Means-tested Age-Pension and Saving

Sang-Wook (Stanley) Cho
Renuka Sane

Means-tested Age-Pension and Saving

Sang-Wook (Stanley) Cho
School of Economics, UNSW
Sydney, Australia

Renuka Sane*
Indian Statistical Institute, Delhi
India

January 2014

Abstract

We investigate whether households adjust retirement savings decisions in response to changes in the means-tested public pension plans. The policy in question lowered the taper rate of the assets test on the age pension in Australia in 2007. We use HILDA, a detailed micro panel data-set for Australian households and focus on the age group between 50 and 64 in 2006, prior to the reform. We compare savings behaviours of those who were constrained to increase financial wealth because of the assets test prior to the reform with those who were not constrained, and find that assets tests do have a perverse impact on saving.

JEL classification : H55, J14, J26
Keywords: means test, age pension, savings, portfolio choice

*Corresponding Author: Renuka Sane, renuka@saner.org.in (email). We thank Garry Barrett and John Piggott for useful discussions.

1 Introduction

Economists have devoted considerable attention to evaluating the incentive effects of means tested welfare programs. Hubbard, Skinner, and Zeldes (1994, 1995) in two influential papers suggest that assets tests act as an implicit tax on wealth and discourage saving to qualify for welfare programs. Since then the focus has been on measuring the extent to which means tests crowd out saving, or discourage labour force participation (Neumark and Powers, 1998; Powers, 1998; Gruber and Yelowtiz, 1999). Although there is broad agreement that means-testing leads to perverse incentives, differences remain over the extent of this impact, and the categories of households who respond to such incentives. Recent literature seems to suggest that means tests do not lead to a crowd out of savings of the poorest households, but there is substantial crowd-out for moderately poor and moderately wealthy households (Hurst and Ziliak, 2006; Maynard and Qiu, 2009; Gittleman, 2011).

The issue of means test design assumes importance in the context of pension system design as countries move towards greater targeting of benefits to reduce their fiscal burdens. This implies that households will bear additional responsibility of building their retirement corpus over their working lives. The role of the means test in influencing how households save for old age will thus become an increasingly critical input into pension policy. Research on the trade-offs of various means tests in such a context is relatively recent and therefore limited (Sefton, van de Ven, and Weale, 2008; Kudrna and Woodland, 2008; Tran and Woodland, 2011; Sefton and van de Ven, 2009; Tran and Woodland, 2010). Given the scale and size of pension programs around the world and the current and potential restructuring towards more targeted benefits, it is important to understand the impact of means tests on household saving and portfolio choice.

In this paper, we use the Household, Income and Labour Dynamics in Australia (HILDA) panel data to investigate the impact of the assets test on financial and non financial saving of households a few years away from retirement. The setting is the age pension in Australia which is a means tested (income and assets tests) program and constitutes the major source of income for those on the pension. A distinctive feature of this program is that the owner-occupied home is exempt from the assets test.

Financial decisions made before reaching the retirement age determine pension access, and can therefore be studied to measure perverse incentives of the assets test, if any. The

assets test taper rate on the age pension was lowered in 2007, and we exploit this policy change to evaluate if those constrained to increase financial wealth because of the assets test prior to the reform save more relative to those who were not constrained by the same. The exclusion of housing from the assets test should lead people to increase their home equity at the expense of their equity holdings in other income-yielding assets while still being eligible to receive the age pension benefit. As such, changes in the taper-rate may have less of an impact on savings on the family home.[1]

We use the propensity score matching methodology to estimate the average treatment effect of the policy change on financial and non-financing savings. Our treatment group consists of households between age 50 and 64 that was the closest to the threshold before the taper rate change. Saving is measured as the change in real wealth between 2010 and 2006. We find that the treatment group had 1.8 percent more financial savings than the control group. This group saved $110,926 more than those of the same age but not earlier constrained by the assets test. This is 16 percent of the average wealth of the overall sample in 2006, and 31 percent of the average financial wealth of those constrained by the assets test prior to the reform. There is however no significant difference in the saving on the family home of both the groups. These results are consistent across different matching specifications.

The assets test is more likely to bind to those closer to the retirement age of 65. Those in the treatment group between age 58-64 have 1.5 percent more financial savings than those not constrained by the assets test in the same age group. Our results demonstrate that the relaxation of the assets test taper rate leads to a significant increase in financial savings of those most likely to be constrained by the thresholds. This implies that at the margin, households do respond to incentives stemming from the design of welfare programs.

The results are contrary to those of Cobb-Clark and Hildebrand (2011) who use cross-sectional HILDA data to find limited evidence that average households respond to the incentives embedded in the means test of the age pension in Australia by reallocating their assets. Similarly Hulley et. al. (2013) find that the more wealthy households in retirement in Australia continue to accumulate financial (non-housing) wealth between 2002 and 2006 despite being subject to the incentives of a stricter means-test for pension

[1]Cho and Sane (2012) conduct a counter-factual experiment of abolishing the exemption of owner-occupied housing and find that this increases aggregate wealth accumulation while resulting in a lower housing investment and home-ownership.

payments. Their results are specific to the already retired households, who may behave differently from those on the verge of retirement.

The treatment group in our setting is a moderately wealthy group near the eligibility threshold a few years away from retirement, responding to a change in taper-rate that allows for larger non-housing wealth. Our results are consistent with Maynard and Qiu (2009) who find that crowd-out of savings from the Medicaid program in the US is surprisingly large for moderately wealthy households. The results are also consistent with research specifically evaluating the impact of an increase in asset limits (Powers, 1998; Hurst and Ziliak, 2006; Sullivan, 2006).

This paper raises important questions around what are the appropriate thresholds for an assets test, especially if they generate perverse incentives for the relatively well-to-do segments of the population. Policy debate in Australia has begun discussing the possibility of setting a limit on the value of the exemption of owner-occupied housing. The results of this paper contribute to this debate by showing that pension policy ought to consider the taper-rate design applicable to relatively wealthy households as well.

Section 2 describes the institutional structure of the age pension in Australia. We present the identification strategy in section 3 and the data in section 4. We test the reliability of our estimation model in section 5 and present the main results in section 6. Section 7 checks for the robustness of our results while section 8 concludes.

2 The age pension in Australia

The Australian public pension (called age pension) is a publicly managed "safety net" welfare program since 1910, where the eligibility depends on age and residency status. The welfare payment is means-tested, that is, people with income and assets falling below certain thresholds are eligible to receive different amounts of welfare payment. The applicable thresholds for the income and assets tests vary depending on the marital status and home-owership of the individual.

In 2006, the eligibility age for men was 65, and for women 62. The maximum payment rates for singles was set at $499.7 per fortnight and couples (both eligible for pension) was set at $417.2 each per fortnight. However, the income and assets tests would reduce the payment, and the test that results in the lower pension (or zero pension) continues to

get used for the final pension payout.

For the income test, employment income of more than $128 per fortnight for singles, $240 per fortnight for couples was considered part of the assessable income. The assessable income also included (and continues to include) financial asset income with a deeming rule where financial assets upto a threshold amount of $38,400 for singles and $63,800 for couples were assumed to earn an interest of 3 percent and any amount higher than the threshold was deemed to earn an interest of 5 percent. All combined together, for fornightly assessable income higher than the full pension threshold, the pension was reduced by 50 cents in the dollar, such that the pension got reduced to zero for assessable income level higher than $1557.75 per fortnight for single and $2,602 per fortnight for couple households.

For the assets test, only financial assets (less liabilities) and other non-financial assets (such as real estate investments and personal assets) were counted while owner-occupied housing equity was not. Different asset test schedules were applied to homeowners and non-homeowners. For homeowner couples, every additional $1,000 above the threshold financial asset of $229,000 reduced the fortnightly benefit by $3, while for renters, this full-pension cut-off threshold was $346,000. For homeowner couples with financial asset more than $509,500, they were no longer eligible for age pension benefit, whereas for renter couples, this part-pension cut-off threshold was $626,500.

The lower assets test thresholds (for full age pension eligibility) are adjusted annually on 1 July. The upper assets test thresholds (determining part age pension eligibility) are adjusted every six months (March and September) in line with the six-monthly age pension adjustments.[2]

Currently, most retired Australians qualify for some income support with around 77 per cent of individuals over the age of 65 (2 million people) receiving all or part of the age pension (Harmer, 2008). In addition, around 73 per cent of pensioners receive over 90 per cent of the maximum pension rate and only 3 per cent receive less than 25 per cent of the maximum rate (Henry, 2009). On average, a single pensioner who is a homeowner, is paid around 28 percent of Male Total Average Weekly Earnings (MTAWE), whereas similar partnered pensioners receive 21 percent of MTAWE per person.[3]

[2] `http://www.superguide.com.au/how-super-works/age-pension-government-retirees`

[3] For the most recent age pension entitlement, `http://www.netactuary.com.au/calculators/agepension.aspx` calculates an estimate of age pension.

The age-pension also entitles the holder to various other subsidies (such as health-care, rent, pharmaceutical, utilities and public-transport) which provides additional incentive to ensure that income and wealth stay within the bounds of the respective test thresholds.

In Australia, the income-test is said to bind more than the assets-test: 32 percent are disqualified because of the income test thresholds while only 9 percent because of the assets test thresholds. However, assets are deemed to earn a certain fixed rate of interest which is counted in the income test. Therefore, the quantity and composition of assets will impact the receipt of the pension through both tests.

2.1 The age pension reform in 2007

On 20th September 2007, the Government relaxed the age pension assets test by halving the taper rate applied to eligible assets. Instead of losing $3 a fortnight ($78 a year) for every $1,000 of assets above a certain threshold, singles and couples now lose $1.50 a fortnight ($39 a year) for every $1,000 of assets above the asset threshold. With the lowering of the taper rate on assets, the part pension thresholds increased from $509,500 to $856,500 for homeowner couples and from $330,000 to $549,250 for single homeowners, with an estimated 300,000 extra people to benefit from the new age pension. For non-homeowners, the new thresholds became $664,750 and $981,500 for couples and singles, respectively.[4]

Figure 1 provides a graphical representation of the policy change. FP is the full pension threshold, while PP-old is the part pension threshold until September 2007. PP-new is the part pension threshold after the reduction in taper-rate.

This change to the assets test effectively doubled the value of assets allowed for eligible part-age pensioners, and enabled Australian retirees who would not have been eligible to receive the age pension previously to receive a part pension. It also enabled those on the verge of retirement with assets close to the part pension thresholds to be able to accumulate more wealth without the potential loss of the pension on retirement. It was also projected that the reduction in the assets test taper rate would increase incentives for labour participation and saving especially for those close to retirement who will still rely on the age pension to finance part of their retirement.

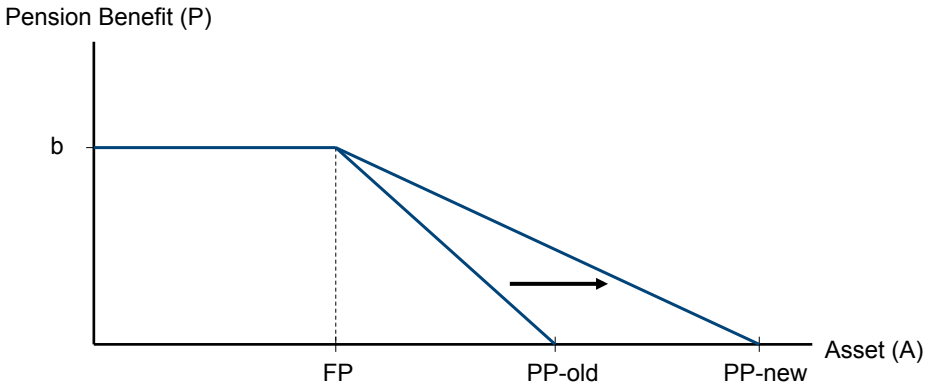
[4]Australias future tax system, August 2008, Table 2.21.

Figure 1 Assets-test taper-rate reform

FP is the full pension threshold, while PP-old is the part pension threshold until September 2007. PP-new is the part pension threshold after the reduction in taper-rate. Pension benefit (P) can be represented by the following equation:

$$P = \max\{0, \min\{b, b - t(A - FP)\}\}$$

where A is the eligible asset, b is the maximum pension benefit, t is the taper ratio.



3 Identification

Our hypothesis is that the structure of the assets test has an adverse impact on the level of savings. To test this hypothesis, we want to estimate the impact of the potential receipt of the age-pension on the saving behaviour of households. One estimation strategy is to identify households who are likely to be pensioners and observe the difference between their savings and of those who are not likely to be pension recipients. However, given that pension receipt is a function of financial resources only (apart from age), predicting pension status is likely to lead to problems of endogeneity.

The change in the assets-test taper rate allowed households to hold more financial wealth and yet be on the pension. If the age-pension has a perverse incentive on financial saving, then households affected by the taper-rate change should exhibit greater financial saving, as doing so does not disqualify them from the pension.

Let P_{it} be an indicator if household i was affected by the change in the assets-test taper rate. Let S^1_{it+s} be the saving of household i at time $t+s, s > 0$. Let S^0_{it+s} be the saving of household i had the assets-test taper rate change not been implemented. The causal

effect of the taper-rate change on household i at time period $t+s$ is:

$$S^1_{it+s} - S^0_{it+s}$$

The problem however is that S^0_{it+s} is unobservable.

To overcome this problem we define the average effect of the assets-test taper rate on households affected by it as

$$E(S^1_{t+s} - S^0_{t+s}|P_{it} = 1) = E(S^1_{t+s}|P_{it} = 1) - E(S^0_{t+s}|P_{it} = 1)$$

where causal inference relies on the construction of the counter-factual for the outcome the affected households would have experienced, on average, had they not been been affected by the change in policy (Heckman et. al., 1997; Dehejia and Wahba, 2002). This is estimated by the saving of households that were unaffected by the taper-rate change i.e. $E(S^0_{it+s}|P_{it} = 0)$.

The construction of the counter-factual requires the selection of a valid control group such that contemporaneous effects correlated with the household are controlled for. We use the assets-test taper rate as a natural experiment to create the counter-factual. The assets-test taper rate change was unexpected, therefore allowing us to compare outcomes with households that were unaffected. This is the crux of the identification strategy of this paper.

The approach is to apply matching techniques. The procedure is to pair each household affected by the policy change to one not affected by the change on the basis of some observable variables. This type of matching procedure is preferable to randomly selecting the control group as it is less likely to lead to estimation bias by picking households with completely different characteristics. We first describe the treatment group and then outline the methodology used to obtain the control sample.

3.1 The treatment group

As mentioned in Figure 1 households in the pre-retirment age group with non-housing assets between the FP and PP-old thresholds prior to the 2007 taper-rate change would not have qualified for the age pension once they reached the pension access age, if they

acquired wealth greater than the threshold. The saving of this group was therefore constrained prior to the reform.

After the taper-rate change, households can now own greater non-housing or financial wealth and yet qualify for the age pension. If means testing leads to perverse effects on saving, then we should see an increase in the saving of these households after the policy change.

The treatment group for our analysis is therefore the set of households with non-housing wealth between the full and part pension thresholds prior to the reform, and in the pre-retirement age bracket.

3.2 The control group

A valid control group is that which is similar across various observable characteristics to the treatment group except that it is unaffected by the policy change. In our setting, households not between the two thresholds constitute the control group. One could argue that a change in the taper rate affects all households, and not just the ones between the two thresholds. However, the assets test is most binding for those between the thresholds. This is the group that was closest to not obtaining the pension. It is this differential that justifies the use of those oustide of the full and part pension thresholds as a control sample.

It is desirable to perform the matching on the basis of a single index that captures all the information from the observables. We adopt the propensity score matching method outlined by Rosenbaum and Rubin (1983) which suggests the use of the probability of receiving treatment (being affected by the taper-rate reform in our context) conditional on pre-reform characteristics to reduce the dimensionality problem.

We first identify the probability (or propensity score) of being in the treatment group T, using a probit model

$$P(T = 1) = F(X_{it-1})$$

where, X is a vector of covariates observed in the time period before the taper-rate reform. That is, X includes the variables in 2002, and between 2002 and 2006 that may influence the likelihood of the household being in between the full and part pension thresholds in 2006, before the taper rate reform. The variables are as follows:

- Demographic variables such as age, gender, marital status, education, labour force

and health status

- Household disposable income
- Proportion of wealth in equity markets: direct equity as well as superannnuation.
- Worsening of finances between 2002 and 2006.
- Events such as death of spouse, or injury between 2006 This variable also captures events between the years that may have an impact on saving.
- Home-ownership. The thresholds facing home-owners and renters are different and may imply a differentialy incentive to save.

In addition, we also use the value of non housing equity and the proportion of wealth in equity markets in 2006 because the period between 2006 and 2010 was a period of the financial crisis, and larger investments in equity in 2006 might have a bearing on financial savings reported in 2010. Similarly, events between 2006 and 2010 may have a bearing on the financial savings reported in 2010, are therefore used to finally estimate the average treatment effect.

3.3 Matching estimator of causal effect

The propensity score estimator allows the estimation of p_i, the estimated probability of individual i being in the treatment group (group T) i.e. between the full and part pension thresholds before the taper-rate reform. Let p_j denote the predicted probability of individual j being in the control group (group C), and S denote the savings.

A standard matching estimator of the causal effect of the taper-rate reform can be written as

$$\mu = \sum_{i \in T} (S_i - \sum_{j \in C} g(p_i, p_j) S_i)$$

Here $g(.)$ is a function assigning the weights to be placed on the comparison household j while constructing the counterfactual for the treated household i. The weights used in our estimation are the inverse of the variances. We employ a one-one matching estimator with replacement.

4 Data

The data used for the analysis comes from the Household, Income and Labour Dynamics in Australia (HILDA) panel data-set, which is a nationally representative random-sample survey of private households in Australia.[5] It encompasses approximately 13,000 individual respondents, living in more than 7,000 households. The data has been collected each year from 2001, and as of 2013 is in its thirteenth year.[6]

We are interested in estimating the impact of the potential receipt of the age-pension on the saving behaviour of households. We expect households to save less if by doing so they stand to receive a greater amount of transfer income. We also expect that households may find it beneficial to trade-up their owner-occupier homes as the family home is exempt from the assets test. These incentives are more relevant to those closer to the eligibility age for the pension, as not only can these people form better predictions of their post-retirement assets, but also are more likely to pay greater attention to the receipt of the age-pension. We therefore focus on individuals who are between the age of 50 and 64 in 2006, just prior to the 2007 taper rate reform and are not already on an age-pension.[7] Household age is that of the oldest person in the household.

The income and wealth variables are that of the household the respondent belongs to. Detailed wealth estimates of households were collected in the second wave in 2002, sixth wave in 2006, and the tenth wave in 2010. These include total assets, total housing equity (value minus debt on the home), and non-housing equity (which we calculate as total non-housing assets minus non-housing debt minus debt on the home). Total assets include financial wealth, business wealth, and real estate including the family home.[8] Since wealth in all forms except the family home is considered towards the assets test, we use non-housing equity as the main wealth variable. Saving is defined as the difference in the real value of housing and non-housing equity between 2002 and 2006 (pre-reform savings) and between 2006 and 2010 (post-reform savings). Non-housing wealth is deflated by the consumer price index, and housing wealth by the house-price index.

We classify households in 2006, just before the taper-rate reform, into three types. The first category is those households with non-housing assets lower than the full-pension

[5] For more details see `http://www.melbourneinstitute.com/hilda/`

[6] The data is released with a two-year lag. This implies that the 2013 data will get released in 2015.

[7] We further narrow the analysis to individuals in the 58-64 age category as well.

[8] Detailed definitions are presented in the Appendix.

Table 1 Descriptive statistics of the estimation sample (in 2006)

This table reports the descriptive statistics of the sample as of 2006. Median values are reported for the wealth and income variables. The rest are percentages. For example, the median home-equity of those in between the full and part pension thresholds (PP) was $320,000 while that of the group above the part pension thresholds (NP) was $430,000. The table also shows that of those in the PP group, 65 percent were men, while 70 percent were men of those in the NP group.

	Pension eligibility		
	FP	PP	NP
Number of observations	288	130	281
Home equity ($'000)	210	320	430
Non-housing equity ($'000)	73	368	931
Income ($'000)	40	66	76
Proportion of equity	0	1	2
Proportion of superanuation	45	53	36
Proportion of interest income	0	0	0.1
Age	58	58	58
Male	55	65	70
Couple family	32	57	49
Education: Certificate	42	46	46
Education: School	50	36	31
Education: Graduate	8	18	23
Employed: Full time	40	52	56
Employed: Part time	19	26	18
Employed: Not working	41	21	26
Health problem	50	37	26
Home-owner	72	93	94
Area: City	58	58	57
Area: Regional	39	39	40
Area: Remote	2	3	2
Event: Worsening of finances	9	10	11
Event: Death of spouse	1	1	1
Event: Moved home	19	11	13
Event: Injured	26	21	16

Source: HILDA

threshold (and thus eligible for full-pension and labelled as FP). The second category is those households between the full and part-pension threshold (and thus eligible for part-pension and labelled as PP), and the third cateogry of households with wealth greater than the part-pension threshold (and thus not eligible for pension and labelled as NP). The non-housing equity of the households, home-ownership and family composition (lone person vs. couples) are used to correctly identify the applicable threshold for the year in consideration. The applicable thresholds in 2006 are presented in the Appendix. The households between the full and part pension thresholds (PP) consist the treatment group. Table 1 provides the descriptive statistics of the sample.

A large fraction of households is below the full pension thresholds applicable in 2006. This implies that were they to become age-eligible in 2006, they would qualify for the full pension through the assets-test thresholds. The household income of this group is also considerably lower than that of those above the threshold. A large proportion, almost 60 percent are not employed full time, which may explain the lower income. More than 70 percent are home-owners, and about half of them live in cities. Home-ownership is higher in the latter two groups.

The median value of home-equity of those in the treatment group, that is between the full and part pension thresholds, is about a $100,000 lower than that of those above the part pension thresholds. Their non-housing equity is much lower than that of the latter group. The distributions of the non-housing equity of the three groups are presented in Figure 2. The differences in the distributions underscores the need for a matching model that will lead to a comparison of households that are similar in their wealth and demographic details.

Figure 3 presents the distribution of non-housing equity in 2006 and 2010 for the group that is between the thresholds and above the thresholds in 2006. We find that the distribution of non-housing equity is flatter at the center for the group between the thresholds, but has more mass at the tails. The distribution has not changed much for the group that has assets beyond the thresholds. This suggests that there are incentives that operate for the group that is closer to the thresholds and more likely to be recipients of the age pension. In fact, 42 percent of the between the thresholds group that came of access age in 2010 was a recipient of the age pension in 2010. The corresponding number for the greater than thresholds group was 22 percent.

We further investigate the impact of the potential receipt of the age-pension on the sav-

Figure 2 Non-housing equity

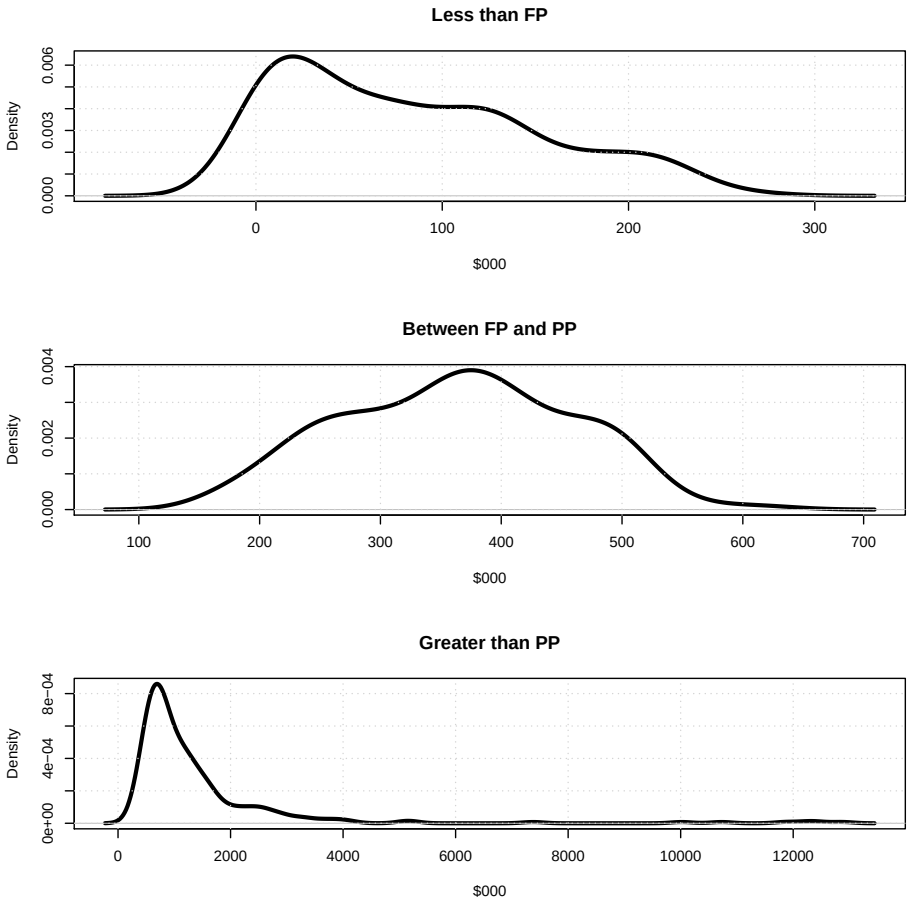


Figure 3 Non-housing equity: 2006 and 2010

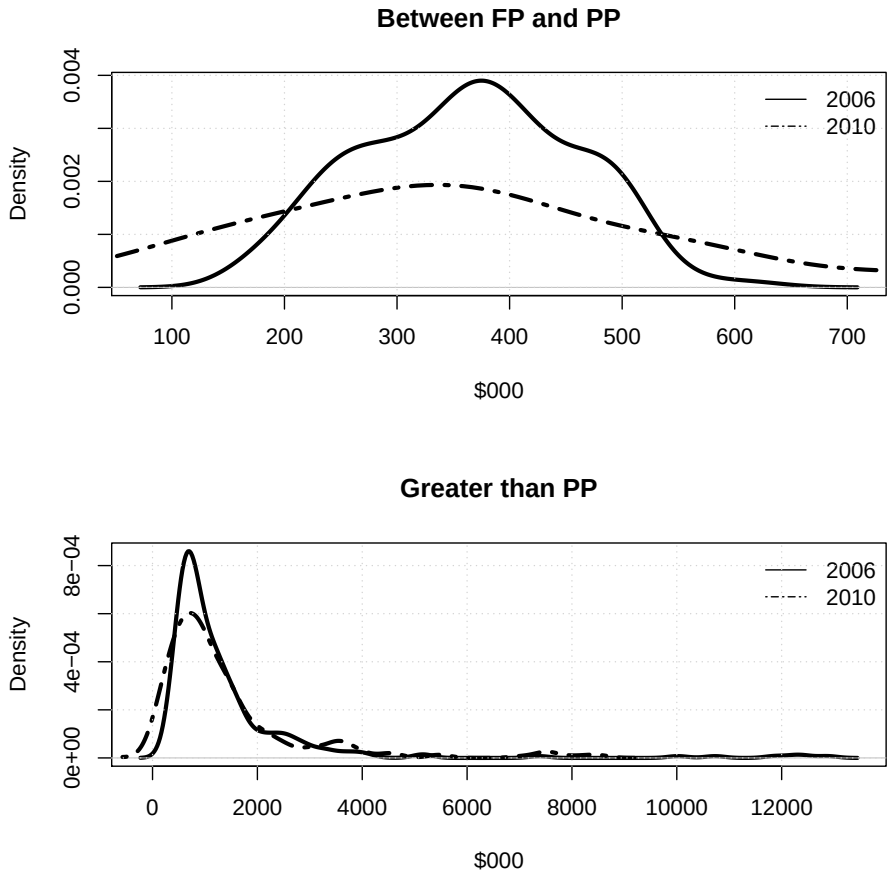


Table 2 Regression: Financial and house savings

	Financial saving			Saving on the home		
	Estimate	Std. error		Estimate	Std. error	
2010-2006						
Less than FP threshold	121,501	55,664	**	-22,386	26,554	
Between FP and PP thresholds	141,296	62,375	*	16,038	29,756	
N	696			696		
R-square	0.12			0.06		
2006-2002						
Less than FP threshold	-333,149	52,377	***	133,302	41,393	***
Between FP and PP thresholds	-286,709	56,257	***	84,089	44,459	*
N	697			697		
R-square	0.15			0.16		

Notes: FP: full-pension, PP: part-pension
Significance code: ***1%, **5%, *10%

ing behaviour of households with two regressions. Table 2 we show the results of an OLS regression on savings in 2010 (and 2006) of households in the three categories described above. The regression controls for demographic characteristics and income of the households in 2006 (and 2002), as well as events that occurred between 2006 and 2010 (and 2002 and 2006).

We find that those who have assets below the full pension threshold and those who have assets between the two thresholds save more than those who have assets above the full-pension thresholds between 2006 and 2010. This result however does not hold for saving on the family home. Even more importantly, the very same groups were actually dissaving their financial wealth but saving on the family home in the period between 2002 and 2006. This suggests that the reduction in the assets-test taper rate may have resulted in an increased incentive to save, especially for those more likely to be on the pension after the taper-rate reform. The regression results are correlations and do not have a causal interpretation. As well, the group in 2010-2006 regression is older than the group in 2006-2002 which may explain some of the results. We turn next to a more rigourous analysis of the saving behaviour by first testing the reliability of the matching model and then presenting our main results.

5 Testing the reliability of the matching model

We use demographic and wealth data in 2002 and events between 2002 and 2006 to estimate the propensity score model as we expect these to influence if the household has non housing equity between the two thresholds in 2006. The variables include include age, age-square, gender, labour-force status, education, marital status, area of residence, income, proportion of equity investments in total non housing equity, proportion of superannuation investments in total non housing equity. The events include if the household moved residence, or reported an injury, death of spouse or worsening of finances between 2002 and 2006.[9]

In calculating the average treatment effect of the taper-rate change, we match on the propensity score and also additionally match on variables in 2006 and events between 2006 and 2010 that can have a bearing on household savings reported in 2010. These include income, non housing equity, proportion of equity investments and proportion of superannuation investments in non housing equity, and proportion of interest income in total income, all measured in 2006 prior to the reform. The households are grouped according to gender, marital status, home-onwership, area of residence, health, occurrence of events such as injury or worsening of finances residential mobility in 2006. The wealth and income variables are matched within each such grouping.

The matching procedure will result in a reliable method for estimating the impact of the taper-rate change if, the pre-reform variables are balanced between the treatment and control groups. Lack of balance points to a possible mis-specification of the propensity score estimation. It is therefore important to verify if the balancing condition is satisfied by the data.

In Table 3 we report several balance measures for income and wealth variables and not for the discrete variables, as the former are matched within each group of the combination of discrete variables. The first two, in columns (1) and (2) are the means of the treated and control group respectively after matching.

The third, reported in column(3), is the standardised difference as described by (Smith and Todd, 2005). For example, the standardised bias for the income variable is defined as the difference in means between the treatment households (group T) and the appropriately matched comparison households (group C) scaled by the average variances of the income

[9]The results of the first-stage probit model are presented in Table A.3.

Table 3 Balancing tests

This table reports the balance statistics for the various variables after the matching exercise.
Std. diff is the standardized difference defined as the difference in means between the treatment households (group T) and the appropriately matched comparison households (group C) scaled by the average variances of the income variable in the two groups.
t-test p val is the p value from a paired t-test between the treated and control households.
Var ratio is the variance ratio of treatment over control groups.
KS statistic reports the statistic from a test of a significant difference across the entire distribution.

	(1) Mean Treated	(2) Mean Control	(3) Std diff	(4) Var ratio	(5) t-test p-val	(6) KS bootstrap p-val
PSM (2002)						
Age	53.74	54.16	-17.67	0.95	0.13	0.68
Age-square	2893.6	2939	-17.81	0.95	0.13	0.68
log (Income)	10.94	10.81	25.97	0.63	0.05*	0.24
Prop. equity	0.09	0.09	0.81	0.66	0.95	0.31
Prop. superannuation	0.57	0.48	28.85	0.79	0.06*	0.06*
Additional covariates (2006)						
log (Income)	10.94	10.95	-0.27	0.97	0.97	0.49
log (Non housing equity)	12.81	12.74	25.12	0.05	0.55	0.00*
Prop. equity	0.15	0.08	8.79	32.51	0.34	0.08*
Prop. superannuation	0.54	0.47	25.26	1.07	0.04*	0.13
Prop. interest income	0.009	0.004	20.57	1.85	0.05*	0.19

variable in the two groups. The lower the standardised difference, the more balanced or similar the treatment and comparison groups will be in terms of the variable under consideration. We are in range of 20 for most our variables, and a little over 20 for some.

In column (4) we report the variance ratio of treatment over control, which should be equal to 1 if there is perfect balance. We find that the Var ratio is close to 1 for most of the variables, except most notably for non housing equity where it is as low as 0.05.

Further, for each variable entering the propensity score mode, we perform a formal paired t-test between the treated and control households to confirm that no significant difference exists. The result is reported in column (5). The p-value of the bootstrapped Kolmogorov-Smirnov (KS) test statistic which tests for a significant difference across the entire distribution is reported in column (6).

We find that several of our variables report a t-statistic p-value less than 0.10. However, when tested for a significant difference in the distribution through the KS-statistic, we find a statistically significant difference between the two groups for three variables: proportion of wealth in superannuation in 2002, non housing equity and proportion of wealth in equity in 2006.[10] The same variables report a reasoanble standardized difference, and two of the variables, non housing equity and proportion of equity in total wealth in 2006 both report t-stats that are not signficant.

From the tests reported in this section we find that no variable consistently performs poorly on all tests. The treatment and control groups are a function of the non-housing equity, and by design it is not possible to have complete balance on this variable. While we have not achieved perfect balance on our variables we are satistified with the balance properties from the propensity score matching exercise, and proceed to discuss the results.

6 Results

The results on saving from the matching estimation are presented in Table 4.

Column (1) reflects the average treatment effect on the treated from the preferred PSM model described in the previous section. This is the sample for which the balance statistics were reported earlier. The coefficient described here is the causal impact of the change

[10] We conduct several robustness checks to see if we can overcome this problem.

Table 4 Results: Savings

	log(Financial saving)	log(Saving on the home)
Estimate	0.018**	0.02
SE	0.009	0.02
p-val	0.04	0.18
No. of observations	699	699
No. of treated obs.	130	130
No. of matched treated obs.	108	108
No. of matched control obs.	108	108
Significance code: ***1%, **5%, *10%		

in the assets test taper rate on the savings of the treated group relative to the control group. If assets test inhibit savings, then we should find that those most constrained by the thresholds prior to the reform, are able to save more once the taper rate has been relaxed.

The estimates show that the average causal effect of the taper-rate reform is a differential increase in financial saving of 2 percent significant at the 5 percent level. When the absolute values are used instead of log values, the differential saving is $110,926, significant at the 5 percent level. This is about 16 percent of the average non housing equity of the households, and 31 percent of the average non housing equity of households in the treatment group. As the samples are matched across different categories of households, this estimate takes into account the heterogeneity of households and the events that can potentially affect saving. This finding suggests that the means-tests design does indeed constrain saving. A relaxation of the taper-rate allowed households prior to the age pension to accumulate more non-housing wealth.

We estimate the same model, but for saving on the family home. As the family home is exempt from the means-test, relaxation of the taper-rate should not have any impact on the saving made on the family home. We find that, as reported in column (2), the ATT for the saving on family home is 2 percentage ponints, but is not statistically significant.

We then focus specifically on various groups identified as more vulnerable to assets test thresholds. The results are reported in Table 5. We first focus on the age group 58-64. This group is closer to retirement than the group between 55-57. We conduct the same matching experiment and find that the average treatment effect for this age group is 1.5 percent ($100,847 in absolute terms), significant at the 10 percent level. This implies that those between the two thresholds in the 58-64 age group in 2006, had higher financial

Table 5 Results: Savings of various groups

	log(Financial saving)		
	(1)	(2)	(3)
	Age 58-64	Married	Poor health
Estimate	0.015*	0.016	0.007
SE	0.008	0.011	0.008
p-val	0.06	0.14	0.40
No. of observations	410	478	264
No. of treated obs.	72	108	48
No. of matched treated obs.	55	86	38
No. of matched control obs.	55	86	38

Significance code: ***1%, **5%, *10%

savings relative to those above the two thresholds. It appears that the disincentive to asset accumulation is thus greater for those who are closer to the retirement age.

Cobb-Clark and Hildebrand (2011) point out that the incentives embedded in the age pension are strongest for single households in poor health, and do not influence the asset allocation of couple or healthy households. We therefore focus on two categories of households: those married in 2006 and those who claim to have a long-term health problem in 2006. Our matching exercise shows that financial savings increased for married households post the taper rate reform but is not statistically significant. The results are not significant for households in poor health as well. However, the number of treated observations in both these groups is very small consistent with Table 1 which shows that a large number of less healthy households are actually below the full pension threshold. This implies that households in poor health are more likely to find themselves below the full pension threshold and therefore quite a distance away from the threshold that makes them ineligible for a pension. If the threshold is not binding, then it is not likely to exert a disincentive effect on savings. This is consistent with the findings of (Hurst and Ziliak, 2004) who show that the assets test do not affect savings of female headed households with children, a disproportionately poor population.

7 Robustness checks

In this section we address alternative explanations that might explain increase in financial saving of those between the two thresholds, rather than the change in the taper rate. Two alternatives can be offered:

Table 6 Placebo: log(Financial Saving)

	Placebo treatment
Estimate	-0.14
SE	0.13
p-val	0.28
No. of observations	699
No. of treated obs.	130
No. of matched treated obs.	114
No. of matched control obs.	114
Significance code: ***1%, **5%, *10%	

1. The effect is not peculiar to the treatment group.
2. The results are sensitive to the matching strategy.

7.1 Placebo treatment

In order to check the robustness of our findings, we consider the effect of the policy on a group that does not directly benefit from the same. We randomly choose 130 observations and assign them as the treatment group. We conduct the same matching exercise, and estimate the average savings of the two groups. Results are presented in Table 6. We are not able to reject the null that there is no difference in financial saving of the treatment and control groups.

We perform this experiment a 1000 times. We find that we are not able to reject the null of no difference in saving at a significance level of 10 percent in only 5 percent of the experiments. This reassures us that the previously measured policy effect for potential age pension recipients is not spurious.

7.2 Sensitivity to the matching technique

We first check for the sensitivity of the results by making a slight change to the originial specification (Dehejia, 2005). The results are presented in Table 7. The first change, presented in column (1) is a 1-2 matching procedure instead of the 1-1 matching in the original model. In the 1-2 matching procedure, two control observations are chosen for every treated observation.

Table 7 Robustness checks: log(Financial Saving)

	(1) 1-2 matching	(2) Modified PSM	(3) Genetic matching	(4) GM with caliper=0.20
Estimate	0.02**	0.018**	0.017*	0.01***
SE	0.01	0.009	0.009	0.002
p-val	0.04	0.04	0.06	0.00
No. of observations	699	699	699	697
No. of treated obs.	130	130	130	130
No. of matched treated obs.	108	108	108	12
No. of matched control obs.	196	108	108	12

Significance code: ***1%, **5%, *10%

In the second model we add a quadratic income term to the orginial list of covariates, the results of which are presented in column (2).

A concern in the estimations above is that we have not achieved match balance on the non housing equity variable, and this may imply that differences in non housing equity are so large that they drive the differences in saving. Towards this end, we conduct a genetic matching exercise as described in (Diamond and Sekhon, 2012), and present the results in column (3). This is a method of multivariate matching, that uses an evolutionary search algorithm to determine the weight each covariate. We are unable to achieve perfect balance on some of the variables including non-housing equity even with genetic matching: the KS bootstrapped p-value continues to be significant at the 1 percent level. The variance ratio has improved from 0.05 (in Table 3) to 0.06, and the standardized difference is now -19.74.[11] It is however important to remember that because the treatment group is determined by the value of the non-housing equity, we will not be able to achieve complete balance on this variable. We have however managed to compare as similar observations as possible, including the wealth variable.

While we have improved our balance statistics with the use of genetic matching, we are still not able to achieve match balance on the proportion of superannuation in total wealth. It could be argued that it is the differences in superannuation wealth that are the drivers of the saving behaviour. We therefore conduct another matching exercise in which we make two changes. First we use the square of the proportion of superannuation in total non-housing wealth as our match variable. Second use a caliper of 0.2, which implies that all that all matches not equal to or within 0.20 standard deviations of each covariate

[11]The full set of balance statistics is presented in Table A.4 in the Appendix.

are dropped. Results are presented in column (4). We find that while our estimate has remained the same, the standard error has decreased dramatically, improving the significance of our result. Even though we retain only 12 matched observations, we are able to achieve match balance on all the variables of interest in 2006 (See Table A.5 in the Appendix for full match balance statistics.)

The results of the four modified models are comparable to that of the original preferred model, both in magnitude and statistical significance. This is further illustration of the appropriateness of propensity score specification, and the robustness of our results.

8 Conclusion

This paper contributes to the literature on the impact of means-testing on saving behaviour. We argue that a relaxation of the taper rate of the age pension assets test in Australia allows for larger saving of those who are most constrained by the thresholds.

We use data from the Household, Income and Labour Dynamics in Australia (HILDA) database on the 50-64 age group in 2006, a few years away from the pension access age of 65. Part of this group becomes age-eligible in 2010, and saving between the years 2010 and 2006 will have an impact on eligibility through the assets-test. We exploit the change in the taper rate of the assests test on the age pension in 2007 to evaluate if those in the 50-64 age group most constrained by the test prior to the reform increased their savings relative to those not constrained by the assets test.

We find that the those constrained by the thresholds had 1.8 percent more financial savings in 2010 than the control group. This group saved $110,926 more than those of the same age but not earlier constrained by the assets test. This is 16 percent of the average wealth of the overall; sample in 2006, and 31 percent of the average financial wealth of those constrained by the assets test prior to the reform.

Our results are consistent with Maynard and Qiu (2009) who find that crowd-out of savings from the Medicaid program in the US is surprisingly large for moderately wealthy households. The results are also consistent with research specifically evaluating the impact of an increase in asset limits (Powers, 1998; Hurst and Ziliak, 2006; Sullivan, 2006). This result has important implications for pension policy design as countries move towards toward greater targeting of benefits to reduce their fiscal burdens.

The income test in Australia is so structured that certain assets are deemed to have a specified rate of return, regardless of the actual rate of return earned on the market. Poor households, i.e. those below the thresholds should have an incentive to dissave, or reallocate wealth in to the family home, to minimize the 'deemed rate of return' on the income test. We do not see large draw-downs in wealth by such households. This may be because deeming rules may lead households to concentrate financial wealth in assets of a particular kind. Means-test will then not only have an impact on allocation between housing and non-housing assets, but also between financial assets of various kinds. This is left for future research.

References

Cho, Stanley, and Renuka Sane (2012): "Means-Tested Age Pension and Homeownership: Is There a Link?," *Macroeconomic Dynamics*.

Cobb-Clark, Deborah A., and Vincent A. Hildebrand (2011): "Portfolio Allocation In The Face Of A Means-Tested Public Pension," *The Review of Income and Wealth*, 57(3), 536–560.

Dehejia, Rajeev (2005): "Program evaluation as a decision problem," *Journal of Econometrics*, 125(1-2), 141–173.

Dehejia, Rajeev H., and Sadek Wahba (2002): "Propensity score matching methods for nonexperimental causal studies," *The Review of Economics and Statistics*, 84(1), 151–161.

Diamond, Alexis, and Jasjeet S. Sekhon (2012): "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies," *Review of Economics and Statistics*, (forthcoming).

Gittleman, Maury (2011): "Medicaid and Wealth: A Re-Examination," *The B.E. Journal of Economic Analysis & Policy*, 11(1).

Gruber, Jonathan, and Aaron Yelowtiz (1999): "Public Health Insurance and Private Savings," *The Journal of Political Economy*, 107(6), 1249–1274.

Harmer, Jeff (2008): "Pension Review: Background Paper," Department of Families, Housing, Community Services and Indigenous Affairs.

Heckman et. al. (1997): "Matching as an econometric evaluation estimator: evidence from evaluating a job training programme," *Review of Economic Studies*, 64, 605–654.

Henry, Ken (2009): "Australia's Future Tax System - The Retirement Income System: Report on Strategic Issues," Australia Treasury.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes (1994): "Expanding the Life-Cycle model: Precautionary Saving and Public Policy," *The American Economic Review*, 84(2), 174–179, Papers and proceedings of the Hundred and Sixth Annual Meeting of the American Economic Association.

——— (1995): "Precautionary Saving and Social Insurance," *The Journal of Political Economy*, 103(2), 360–399.

Hulley et. al. (2013): "Means-Tested Public Pensions, Portfolio Choice and Decumulation in Retirement," *The Economic Record*.

Hurst, Erik, and James Ziliak (2004): "Do Welfare Asset Limits Affect Household Saving? Evidence from Welfare Reform," Working Paper 10487, NBER.

——— (2006): "Do Welfare Asset Limits Affect Household Saving? Evidence from Welfare Reform," *The Journal of Human Resources*, 41(1), 46–71.

Kudrna, George, and Alan Woodland (2008): "A General Equilibrium Analysis of the Australian Means-Tested Age Pension," Working paper, University of New South Wales.

Maynard, Alex, and Jiaping Qiu (2009): "Public insurance and private savings: who is affected and by how much?," *Journal of Applied Econometrics*, 24(2), 282–308.

Neumark, David, and Elizabeth Powers (1998): "The effect of means-tested income support for the elderly on pre-retirement saving: evidence from the SSI program in the US," *Journal of Public Economics*, 68, 181–206.

Powers, Elizabeth T. (1998): "Does means-testing welfare discourage saving? evidence from a change in AFDC policy in the United States," *Journal of Public Economics*, 68, 33–53.

Rosenbaum, Paul R., and Donald R. Rubin (1983): "The central role of the propensity score in observational studies for causal effects," *Biometrika*, 70, 41–55.

Sefton, James, and Justin van de Ven (2009): "Optimal design of means tested retirement benefits," *Economic Journal*, 119.

Sefton, James, Justin van de Ven, and Martin Weale (2008): "Means Testing Retirement Benefits: Fostering equity or discouraging savings?," *The Economic Journal*, 118, 556–590.

Smith, Jeffrey A., and Petra E. Todd (2005): "Rejoinder," *Journal of Econometrics*, 125, 365–375.

Sullivan, James X. (2006): "Welfare Reform, Saving, and Vehicle Ownership: Do Asset Limits and Vehicle Exemptions Matter?," *The Journal of Human Resources*, XLI(1).

Tran, Chung, and Alan Woodland (2010): "Means-Tested Pension as Social Insurance," mimeo, University of New South Wales.

——— (2011): "Trade-Offs in Means Tested Pension Design," *ANUCBE School of Eoc-nomics Working Paper, Australian National University, Canberra*.

Appendix

Table A.1 Means test rules for the age-pension

	Single	Couple
Asset-test		
Homeowners		
Full pension cut-off	$161,500	$229,000
Part pension cut-off	$330,000	$509,500
Non home-owners		
Full pension cut-off	$278,500	$346,000
Part pension cut-off	$447,000	$626,500
Income test		
Full pension cut-off (p.f.)	$128	$240
Part pension cut-off (p.f.)	$1,557.7	$2,602

Note: p.f.: per fortnight
Source: Centrelink

Table A.2 Variable definitions

Variable	Definition
Housing assets	Value of family home
Non-housing assets	Total of equity investments, cash investments, trusts, own bank accounts, joint bank accounts, children's bank accounts, redeemable insurance policies, retirees superannuation, non-retirees superannuation, other property value, collectibles , businesses, vehicles
Housing equity	Housing assets − mortgage on home
Non-housing equity	Non-housing assets − debt − mortgage on home
Net-worth	Housing + non-housing assets − total debt
Disposable income	Total household income − taxes
Education	Grad if Highest education level achieved answered as Bachelor or hounours, or Postgrad Certificate if Highest education level achieved answered as Grad diploma, grad certificate, Adv diploma, diploma, Cert III or IV or Cert I or II or Cert not defined School if Highest education level achieved answered as Year 12 or Year 11 and below
Long-term health condition	1 if long-term health condition, disability or impairment
Areas	Major City if Remoteness Area answered as Major City Regional if Remoteness Area answered as Inner Regional Australia or Outer Regional Australia Remote if Remoteness Area answered as Remote Australia or Very Remote Australia

Table A.3 First state probit regressions: Pr(between thresholds in 2006)

	Estimate	Std. Error	z value	
(Intercept)	-60.6155	30.4350	-1.99	**
Age	2.2093	1.1299	1.96	**
Age square	-0.0206	0.0105	-1.97	**
Male	-0.1383	0.1356	-1.02	
Education: Grad	-0.0477	0.1692	-0.28	
Education: School	-0.0839	0.1301	-0.64	
Never married	-0.3767	0.2772	-1.36	
Separated	-0.3394	0.1839	-1.85	*
Widowed	-0.2585	0.3241	-0.80	
Health problem	0.0690	0.1449	0.48	
log(income)	0.0103	0.3431	0.03	
log(income square)	2966.6014	18653.3444	0.16	
Employed PT	-0.1054	0.1687	-0.62	
Not working	-0.3112	0.1819	-1.71	*
Unemployed	-0.0717	0.4381	-0.16	
Financial saving (2006-2002)	-0.0000	0.0000	-2.88	***
Prop. of equity	-0.1336	0.3511	-0.38	
Prop. of superannuation	-0.0014	0.0104	-0.13	
Home-owner	0.5159	0.2019	2.56	***
Event: Moved	0.0023	0.1526	0.02	
Event: Injured	0.3397	0.1342	2.53	***
Event: Spouse death	-0.3423	0.6102	-0.56	
Event: Financials worsened	0.2315	0.2270	1.02	
Area: Regional	0.0897	0.1220	0.74	
Area: Remote	0.2224	0.4103	0.54	
Log-likelihood	-313.39			
N	699			

Table A.4 Balancing tests: Genetic matching

This table reports the balance statistics for the various variables after the genetic matching exercise.
Std. diff is the standardized difference defined as the difference in means between the treatment households (group T) and the appropriately matched comparison households (group C) scaled by the average variances of the income variable in the two groups.
t-test p val is the p value from a paired t-test etween the treated and control households.
Var ratio is the variance ratio of treatment over control groups.
KS statistic reports the statistic from a test of a significant difference across the entire distribution.

	(1) Mean Treated	(2) Mean Control	(3) Std diff	(4) Var ratio	(5) t-test p-val	(6) KS boot-strap p-val
PSM (2002)						
Age	53.74	53.97	-9.81	1.14	0.41	0.67
Age-square	2893.6	2917.8	-9.55	1.13	0.43	0.67
log (Income)	10.94	10.82	22.83	0.63	0.10	0.15
Prop. equity	0.09	0.11	-14.43	0.75	0.26	0.51
Prop. superannuation	0.57	0.48	30.12	0.97	0.03*	0.07*
Additional covariates (2006)						
log (Income)	10.94	10.98	-5.99	0.88	0.61	0.37
log (Non housing equity)	12.81	12.86	-19.74	0.06	0.57	0.00*
Prop. equity	0.15	0.09	6.66	25.19	0.47	0.10
Prop. superannuation	0.54	0.47	22.92	1.19	0.04*	0.01*
Prop. interest income	0.009	0.004	21.00	1.83	0.04*	0.2

Disclaimer

This paper uses 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 Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this thesis, however, are those of the authors and should not be attributed to either FaHCSIA or the Melbourne Institute.

Table A.5 Balancing tests: Genetic matching with caliper

This table reports the balance statistics for the various variables after the genetic matching exercise with a caliper of 0.25.
Std. diff is the standardized difference defined as the difference in means between the treatment households (group T) and the appropriately matched comparison households (group C) scaled by the average variances of the income variable in the two groups.
t-test p val is the p value from a paired t-test etween the treated and control households.
Var ratio is the variance ratio of treatment over control groups.
KS statistic reports the statistic from a test of a significant difference across the entire distribution.

	(1) Mean Treated	(2) Mean Control	(3) Std diff	(4) Var ratio	(5) t-test p-val	(6) KS boot-strap p-val
PSM (2002)						
Age	53.08	53.67	-24.78	1.20	0.50	0.66
Age-square	2822.9	2884.3	-24.3	1.21	0.51	0.66
log (Income)	10.93	10.88	14.51	1.15	0.57	0.99
Prop. equity	0.04	0.21	-235.8	0.06	0.06*	0.07*
Prop. superannuation	0.58	0.46	39.47	1.13	0.25	0.49
Additional covariates (2006)						
log (Income)	11.07	11.07	0.90	.86	0.93	0.99
log (Non-housing equity)	12.87	12.92	-18.04	0.46	0.30	0.02*
Prop. equity	0.02	0.02	-2.07	1.07	0.92	0.97
Prop. superannuation square	0.39	0.39	2.6	0.95	0.74	0.98
Prop. interest income	0.001	0.001	-45.13	0.11	0.63	0.42