APA) cohort (and APA cohort and age). In these specifications, Atalay and Barrett (2015) divide the after cohort into five '*AC*' cohorts, AC_1 (APA $\in \{60.5, 61.0\}$) to AC_5 (APA $\in \{64.5, 65.0\}$), and estimate separate treatment effects for women in each *AC* cohort. Like in Atalay and Barrett's (2015) baseline regressions, these specifications compare the labor force participation (LFP) of women in each *AC* cohort at ages 60–64 with the LFP of women in the before cohort, using the corresponding male cohorts as comparison groups. As such, these specifications rely on five parallel-trend assumptions. These assumptions are closely related to the baseline parallel-trends assumption in equation (1); the counterfactual for women in the after cohort in the baseline regressions is a weighted average of the counterfactuals for women in the five *AC* cohorts in the expanded specifications. As such, if the baseline parallel-trends assumption does not hold, at least one of the parallel-trend assumptions in the expanded specifications must not hold.

A.2 Constructing the probability that individuals are in the after cohort

For the pre-reform sample, I construct the probability that each individual is in the after cohort, $\Pr(CohortA_i^* = 1)$, in four steps. *Step 1*: I estimate the probability that each individual is a certain age in years (i.e., $\Pr(a_i)$) at the time of the survey, for each possible age, using annual population counts by single year of age and gender from the Australian Bureau of Statistics.[A1] For example, for a woman aged 45–49, I construct their probability of being 45 years old as the total number of 45-year-old women in Australia in the relevant year divided by the total number of women aged 45–49. *Step 2*: For each possible age, I estimate an individual's conditional probability of being in the after cohort (i.e., $\Pr(CohortA_i^* = 1|a_i)$) under the assumption that

[A1] As the population data is collected in June and the survey data is collected between September and December, I use linear interpolation to estimate annual population counts by single year of age and gender at the midpoint of each survey period.

their date of birth is uniformly distributed within the possible window (of approximately fifteen months), as per Atalay and Barrett's (2015) approach. *Step 3*: For each possible age, I multiply the probability that the individual is that age (i.e., the $\Pr(a_i)$ terms from *Step 1*) by the probability that an individual of that age is in the after cohort (i.e., the $\Pr(CohortA_i^* = 1|a_i)$ terms from *Step 2*). *Step 4*: I sum up the terms in *Step 3* across each possible age to get $\Pr(CohortA_i^* = 1)$, the overall probability that individual i is in the after cohort. That is:

$$\Pr(CohortA_i^* = 1) = \sum_{a_i \in \mathbb{A}_i} \Pr(a_i) \times \Pr(CohortA_i^* = 1|a_i)$$

where $\mathbb{A}_i$ is the set of possible ages for individual i.

A.3 Additional robustness checks of the placebo estimates in Table 2

In this section, I show that the placebo estimates in Table 2 are only marginally affected by certain idiosyncrasies of the pre-reform surveys, which are discussed in detail by Siminski et al. (2003) and include (i) changes in the measurement of labor force status, (ii) seasonality in labor force participation (LFP) rates, and (iii) the lack of individual-level age controls.

Changes in the measurement of labor force status across surveys. In the 1982 survey, labor force status was categorized differently to the surveys in 1986 and 1990. Table A1 shows how labor force status was categorized in each survey:

Table A1: Categorization of labor force status in each survey

1982	1986 & 1990
Current employment status:	Labor force status in current main and second jobs:
(i) Wages & salaries full-time worker	(i) Not applicable
(ii) Own business (limited liability) full-time worker	(ii) Permanently unable to work
(iii) Wages & salaries part-time worker	(iii) Still at school
(iv) Own business (limited liability) part-time worker	(iv) Studying full time
(v) Self employed	(v) Unpaid voluntary worker
(vi) Unemployed	**(vi) Employed full-time**
(vii) Not in labor force	**(vii) Employed part-time**
	(viii) Unemployed
	(ix) Not in the labor force

Notes: I define individuals as in the labor force if they are in any of the categories in boldface: categories (i)–(vi) in the 1982 survey and categories (vi)–(viii) in the 1986 and 1990 surveys.

Due to the different categories in each survey wave, it is not possible to construct a measure of LFP that is perfectly consistent across survey waves.[A2] For example, categories (i)–(v) in the 1986 and 1990

[A2]For the 1982 survey, I define individuals as in the labor force if they are in categories (i)–(vi) (i.e., any category but 'not in surveys — people categorized as 'not applicable', 'permanently unable to work', 'still at school', 'studying full time' or 'unpaid voluntary worker' — do not have an obvious analogue in the 1982 survey. However, as just 0.6% of the sample are in any of these categories, the impact of these inconsistencies is likely to be minor compared to the size of the placebo estimates in Table 2 (15.7–24.6 percentage points).

A more important concern is that the pre-reform surveys may either over-represent or under-represent labor force participants in different ways over time for men and women. For example, if the later surveys over-represent female labor force participants but not male participants, the placebo estimates in Table 2 may be biased upwards. While Figures 3a and 3b show that the LFP rates in the pre-reform surveys are broadly consistent with the aggregate data over time, Siminski et al. (2003) present evidence that female labor force participants are slightly over-represented in later surveys. Table A2 shows the ratio of labor force totals from the pre-reform surveys to the totals in the aggregate data from the Labour Force Survey for men and women of all ages (from Table 3 in Siminski et al. (2003)); numbers less than one suggest that labor force participants are under-represented in the survey data, while numbers greater than one suggest that labor force participants are over-represented in the survey data. The table shows that (i) labor force participants are under-represented in the 1982 survey and over-represented in the 1986 and 1990 surveys, and (ii) the degree of under/over-representation is larger among women.

Table A2: Ratio of labor force totals in pre-reform surveys to totals in Labour Force Survey

	Men	Women
1982	0.991	0.979
1986	1.023	1.036
1990	1.004	1.019

Source: Siminski et al. (2003).

To examine the extent of any bias caused, I re-estimate the placebo estimates in Table 2 in a way that accounts for the degree of under/over-representation of male and female labor force participants in each survey. Specifically, I re-estimate the placebo regressions by weighted least squares, where the weight for individual i, w_i, is defined as:

$$w_i = \begin{cases} 1 \text{ if } LFP_i = 0 \\ 1/r_i \text{ if } LFP_i = 1 \end{cases}$$

where r_i is the relevant ratio in Table A2. Intuitively, this reduces (increases) the weight of male and female labor force participants by the degree to which these types of individuals are over-represented (under-represented) in the survey data relative to the aggregate data.

labor force'). For the 1986 and 1990 surveys, I define individuals as in the labor force if they are in categories (vi)–(viii) (i.e., employed full-time, employed part-time or unemployed).

Table A3 shows that this adjustment has only a modest impact on the placebo estimates. While the estimates fall by 5–10%, they remain positive, large and statistically significant at the 1% level.[A3]

Table A3: Impact of adjusting the placebo estimates for differential over-representation of female labor force participants over time

	Age 45–49 sample (1)	Age 50–54 sample 1 (2)	Age 50–54 sample 2 (3)
Panel A: Baseline estimates in Table 2			
Placebo effect (Female × After)	0.246***	0.157***	0.161***
	(0.038)	(0.027)	(0.027)
R-squared	0.163	0.206	0.207
Observations	5,085	4,763	3,733
Panel B: With adjustment			
Placebo effect (Female × After)	0.228***	0.148***	0.153***
	(0.040)	(0.027)	(0.027)
R-squared	0.163	0.208	0.207
Observations	5,085	4,763	3,733

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table shows the impact on the placebo estimates in Table 2 of adjusting for the degree of over-representation of male and female labor force participants in each survey relative to the aggregate data. Column (2) includes the full sample of 50–54-year-olds. In column (3), I remove anyone with a positive probability of being misclassified into the before and after cohorts, which restricts the sample to individuals in the 1982 and 1990 surveys. In Panel B, I re-estimate the placebo regressions using weighted least squares, where I reduce (increase) the weight of male and female labor force participants by the degree to which these types of individuals are over-represented (under-represented) in the relevant survey. The sample comes from the 1982, 1986 and 1990 Income Distribution Surveys. See equation (1) and Section 3 in the text for details on the sample and regression specification.

Seasonality in labor force participation rates. The pre-reform surveys were conducted over a shorter period of time (less than three months) than the post-reform surveys used by Atalay and Barrett (2015), which were conducted over the course of one year. As shown in Table A5 below, each of the pre-reform surveys was conducted between September and December of the relevant year.

Given the dates of the pre-reform surveys, seasonality may affect the LFP rates in these surveys. However, this impact is likely to be modest; among individuals aged 45–54 from 1982 to 1990, the average LFP rate was 1.30 percentage points higher between September and December than the rest of the year for

[A3]The estimates also remain similar if I use the sample weights included in each survey (see Table A4).

Table A4: Impact of sample weights on the placebo estimates

	Age 45–49 sample (1)	Age 50–54 sample 1 (2)	Age 50–54 sample 2 (3)
Panel A: Without sample weights			
Placebo effect (Female × After)	0.246***	0.157***	0.161***
	(0.038)	(0.027)	(0.027)
R-squared	0.163	0.206	0.207
Observations	5,085	4,763	3,733
Panel B: With sample weights			
Placebo effect (Female × After)	0.221***	0.135***	0.147***
	(0.043)	(0.028)	(0.029)
R-squared	0.158	0.203	0.208
Observations	5,085	4,763	3,733

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses.

Notes: This table shows the impact of using the sample weights on the placebo estimates in Table 2, which test for an effect of the reform on the labor force participation of women in the after cohort before the reform was announced. The sample comes from the 1982, 1986 and 1990 Income Distribution Surveys. Column (2) includes the full sample of 50–54-year-olds. In column (3), I remove anyone with a positive probability of being misclassified into the before and after cohorts, which restricts the sample to individuals in the 1982 and 1990 surveys. See equation (1) and Section 3 in the text for details on the sample and regression specification.

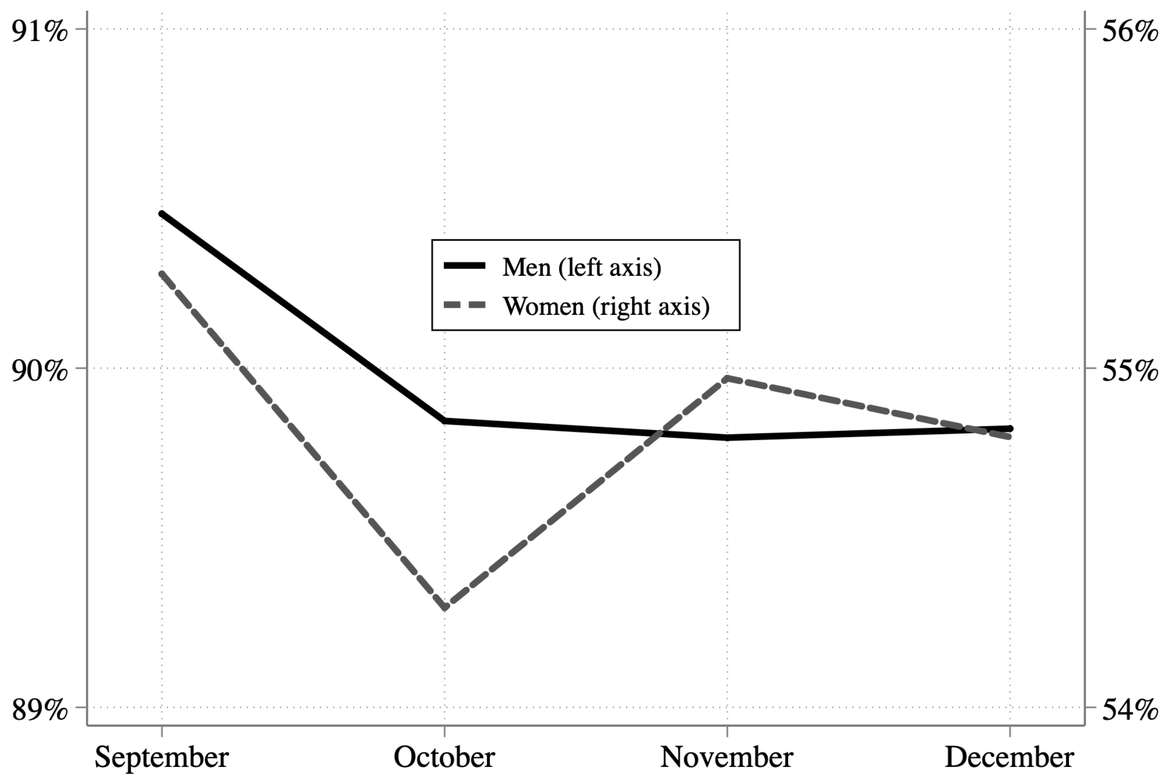
Table A5: Exact dates of pre-reform surveys

1982	Sometime between September and November
1986	September 14 to December 8
1990	October 7 to December 15

Source: Siminski et al. (2003).
Notes: Exact dates are not available for the 1982 survey.

women and 0.18 percentage points higher for men.[A4] Thus, the impact of seasonality on the LFP rates in the pre-reform surveys is likely to be much smaller than the size of the placebo estimates in Table 2, which range from 15.7 to 24.6 percentage points. Moreover, seasonality is likely to have an even smaller impact on the *trends* in male and female LFP rates between surveys, given that each survey was conducted over similar months of the year. Figure A1 shows that there is little seasonality within the survey months. Figure A1 shows the average LFP rates of men and women aged 45–54 by month in the Labour Force Survey from 1982 to 1990. Unsurprisingly, the LFP rates are similar across the survey months, varying by less than one percentage point. Overall, the impact of seasonality on the placebo estimates is likely to be minor.

Figure A1: Average labor force participation rates by month within the survey months (1982–1990)



Lack of individual-level age controls. Another possible concern is the absence of controls for age. Since age information is only available in five-year groups in the pre-reform surveys, it is not possible to control for the age of individuals in these regressions. There are two reasons why controls for age may be useful in Atalay and Barrett's (2015) differences-in-differences set-up: (i) as older individuals are generally less likely to be in the labor force, including age controls may improve the precision of estimates of the treatment effect and (ii) the age distribution may change in different ways over time for men and women,

[A4] For this calculation, I use monthly data from the Labour Force Survey.

biasing estimates of the treatment effect. However, age controls have only a minor impact on the estimates. For example, with age controls, Atalay and Barrett's (2015) baseline estimate falls by less than 3% (from 8.0 percentage points to 7.8 percentage points, see Table A7), and the standard error is almost identical to three decimal places. Moreover, I find that the placebo estimates are similar with gender-specific controls for the average age of individuals in the population (of the relevant ages) at the time of each survey; as shown in Table A8, the placebo estimates fall by just 2%.[A5] This minor impact of controlling for age is not surprising given that (i) the identification strategy relies on the corresponding male cohorts as a comparison group for female cohorts and (ii) among individuals aged 45–49 and 50–54, the average age of men and women changed in similar ways over time, as shown in Table A6 below.[A6]

Table A6: Average age of men and women at the time of each (pre-reform) survey

	1982	1986	1990
Average age (among those aged 45–49): Women	47.45	47.44	47.39
Average age (among those aged 45–49): Men	47.45	47.43	47.40
Average age (among those aged 50–54): Women	52.52	52.45	52.45
Average age (among those aged 50–54): Men	52.52	52.45	52.44

Notes: This table presents the average age of men and women aged 45–49 and 50–54 at the time of each survey. I construct these averages using annual populations counts by single year of age and gender from the Australian Bureau of Statistics.

Table A7: Impact of age controls on Atalay and Barrett's (2015) estimates

	No age controls (1)	With age controls (2)
Treatment effect (Female $\times$ After)	0.080***	0.078***
	(0.021)	(0.022)
R-squared	0.103	0.121
Observations	11,497	11,497

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table shows the impact of individual-level age controls on Atalay and Barrett's (2015) baseline estimate of the effect of the reform on female labor force participation. The estimates come from estimating equation (1). The sample consists of men and women aged 60–64 from the eleven Income and Housing Costs Surveys between 1994/95 and 2009/10.

[A5]Due to perfect multicollinearity, it is not possible to include age controls for the sample in column 3 of Table 2, where there is no misclassification of individuals into the before and after cohorts.

[A6]Unsurprisingly, the age distribution also changed in similar ways over time for men and women (see Table A9).

Table A8: Impact of age controls on the placebo estimates

	Age 45–49 sample (1)	Age 50–54 sample 1 (2)
Panel A: Without age controls		
Placebo effect (Female × After)	0.246***	0.157***
	(0.038)	(0.027)
R-squared	0.163	0.206
Observations	5,085	4,763
Panel B: With age controls		
Placebo effect (Female × After)	0.241***	0.154***
	(0.039)	(0.027)
R-squared	0.164	0.206
Observations	5,085	4,763

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table shows the impact of age controls on the placebo estimates in Table 2, which test for an effect of the reform on the labor force participation of women in the after cohort before the reform was announced. Column (2) includes the full sample of 50–54-year-olds. In Panel B, I include gender-specific controls for the average age of individuals in the sample at the time of each survey. The sample comes from the 1982, 1986 and 1990 Income Distribution Surveys. See equation (1) and Section 3 in the text for details on the sample and regression specification.

Table A9: Age distribution of men and women at the time of each (pre-reform) survey

	1982	1986	1990
Age 45–49 sample:			
Proportion aged 45: Women	0.210	0.214	0.223
Proportion aged 45: Men	0.210	0.215	0.221
Proportion aged 46: Women	0.208	0.205	0.213
Proportion aged 46: Men	0.208	0.207	0.212
Proportion aged 47: Women	0.199	0.198	0.195
Proportion aged 47: Men	0.198	0.197	0.195
Proportion aged 48: Women	0.191	0.193	0.185
Proportion aged 48: Men	0.192	0.192	0.187
Proportion aged 49: Women	0.192	0.190	0.184
Proportion aged 49: Men	0.192	0.189	0.185
Age 50–54 sample:			
Proportion aged 50: Women	0.190	0.215	0.211
Proportion aged 50: Men	0.192	0.215	0.213
Proportion aged 51: Women	0.204	0.201	0.203
Proportion aged 51: Men	0.202	0.201	0.206
Proportion aged 52: Women	0.201	0.195	0.200
Proportion aged 52: Men	0.201	0.195	0.200
Proportion aged 53: Women	0.201	0.196	0.195
Proportion aged 53: Men	0.202	0.195	0.194
Proportion aged 54: Women	0.203	0.193	0.191
Proportion aged 54: Men	0.203	0.194	0.187

Notes: This table shows the age distribution among men and women aged 45–49 and 50–54 at the time of each survey. I construct this table using annual populations counts by single year of age and gender from the Australian Bureau of Statistics.

A.4 Examining the key identifying assumption in Section 4 ("Re-estimating the impact on female participation")

To test the validity of the parallel-trends assumption underpinning equation (2), I examine how women's labor force participation (LFP) changes leading up to their Age Pension Age (APA). If the parallel-trends assumption is suitable, we would not expect to see any trend in women's LFP leading up to the APA after controlling for age, cohort and individual factors. That is, after controlling for age, cohort and individual factors, women who are one year below the APA should have a similar level of LFP to women who are two, three or four years below the APA. To test whether this is true, I estimate regressions based on the following specification for all women aged 55–64 in the surveys used by Atalay and Barrett (2015):

$$\begin{aligned} LFP_i = & \beta x_i + \delta_{-5}\mathbf{1}(age_i < APA_i - 4) + \sum_{\substack{j=-4 \\ j\neq -1}}^{0} \delta_j \mathbf{1}(age_i \in [APA_i + j, APA_i + j + 1)) \\ & + \delta_1 \mathbf{1}(age_i \geq APA_i + 1) + FE\_age\_yrs_i + FE\_APA\_cohort_i + u_i \end{aligned} \tag{A1}$$

where these regressions modify equation (2) by replacing $\mathbf{1}(age_i < APA_i)$ with six indicator variables that each specify that a woman is a certain number of years above or below the APA. This divides women into seven mutually exclusive and exhaustive categories based on how many years they are above or below the APA. As information on age is only available in years, there is potential misclassification for some women of (i) the cohort fixed effects, $FE\_APA\_Cohort_i$, and (ii) the indicator variables specifying that a woman is a certain number of years above or below the APA. I address this issue by replacing these variables with (i) the probabilities that woman i is in each APA cohort, $\Pr(APA\_Cohort_i)$, and (ii) the probabilities that woman i is in each of the six categories in equation (2).[A7]

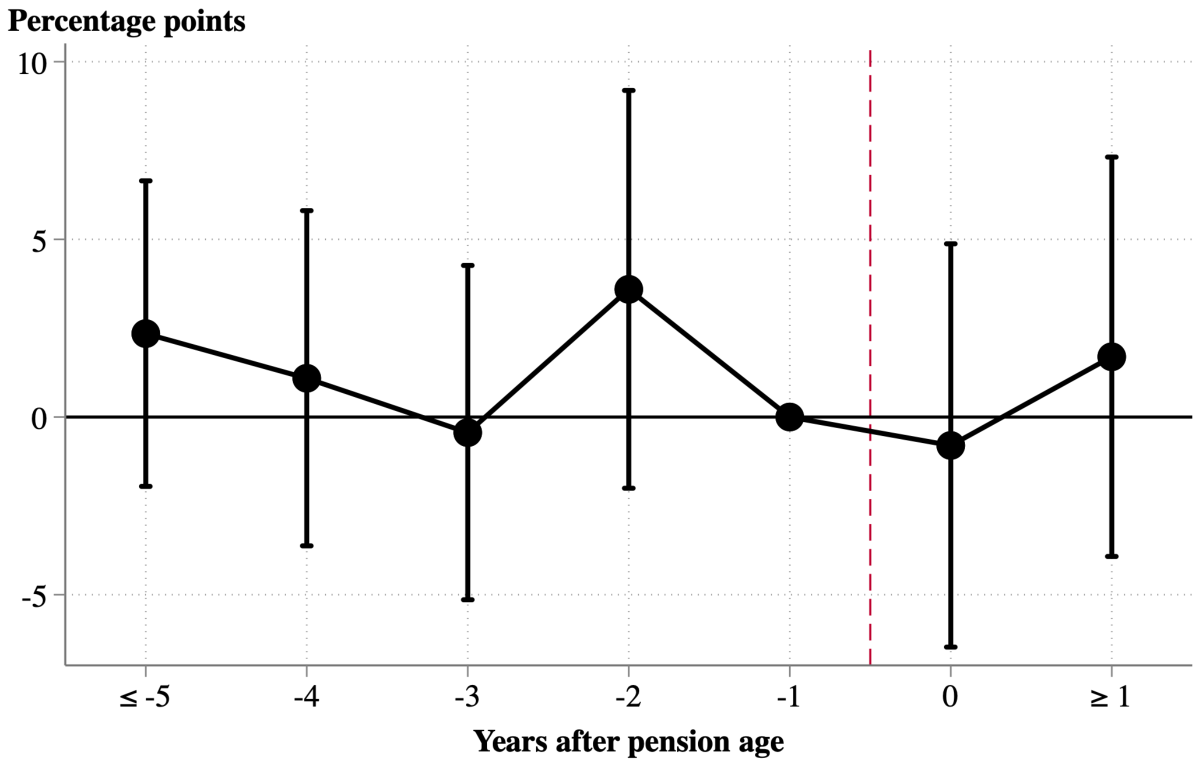
Each of the δ coefficients in equation (A1) should be interpreted as an effect relative to the omitted category, which is women who are less than or equal to one year below the APA. For example, δ_{-2} estimates the effect on LFP from women being 1 to 2 years below the APA, i.e. $age_i \in [APA_i - 2, APA_i - 1)$, relative to them being less than or equal to one year below the APA, i.e. $age_i \in [APA_i - 1, APA_i)$. If the identification strategy is suitable, we would expect the δ_j coefficients to be close to zero and statistically insignificant for $j < 0$. We would also not expect to find large estimates of the δ_0 and δ_1 coefficients given that equation (2) provides only weak evidence that the reform caused any increase in female LFP.

Figure A2 presents the estimated δ coefficients with 95% confidence intervals. The estimates support the identification strategy on both dimensions: (i) the estimated δ_j coefficients are close to zero and statistically insignificant at the 10% level for $j < 0$, both individually and jointly,[A8] providing no evidence of any trend in women's LFP leading up to the APA, and (ii) the estimates of δ_0 and δ_1 are not significant at the 10% level, providing little evidence of any drop in female LFP from women reaching the APA.

[A7] I construct these variables using the same approach as discussed below equation (2). First, I calculate the probabilities that woman i is in each APA cohort, $\Pr(APA\_Cohort_i)$, under Atalay and Barrett's (2015) assumption that birthdates are uniformly distributed within the possible window of fifteen months. Then, I use these probabilities to calculate woman i's expected APA, $\mathbb{E}[APA_i]$. Finally, I calculate the probabilities that woman i is in each category under the assumption that her exact age is uniformly distributed on the interval $[age_y, age_y + 1)$, where age_y is her age in years.

[A8] The p-value on the joint significance test is 0.615.

Figure A2: Estimated effects on female labor force participation relative to year before the pension age

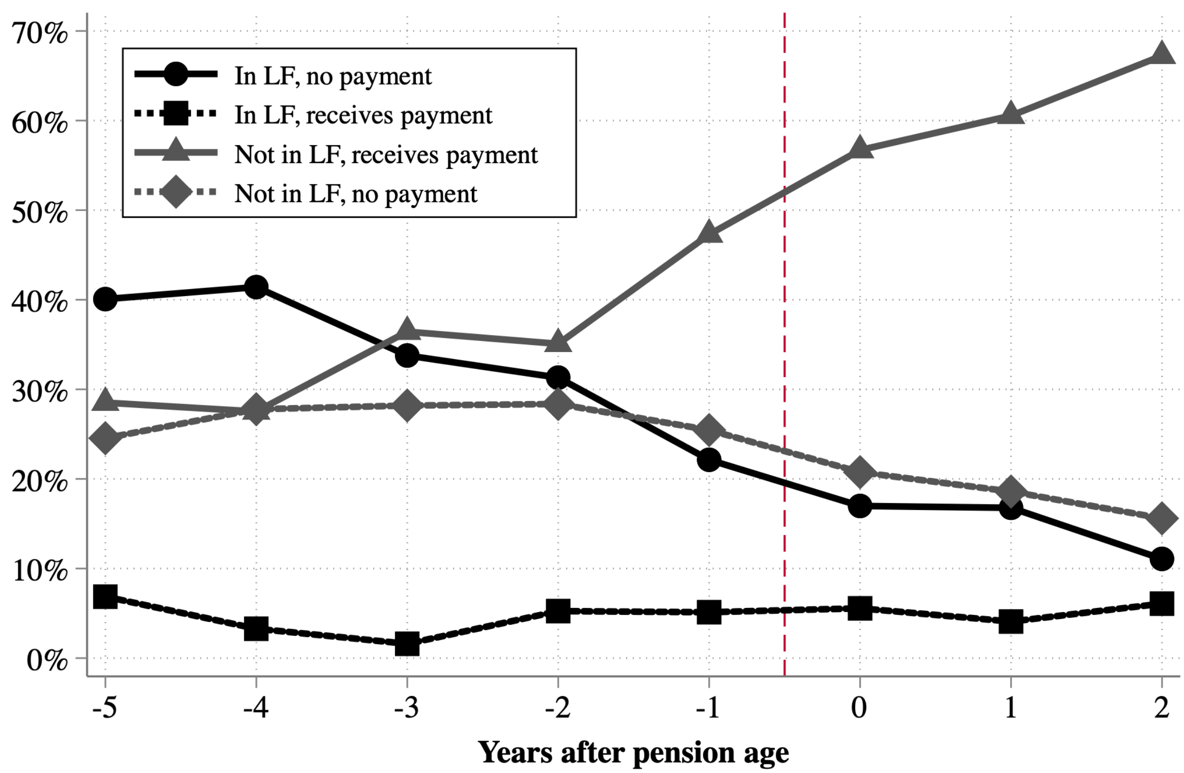


Notes: These figures plot the estimated δ coefficients, with 95% confidence intervals, from estimates of equation (A1). The regression sample consists of women aged 55–64 from the surveys used by Atalay and Barrett (2015) — the eleven Income and Housing Costs Surveys from 1994/95 to 2009/10.

A.5 Assessing whether "not in the labor force" is a suitable proxy for retirement

I follow Atalay and Barrett (2015) and use labor force participation (LFP) as the key variable for analysis. Atalay and Barrett (2015) implicitly use "not in the labor force" as a proxy for retirement behavior. For example, in the abstract, they discuss a "decline in the probability of retirement by 12 to 19 percentage points". While I am slightly more cautious in the interpretation of my estimates, it is worth discussing whether the estimates in this paper are likely to reflect retirement decisions or other factors affecting female LFP. As discussed in the paper, my preferred identification strategy picks up within-cohort changes in LFP around the APA. This variation seems likely to result mainly from decisions to retire, given the low re-entry into the labor market at these ages. To assess this further, I examine whether the decrease in LFP rates around the APA coincides with an increase in the receipt of government payments. Specifically, I divide women into four groups: (i) in labor force and not receiving a government payment (21.9% of the sample); (ii) in labor force and receiving a government payment (4.9%); (iii) not in labor force and receiving a government payment (51.2%); (iv) not in labor force and not receiving a government payment (22.0%). Evidently, most women are either in the labor force or receiving income support but not both. Moreover, Figure A3 shows that, around the APA, women mainly transition from group (i) ‘in labor force, no payment’ to group (iii) ‘not in labor force, receives payment’. Thus, within-cohort changes in LFP around the APA are likely to mainly reflect decisions to retire with the support of government payments.

Figure A3: Changes in female labor force participation and government support around the pension age

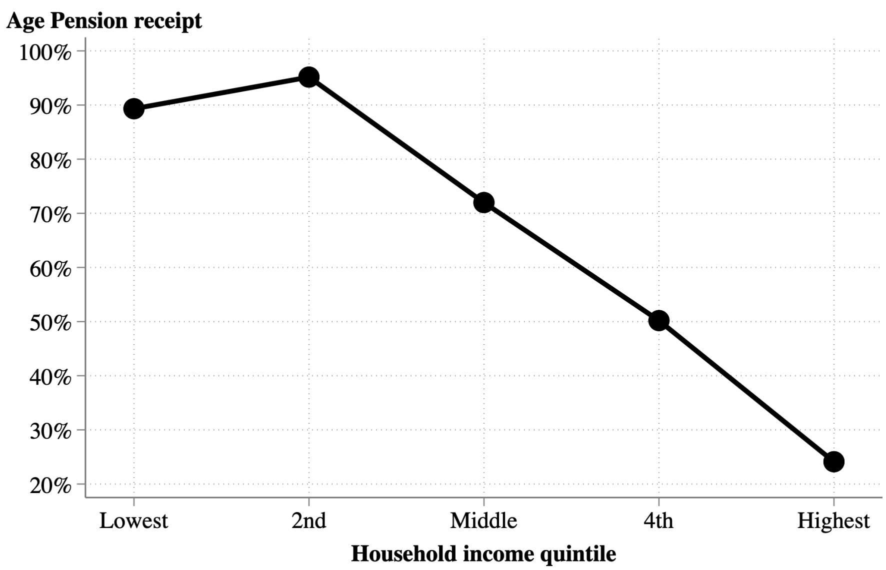


Notes: This figure shows the proportion of women in each of the four groups (based on their labor force participation and receipt of government payments) by the number of years to their pension age. I assign women’s number of years to pension age based on their most likely group, assuming birthdates are uniformly distributed within the possible 15-month windows. The sample comes from the eleven Income and Housing Costs Surveys used by Atalay and Barrett (2015) from 1994 to 2010.

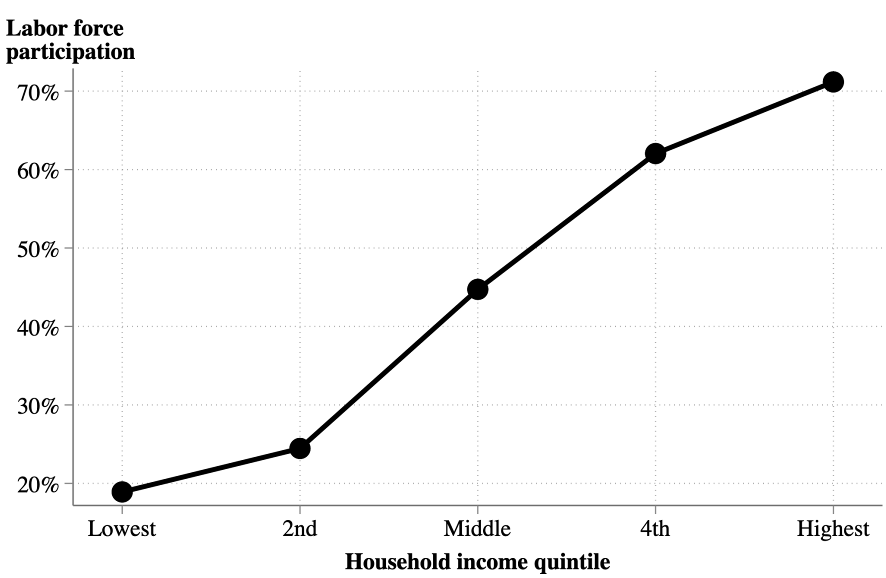
B Additional figures and tables referenced in the main text

Figure B1: Age Pension receipt and labor force participation by household income

(a) Age Pension receipt: Pension-age women

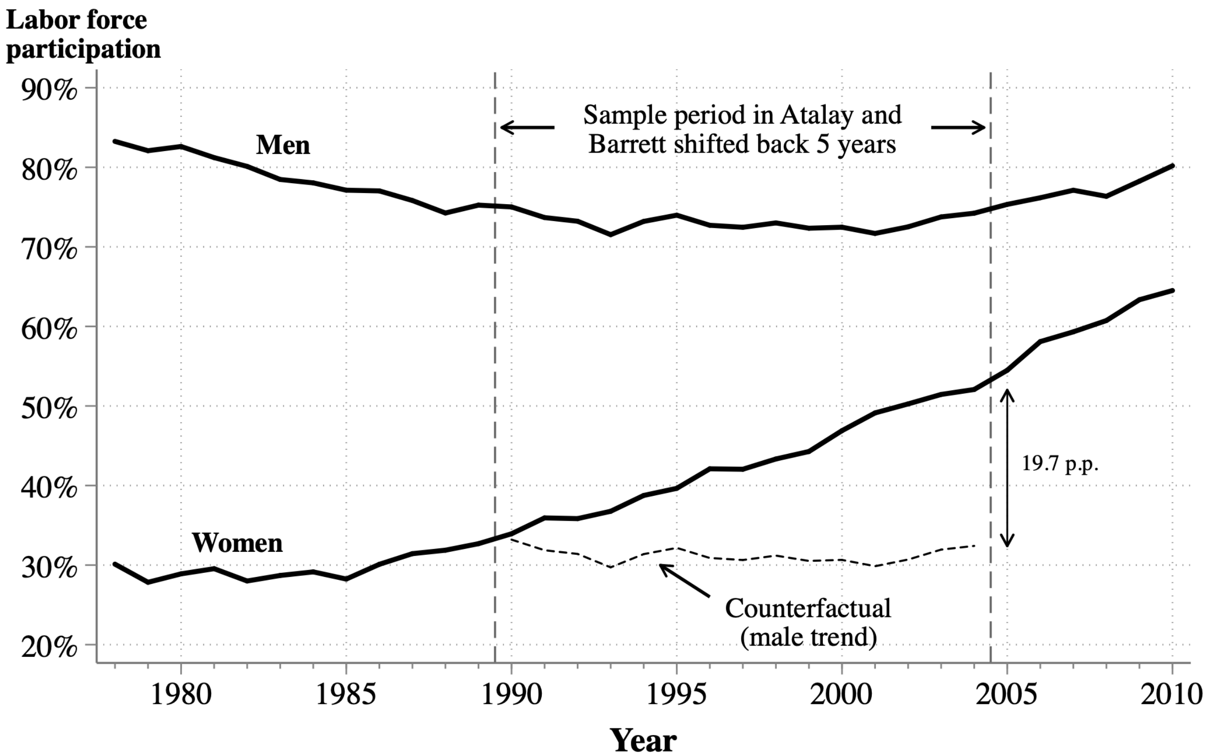


(b) Labor force participation: Women at age 59



Notes: These figures show how Age Pension receipt and labor force participation vary among women based on their household income. I divide women into quintiles based on their equivalised household income, after excluding their own income from the Age Pension and other transfers. In (a), I restrict the sample to women who are above the pension age. In (b), I restrict the sample to women aged 59, who are all below the pension age. The overall sample consists of women aged 59–64 from the eleven Income and Housing Costs Surveys used by Atalay and Barrett (2015) from 1994 to 2010.

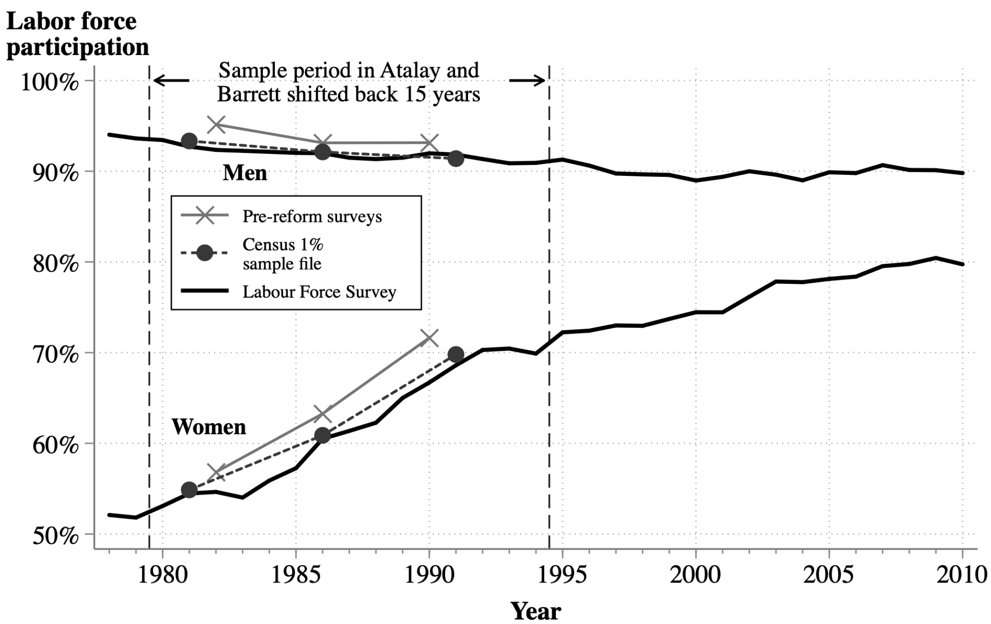
Figure B2: Labor force participation rates of the relavant cohorts at ages 55–59



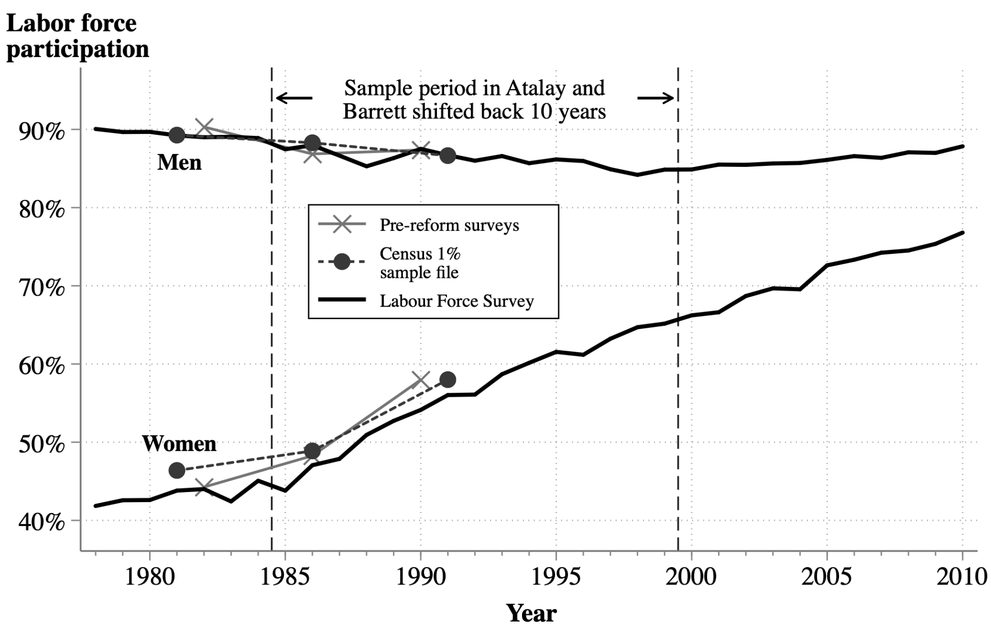
Notes: This figure plots labor force participation (LFP) rates from the Labour Force Survey. Between the vertical dashed lines, the cohorts in Atalay and Barrett's sample were aged 55–59. During these periods, male and female LFP rates follow opposing trends.

Figure B3: Comparison of the participation rates in the pre-reform surveys with Census data

(a) Ages 45–49



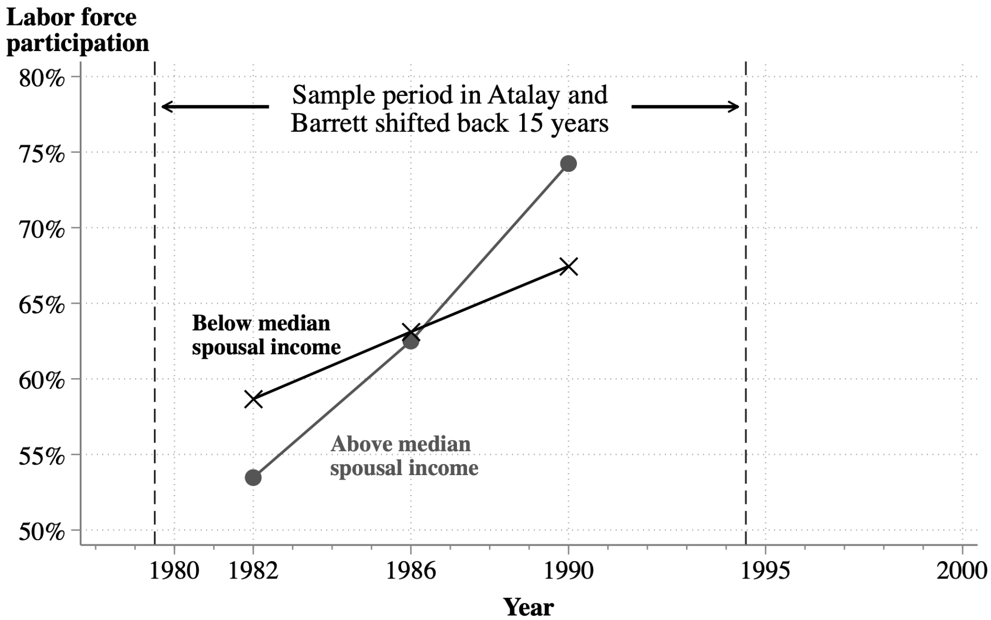
(b) Ages 50–54



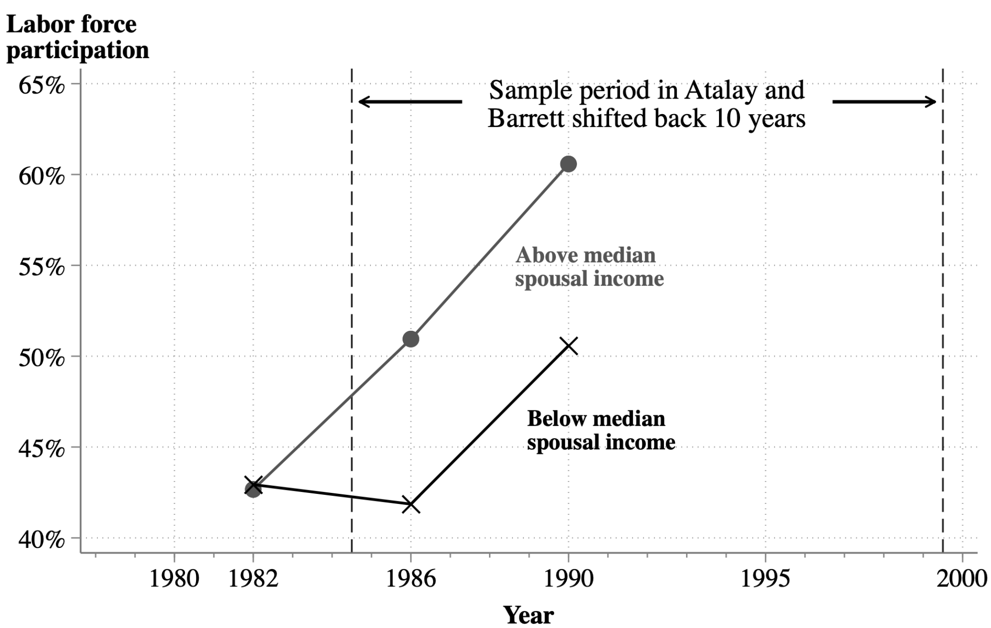
Notes: These figures augment Figures 3a and 3b with the labor force participation (LFP) rates in the 1% sample files for the 1981, 1986 and 1991 Censuses. These figures show that the LFP rates in the pre-reform surveys are consistent with the aggregate data in both the Census and the Labour Force Survey.

Figure B4: Pre-reform labor force participation rates of partnered women by spouses' income

(a) Ages 45–49



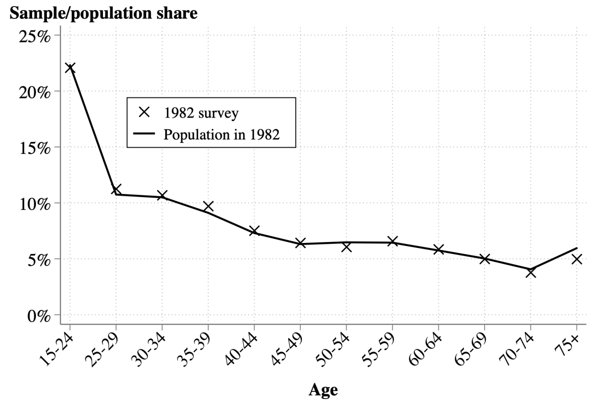
(b) Ages 50–54



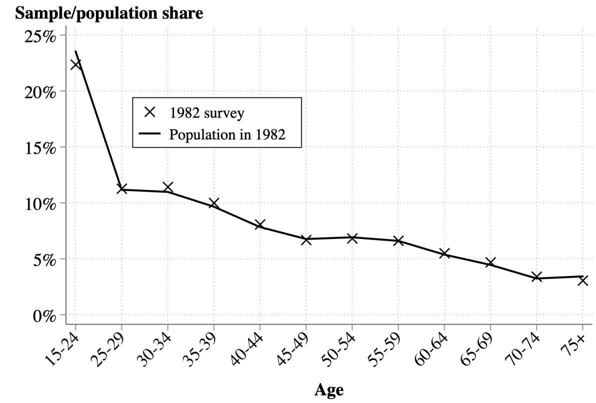
Notes: These figures plot the labor force participation (LFP) rates of partnered women from the 1982, 1986 and 1990 Income Distribution Surveys. I divide women in half based on the median income of their spouses for women of the relevant age in the relevant survey. Between the vertical dashed lines, the cohorts in Atalay and Barrett's (2015) sample were aged 45–49 in (a) and 50–54 in (b). These graphs show that partnered women with low- and high-income spouses contributed to the higher LFP rates among later cohorts, but the increase was larger for those with high-income spouses. See Section 3 for more details.

Figure B5: Representativeness of the pre-reform surveys with respect to age

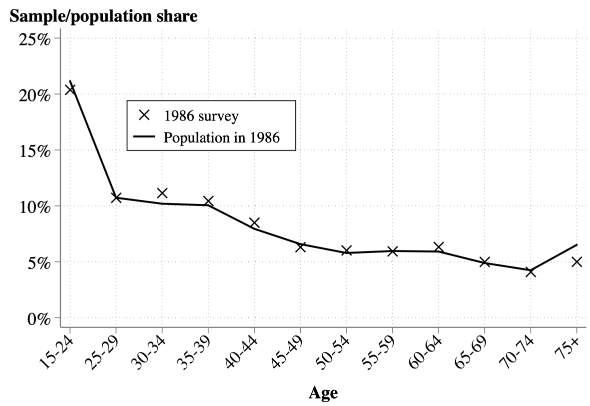
(a) 1982: Women



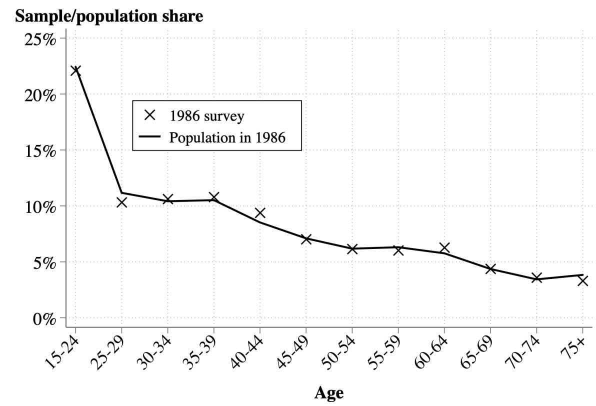
(b) 1982: Men



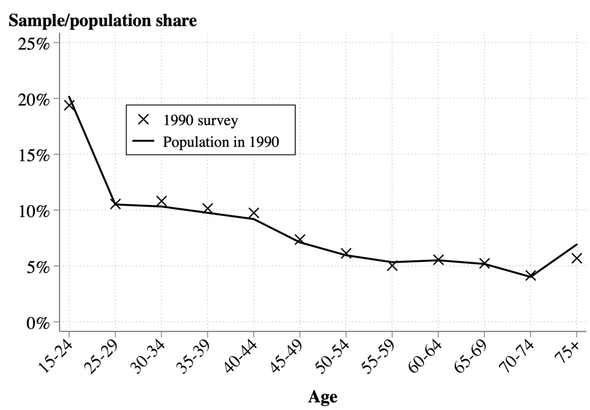
(c) 1986: Women



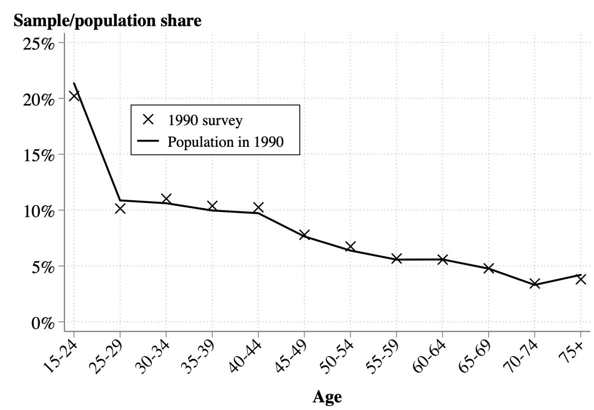
(d) 1986: Men



(e) 1990: Women



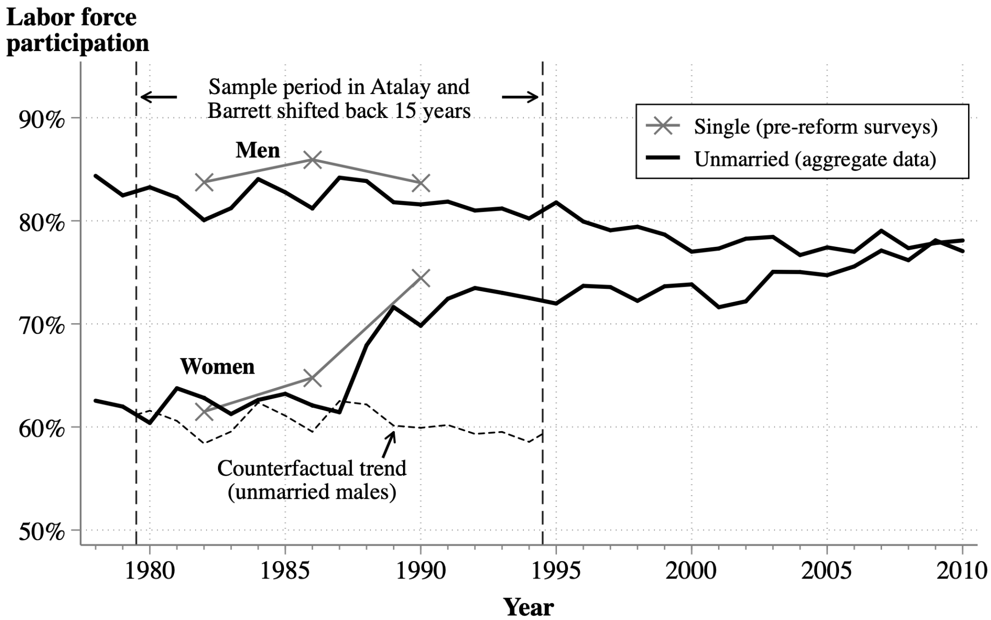
(f) 1990: Men



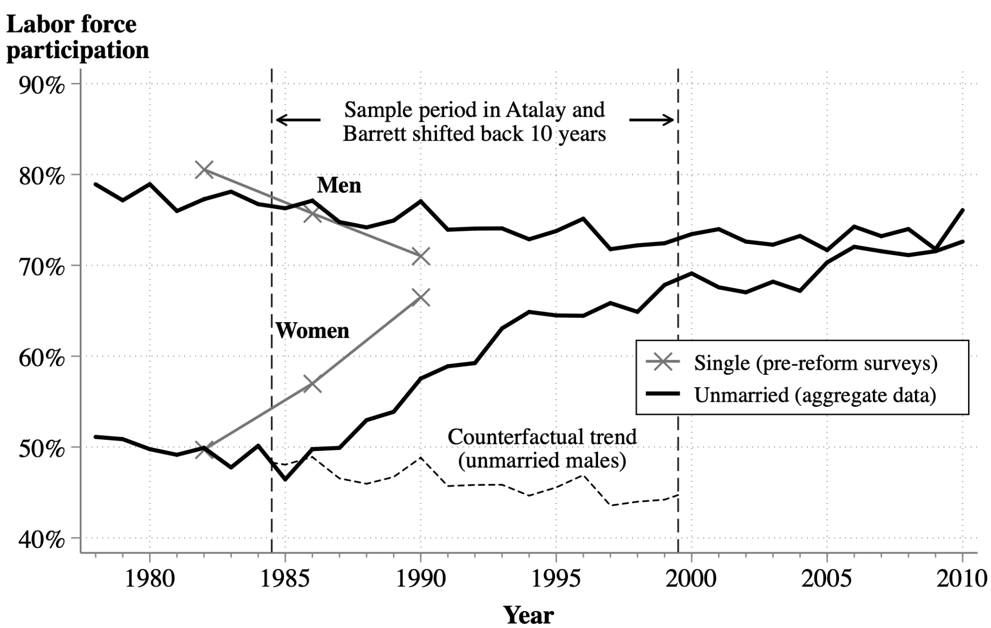
Notes: This graph examines the representativeness of the pre-reform surveys with respect to age. It shows the fraction of male and female respondents from the 1982, 1986 and 1990 Income Distribution Surveys in different age groups. The surveys allow for a consistent definition of the age groups shown above. I use estimates of the total population by single year of age and gender from the Australian Bureau of Statistics to construct the fraction of the 15+ population in each age group in the relevant years (1982, 1986, 1990) for men and women.

Figure B6: Participation rates in the pre-reform surveys of single/unmarried men and women

(a) Ages 45–49

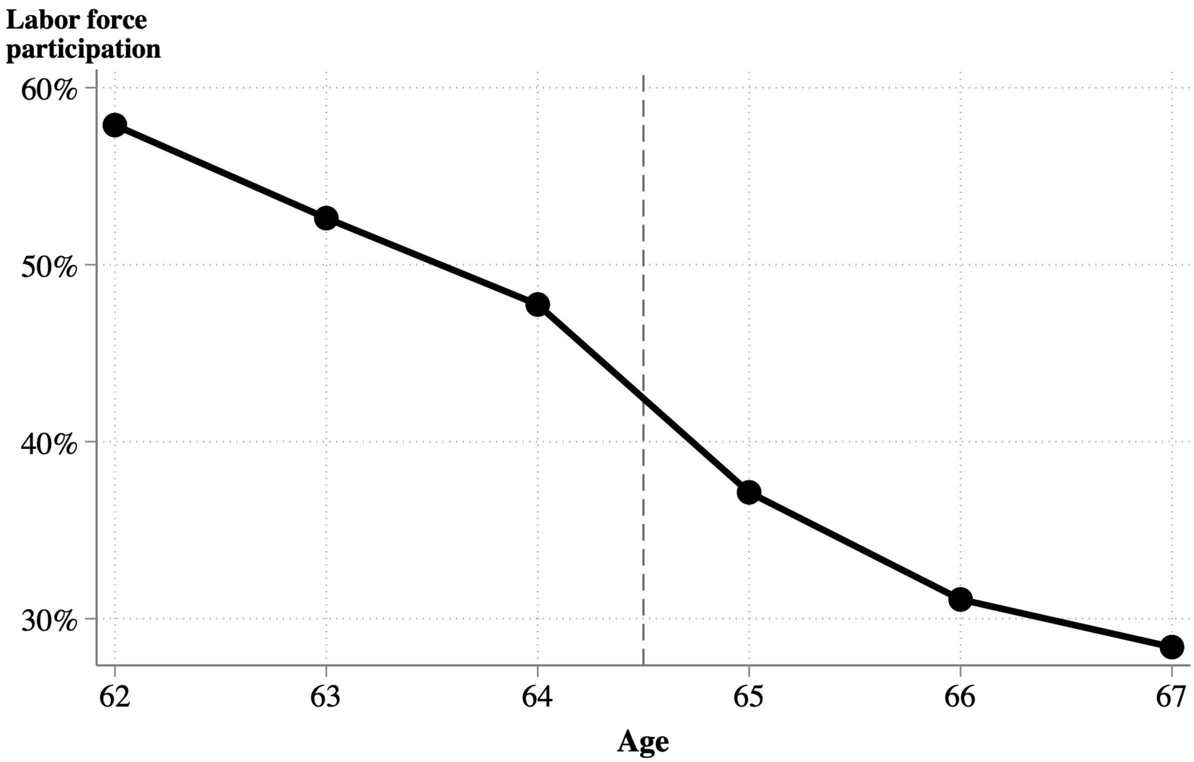


(b) Ages 50–54



<u>Notes</u>: These figures plot labor force participation (LFP) rates from the Labour Force Survey (in black) and the Income Distribution Surveys (in gray). The gray line shows the LFP rates of single men and women, while the black line shows the LFP rates for unmarried men and women. Between the vertical dashed lines, the cohorts in Atalay and Barrett's (2015) sample were aged 45–49 in (a) and 50–54 in (b). During these periods, male and female LFP rates follow opposing trends for both the single and unmarried samples.

Figure B7: Male labor force participation rates around their pension age of 65



Notes: This figure shows how male labor force participation rates change around the pension age, which was constant at 65 over the sample period (2001–15). The sample comes from waves 1–14 of the Household Income and Labour Dynamics in Australia Survey.

Table B1: Comparison of the number of observations and means of key variables in the replication sample to Atalay and Barrett's (2015) sample

	Women		Men	
	Atalay & Barrett (1)	Replication (2)	Atalay & Barrett (3)	Replication (4)
Labor force participation	0.27	0.27	0.52	0.52
Age (years)	61.9	61.9	61.9	61.9
Single	0.29	0.30	0.19	0.19
Bachelor degree +	0.17	0.17	0.19	0.20
Household size	1.73	1.75	1.90	1.91
Receives any government payment				
Full sample	0.49	0.56	0.32	0.37
Below APA		0.39		0.37
Receives payment other than the Age Pension				
Below APA		0.39		0.37
Above APA		0.06		N/A
Equivalised household income per week (2008 AUD)				
Receives any government payment		441		454
Not receiving any government payment		860		894
Observations	5,822	5,863	5,587	5,634

Notes: This table shows that I am able to closely replicate Atalay and Barrett's (2015) sample. The replication sample is similar to Atalay and Barrett's (2015) sample in terms of the number of observations and the means of the key variables. The sample consists of men and women aged 60–64 from the eleven Income and Housing Costs Surveys between 1994/95 and 2009/10.

Table B2: Replication of Atalay and Barrett's (2015) other labor force participation estimates

	Atalay & Barrett (1)	Replication (2)
Estimates in Table 3 on single sample		
Treatment effect (Female × After)	0.135***	0.138***
	(0.043)	(0.043)
Female	-0.218***	-0.226***
	(0.038)	(0.038)
After cohort	-0.011	-0.005
	(0.037)	(0.037)
R-squared		0.090
Observations	2,809	2,825
Key estimates in Table 4		
Female × Pension age ∈ {60.5, 61.0}	0.032	0.031
	(0.026)	(0.033)
Female × Pension age ∈ {61.5, 62.0}	0.068**	0.067**
	(0.029)	(0.029)
Female × Pension age ∈ {62.5, 63.0}	0.028	0.030
	(0.026)	(0.028)
Female × Pension age ∈ {63.5, 64.0}	0.103***	0.106***
	(0.028)	(0.028)
Female × Pension age ∈ {64.5, 65.0}	0.128***	0.127***
	(0.033)	(0.032)
R-squared		0.126
Observations	11,409	11,497
Key estimate in Table 5		
Hypothetical one-year increase in the pension age	0.120**	0.121**
	(0.049)	(0.050)
R-squared		0.131
Observations	11,409	11,497

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table replicates Atalay and Barrett's (2015) estimates in Table 3 for the single sample and the key estimates in Tables 4 and 5. Table 4 presents estimates where the baseline regressions are modified to allow the effect on female labor force participation to vary based on women's pension age. Table 5 presents estimates where the regressions are modified further to allow the treatment effect to vary based on women's pension age and age; Atalay and Barrett (2015) estimate the average response to a one-year increase in the pension age as a weighted average of these estimates. I use the delta rule to construct the standard error of this estimate. The sample consists of men and women aged 60–64 from the eleven Income and Housing Costs Surveys between 1994/95 and 2009/10.

Table B3: Comparison of the placebo estimates in Table 2 to those using the uniform distribution to construct the probability that individuals are in the after cohort

	Age 45–49 sample (1)	Age 50–54 sample 1 (2)
Panel A: Baseline estimates in Table 2		
Placebo effect (Female × After)	0.246***	0.157***
	(0.038)	(0.027)
R-squared	0.163	0.206
Observations	5,085	4,763
Panel B: Estimates using uniform distribution		
Placebo effect (Female × After)	0.246***	0.156***
	(0.040)	(0.026)
R-squared	0.163	0.206
Observations	5,085	4,763

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table shows how the placebo estimates in Table 2 change if $\Pr(CohortA_i^* = 1)$, the probability that individual i is in the after cohort, is constructed under the assumption that each person's birthdate is uniformly distributed within the possible window, rather than the more accurate approach in Table 2 (and described in Appendix A.2), which uses annual population counts by gender and single year of age. Column (2) includes the full sample of 50–54-year-olds. See equation (1) and Section 3 for details on the regression specification.

Table B4: Placebo regression estimates on the pre-reform sample of single men and women

	Age 45–49 sample (1)	Age 50–54 sample 1 (2)	Age 50–54 sample 2 (3)
Placebo effect (Female × After)	0.162	0.240***	0.246***
	(0.113)	(0.068)	(0.071)
Female	-0.297***	-0.324***	-0.337***
	(0.095)	(0.045)	(0.050)
After cohort	-0.028	-0.095*	-0.098**
	(0.082)	(0.050)	(0.050)
Mean: Women in after cohort	0.714	0.665	0.665
Implied percent increase in female participation	29.3%	56.3%	58.7%
R-squared	0.088	0.108	0.121
Observations	759	824	661

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table presents placebo regression estimates on single women's labor force participation using data from the 1982, 1986 and 1990 Income Distribution Surveys. Column (2) includes the full sample of 50–54-year-olds. In column (3), I remove anyone with a positive probability of being misclassified into the before and after cohorts, which restricts the sample to individuals in the 1982 and 1990 surveys. See equation (1) and Section 3 for details on the regression specification.

Table B5: Impact of reducing the sensitivity of the placebo estimates at ages 45–49 to differences between the 1982 survey and subsequent surveys

	Baseline estimate (1)	No excess sensitivity (2)
Placebo effect (Female × After)	0.246***	0.137***
	(0.040)	(0.022)
R-squared	0.163	0.163
Observations	5,085	5,085

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table shows the impact of reducing the sensitivity of the placebo estimates at ages 45–49 to differences between the 1982 survey and subsequent surveys. Column (1) shows the baseline estimate, where I follow Atalay and Barrett (2015) by using the probability of individuals being in the after cohort in the regressions. Column (2) reduces this probability to zero for individuals in the 1982 survey to reduce the sensitivity of the placebo estimates to differences in LFP rates between this survey and subsequent surveys. See equation (1) and Section 3 for details on the regression specification.

Table B6: Impact of controls on the placebo estimates

	Age 45–49 sample (1)	Age 50–54 sample 1 (2)	Age 50–54 sample 2 (3)
Panel A: Controls excluded from regressions			
Placebo effect (Female × After)	0.251***	0.162***	0.166***
	(0.040)	(0.028)	(0.028)
R-squared	0.143	0.181	0.184
Observations	5,085	4,763	3,733
Panel B: Controls included in regressions			
Placebo effect (Female × After)	0.246***	0.157***	0.161***
	(0.038)	(0.027)	(0.027)
R-squared	0.163	0.206	0.207
Observations	5,085	4,763	3,733

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table shows the impact of controls on the placebo estimates in Table 2, which test for an effect of the reform on the labor force participation of women in the after cohort before the reform was announced. Individual-level age controls are not available, though the estimates are similar with gender-specific controls for the average age of individuals in the sample at the time of each survey (see Table A8). The sample comes from the 1982, 1986 and 1990 Income Distribution Surveys. Column (2) includes the full sample of 50–54-year-olds. In column (3), I remove anyone with a positive probability of being misclassified into the before and after cohorts, which restricts the sample to individuals in the 1982 and 1990 surveys. See equation (1) and Section 3 for details on the sample and regression specification.

Table B7: Placebo regression estimates without men as a comparison group

	Age 45–49 sample (1)	Age 50–54 sample 1 (2)	Age 50–54 sample 2 (3)
Panel A: Controls excluded from the regressions			
Placebo effect (After cohort)	0.232***	0.137***	0.137***
	(0.038)	(0.023)	(0.022)
Mean: Women in after cohort	0.688	0.580	0.580
Implied percent increase in female participation	50.9%	31.0%	31.0%
R-squared	0.015	0.015	0.019
Observations	2,538	2,343	1,823
Panel B: Controls included in regressions			
Placebo effect (After cohort)	0.198***	0.115***	0.119***
	(0.039)	(0.023)	(0.024)
Mean: Women in after cohort	0.688	0.580	0.580
Implied percent increase in female participation	40.5%	24.8%	25.7%
R-squared	0.045	0.062	0.060
Observations	2,538	2,343	1,823

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table presents placebo regression estimates on women's labor force participation using data from the 1982, 1986 and 1990 Income Distribution Surveys. Column (2) includes the full sample of 50–54-year-olds. In column (3), I remove anyone with a positive probability of being misclassified into the before and after cohorts, which restricts the sample to individuals in the 1982 and 1990 surveys. The regressions modify equation (1) by dropping men from the sample and omitting the female dummy and the interaction term from the regression.

Table B8: Impact of controls on Atalay and Barrett's (2015) estimates

	Without controls (1)	With controls (2)
Treatment effect (Female × After)	0.087***	0.078***
	(0.022)	(0.020)
R-squared	0.076	0.121
Observations	11,497	11,497

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table shows the impact of controls on Atalay and Barrett's (2015) baseline estimate of the effect of the reform on female labor force participation. The estimates come from estimating equation (1). The sample consists of men and women aged 60–64 from the eleven Income and Housing Costs Surveys between 1994/95 and 2009/10.

Table B9: Impact of controls for Age Pension eligibility on Atalay and Barrett's (2015) estimates

	(1)	(2)
Treatment effect (Female × After)	0.078***	0.036
	(0.021)	(0.023)
Probability of being above pension age		-0.069***
		(0.015)
R-squared	0.121	0.122
Observations	11,497	11,497

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table shows that Atalay and Barrett's (2015) estimates fall by over 50% and become statistically insignificant if equation (1) is augmented with the probability that individuals have reached the APA. To construct this variable, I first calculate the probabilities that woman i is in each APA cohort, $\Pr(APA\_Cohort_i)$, under the assumption that woman i's birth date is uniformly distributed within the possible fifteen-month window, as per Atalay and Barrett's (2015) approach. Then, I use these probabilities to calculate woman i's expected APA, $\mathbb{E}[APA_i]$. Finally, I calculate $\Pr(age_i \geq \mathbb{E}[APA_i])$ under the assumption that each woman's exact age is uniformly distributed on the interval $[age_y, age_y + 1)$, where age_y is her age in years. The sample consists of men and women aged 60–64 from the eleven Income and Housing Costs Surveys between 1994/95 and 2009/10.

Table B10: Robustness of the updated estimates of the impact on female labor force participation

			Baseline without measurement error in		
	Baseline estimate (1)	Time controls, not cohort (2)	Pension eligibility (3)	Pension cohort (4)	Both (5)
Treatment effect	0.041	0.039*	0.038	0.041	0.031
	(0.026)	(0.022)	(0.027)	(0.040)	(0.042)
Cohort controls	X		X	X	X
Year-quarter fixed effects		X			
No misclassification of pension eligibility			X		X
No misclassification of pension cohort				X	X
R-squared	0.093	0.097	0.100	0.090	0.093
Observations	5,863	5,863	5,101	2,655	2,577

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Bootstrapped standard errors, based on 999 replications, in parentheses.

Notes: This table examines the robustness of the preferred estimates of the impact of the reform on female labor force participation from equation (2). The regression sample consists of all women in the replication sample: women aged 60–64 in the eleven Income and Housing Costs Surveys between 1994/95 and 2009/10. All regressions include age fixed effects and other individual-level controls. See equation (2) and Section 4 in the text for details on the sample and regression specification.