



UNSW Business School

Working Paper

Never Stand Still

Business School

Economics

UNSW Business School Research Paper No. 2016 ECON 16

Policy Choice and Product Bundling in a Complicated Health Insurance Market: Do People get it Right?

Nathan Kettlewell

This paper can be downloaded without charge from
The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=2858699>

Policy choice and product bundling in a complicated health insurance market: Do people get it right?*

Nathan Kettlewell[†]
University of New South Wales

October 2016

Abstract

This paper evaluates health insurance policy selection and how this interacts with product bundling by using a discrete choice experiment closely calibrated to the Australian private health insurance market. The experimental approach overcomes some limitations of revealed preference research in this area. The results indicate that consumers are likely to make choices that violate expected utility theory, use heuristic decision strategies, and over-insure relative to minimising out-of-pocket costs. Decision quality is significantly lower when choosing a bundled hospital/ancillaries health insurance policy (compared to stand-alone ancillaries cover), which is the policy type most consumers purchase in Australia.

Keywords: health insurance, heuristics, choice consistency, discrete choice experiment, latent class logit

JEL: I13; D81; D03

*I am grateful to Mikel Berdud, Denzil Fiebig, Ben Greiner, Greg Kaplan, Chengxiang Tang and Hong Il Yoo for useful advice and feedback related to this research. I am especially thankful to Denise Doiron for her guidance and support. This paper has also benefited from feedback by participants at the 13th CEPAR Summer Workshop on the Economics of Health and Ageing and the 89th Health Economists' Study Group Meeting. I thank the University of New South Wales Business School for financial support through their research network funding.

[†]Economics, UNSW Business School, University of New South Wales, Sydney, NSW, 2052, Australia. Email: n.kettlewell@student.unsw.edu.au

1 Introduction

The traditional expected utility model for insurance choice assumes that agents are perfectly informed about the effect of different options on the distribution of expected out-of-pocket costs and choose policies to maximise utility (e.g. Arrow, 1963; Rothschild & Stiglitz, 1976). However, recent evidence questions the underlying assumptions of this model in the case of health insurance. For example, decisions appear to be heavily influenced by choice complexity, information and comparison frictions (Kling et al., 2012; Handel & Kolstad, 2015), inertia (Handel, 2013; Handel & Kolstad, 2015; Polyakova, 2016) and heuristics (Ericson & Starc, 2012; Bhargava et al., 2015). Surveys asking people to answer questions about health insurance contracts demonstrate limited comprehension (e.g. Hibbard et al., 1998; Loewenstein et al., 2013). Experimental research suggests that comprehension is even lower when consumers are presented with larger choice sets (Hanoch et al., 2009; Barnes et al., 2014).

Understanding how consumers choose health insurance and the quality of those choices is crucial for interpreting preferences derived from models that assume agents are traditional expected utility maximisers. In particular, welfare measures may be severely biased if the underlying behavioural assumptions fail. Choice quality also has implications for how (and whether) regulators should intervene in the choice environment. The prevalence of various interventions in insurance markets to simplify the comparison of policies suggests that policy makers are sensitive to this¹.

The primary aims of this paper are to evaluate choice quality (i.e. consistency with expected utility theory and probability of maximising own welfare) and how this is influenced by an important form of complexity – the bundling of different types of health insurance. To achieve this I use results from a discrete choice experiment (DCE) calibrated to the Australian private health insurance market. I first test for violations of expected utility

¹One example is the metallic coding (e.g. silver, gold, platinum) given to policies under the Affordable Care Act to indicate copay rates. Similarly, in Australia policies are labelled as either public, basic, medium or top depending on the number of included health services.

theory. In particular, I exploit the annual caps on benefits for ancillary health services, which create limits on any rational consumer's willingness to pay (WTP) for additional cover. By exploiting these caps I can evaluate choices without knowing risk preferences or expected health costs. Using the preferences elicited from the DCE I then predict policy choices and compare to the choice distribution that would minimise out-of-pocket costs (i.e. maximise expected value). I also discuss the importance of narrow focus choice strategies and what the results imply about preferences for different policy features.

All Australians can receive free treatment in a public hospital through a system of public insurance known as Medicare. Parallel private health insurance can be purchased to subsidise treatment in a private hospital or ancillary health expenses excluded from Medicare. The main ancillaries people insure against are dental, optical and physical health services (e.g. physiotherapy and chiropractic). Several reviews have raised concerns that the large number of features and complicated copay structures for health insurance has a detrimental effect on consumer choices (Deloitte Access Economics, 2012; PHIAC, 2013; ACCC, 2015). However, the effects of complexity on choice quality remain unmeasured. Private hospital and ancillaries health insurance can either be purchased as separate products or a single combined product. Around 85% of policies purchased are combined policies. While bundling can reduce transaction costs, it may also increase complexity and reduce choice quality.

Understanding how product bundling interacts with choice is important for discussing reforms to the regulation of insurance and the choice environment. This issue is not uniquely Australian – most developed countries operate private health insurance markets to cover gaps in public insurance for both in- and out-of-hospital services. In Canada consumers can only purchase supplementary cover, which is similar to Australian ancillaries insurance, and some commentators have argued for expanding insurance to publicly funded services. If insurance was expanded, then it would be important to know how the new insurance products might interact with choice quality.

The DCE involves respondents from an online panel choosing preferred health insurance

policies. Respondents were sampled to match closely the distribution of characteristics within the Australian population and randomly assigned to two treatments (i.e. ancillaries insurance or combined hospital/ancillaries insurance).

Because the choice task is hypothetical, I can design the experiment to exploit structural features of the health insurance market. This leads me to focus particularly on preference for ancillary health insurance. As mentioned above, benefits for ancillaries are subject to annual limits (caps). These provide boundaries for rational consumers' WTP, regardless of risk preferences or expected health care costs. I first show that between 16-23% of respondents directly violate expected utility theory by choosing a financially dominated policy when it is included in a choice task. These violations are partly explained by a narrow focus heuristic, where respondents with high expected utilisation of a particular health service focus narrowly on maximising coverage for that feature.

I next estimate latent class models for policy choice. Following Burton and Rigby (2009) I use the latent classes to identify potential non-compliers (i.e. respondents who do not engage with the DCE) as well as account for heterogeneity in preferences. I make two hypotheses for expected utility maximisers. First, the policy choices should not imply WTP for any capped attribute that exceeds the increase in the cap. For example, in the experiment the annual cap for general dental varies between \$350 and \$700. Nobody should be willing to pay more than \$350 to increase the cap on general dental across this range. Second, assuming independence of utility from ancillary and hospital services, WTP for common attributes (the ancillaries) should be equal regardless of whether the policy covers ancillaries only or ancillaries and hospital services².

Both of these hypotheses are violated in the experiment. Continuing with the general dental example, up to 21% of respondents in the ancillaries treatment and 56% in the combined policy treatment are willing to pay more than \$350 to increase the cap by \$350. These

²In other words, this is conditional on there being no correlation between preference weights for ancillaries and hospital services. I argue that since these services are mutually exclusive they are unlikely to be substitutes and estimate a model that imposes this as a baseline.

results cannot be explained by non-compliance – similar estimates are obtained when the analysis is restricted to classes of respondents who are likely to be compliant. WTP for general dental, physical health services and remedial massage is also significantly higher in the combined policy treatment. The discrepancies across treatments seems to be driven by increased likelihood of deviating from a price to an attribute focus in the combined treatment.

When predicted choices are compared with the distribution that would minimise out-of-pocket costs, I find that respondents are likely to over-insure on the intensive margin, especially with combined policies. This result is consistent with the high profit margins for ancillaries insurance. There is no strong evidence that insurance status (i.e. product experience) improves choices in this regard. Education level also only delivers a modest improvement in the over-insurance rate.

Overall, the results suggest that many consumers do not choose health insurance according to the traditional expected utility model. On average, since deviations from this model result in too much insurance, the negative consequences are likely to be financial rather than to health or risk exposure. The results also indicate that product bundling is one form of complexity with serious implications for choice quality. When choosing combined hospital/ancillaries insurance, respondents are more likely to violate behavioural assumptions of expected utility theory as well as choose a policy more comprehensive than required to minimise out-of-pocket costs.

The paper is organised as follows. Section 2 provides a review of the literature on choice quality in health insurance markets. Section 3 provides background information on the Australian health insurance market. The data and experimental design are discussed in Section 4. Section 5 discusses the modelling strategy and main predictions. Section 6 presents results, which are discussed in Section 7.

2 Previous literature

Most of the literature on health insurance choice quality has focused on the US, particularly the Medicare Part D market (which provides prescription drug coverage for the predominantly elderly) or using administrative data for a single employer. These studies typically find that consumers make mistakes with high frequency. However, it is unclear how far these results extend to other insurance markets. The population of Medicare Part D recipients is older than the general population and pharmaceuticals comprise a narrow group of health expenditures. Similarly, the design of employer insurance schemes are often unique to that employer. The nature of health insurance is also different in the US than most other developed countries, where private health insurance is usually duplicate to universal public insurance and may also cover supplementary health costs. This paper focuses on choices in the Australian context, which is institutionally similar to the health insurance systems in many countries in Western Europe. To the best of my knowledge, this is the first paper to estimate preferences for policy features and evaluate choice quality within the Australian health insurance market.

The Medicare Part D market has attracted attention because of its complexity and large choice sets. Zhou and Zhang (2012) show that in 2009 only 5.2% of Medicare Part D participants chose the plan that minimised their out-of-pocket costs. Abaluck and Gruber (2011), Ketchum et al. (2012) and Heiss et al. (2013) also document significant overspending in this market.

While overspending can indicate choice inconsistency, it may also reflect high degrees of risk aversion. Abaluck and Gruber (2011) use a structural model and find that Medicare Part D participants overweight price relative to out-of-pocket costs, place irrational value on financial features of plans, and do not value reductions in the variance of expected out-of-pocket costs. Harris and Keane (1999) find that elderly Americans choosing supplementary Medicaid cover behave as if the basic plan has the most generous cost sharing arrangements, when it is actually the least generous option available. Violations of expected utility theory

that do not rely on parametric assumptions are demonstrated by markets where people choose to remain in financially dominated plans (i.e. plans unambiguously worse than at least one other option) offered by the same insurer, such as in Sinaiko and Hirth (2011), Handel (2013) and Bhargava et al. (2015).

Some studies attempt to explain departures from the traditional behaviour model. Several researchers identify the importance of inertia and switching costs (Handel, 2013; Handel & Kolstad, 2015; Polyakova, 2016; Ho et al., 2015), while Schmitz and Ziebarth (2016) identify the importance of price framing in switching decisions and (Frank & Lamiraud, 2009) demonstrate that with more competitors consumers are less likely to switch. Handel and Kolstad (2015) focus on information frictions (i.e. consumers failing to understand products) and hassle costs and show that failing to account for these factors biases upwards estimates for risk aversion. Kling et al. (2012) conduct a field experiment to demonstrate that providing people with publicly accessible information on their own expected health costs can overcome so called ‘comparison frictions’ and entice switching to a better value plan. Little is known about the effect of product bundling on choice quality³.

Identifying why people deviate from the traditional behaviour model has been difficult since behavioural factors may be confounded by risk preferences, brand effects or other complicated preference structures. Experimental data can overcome some limitations of revealed preference research in this regard since the researcher exercises control over the choice environment. To date, most experimental research analyses decisions within heavily controlled environments designed to mimic the kind of calculations consumers should make when choosing health insurance (e.g. Schram & Sonnemans, 2011; Besedes et al., 2012a, 2012b; Kairies-Schwarz et al., 2014). This literature finds that participants generally make worse choices when the choice set is larger and often employ heuristic strategies rather than calculating expected values.

³Frank and Lamiraud (2009) control for bundling of basic/supplementary insurance as a proxy for complexity in their study of switching decisions and find no significant effect. However, they do not address the endogeneity of this variable and switching is not necessarily synonymous with choice improvement.

Less work has gone into DCEs that use real world insurance menus and ask respondents to make choices based on their own expectations, to which this paper contributes. A recent exception is Bhargava et al. (2015), although their choice set is unusual since only the premium and excess vary and most policies are financially dominated. They find that reducing search frictions and improving health insurance literacy reduces the probability of choosing a financially dominated policy under strong interventions while subtler interventions are ineffectual⁴. They also find that people with poor self-assessed health often employ the decision strategy of maximising cover – a heuristic that performs poorly in their setting.

There are several ways in which DCEs can improve our understanding of choice quality in health insurance markets. In particular, DCEs incorporate some of the key advantages of both laboratory type experiments and revealed preference research. Since data on policy details are often not available to researchers, DCEs can overcome this constraint. DCEs can also overcome some of the identification issues that have affected revealed preference research, since the researcher manages all information in the choice environment. In laboratory type experimental research, participants are typically bestowed with a particular health risk and required to choose from a menu of insurance policies that bear only loose resemblance to real world health insurance⁵. In contrast, DCEs are usually designed to mimic the actual market and can in principle be used to estimate true preferences and predict real-market behaviour.

There are also important drawbacks to DCE research that need to be acknowledged. The main concern is how well hypothetical choices map into real world choices, especially when choices are not incentive compatible⁶. Kesternich et al. (2013) attempt to quantify

⁴Effective interventions include providing a detailed table for policy comparison and providing a tutorial and rule of thumb for choosing a policy. Ineffective interventions include expressing cost as monthly/yearly, reducing the number of variable attributes from 2 to 1 and providing a warning that poor choice could be costly.

⁵This is usually necessary so that choices can be evaluated against different behavioural theories, such as expected utility theory.

⁶Although respondents were compensated for their time, it was not possible to link compensation to choices since any sequence of choices could reflect true preferences. An attention filter was used to remove non-engaged respondents and the modelling strategy seeks to identify potential non-compliers. It is worth noting that many of the findings in the paper relate to comparison across treatments and engagement effects should cancel out (provided engagement is not significantly affected by the treatment itself).

this by comparing hypothetical choice data to actual choices in the Medicare Part D market. They find that the sign and significance of coefficients are consistent across data types, are also generally not statistically different, and conclude that the hypothetical choice data are effective at predicting demand for product features. Bhargava et al. (2015) also find that the distribution of hypothetical health insurance policy choices obtained from an online survey closely resembles that of the firm they base their experiment on. Another issue is the degree to which the choice environment needs to be modified to meet the feasibility constraints of the DCE. This may reduce realism and weaken external validity.

Overall, while DCEs do not dominate other approaches on all dimensions they do offer important advantages that can improve our understanding of health insurance markets.

3 Institutional background

The Australian health care system consists of universal public insurance (Medicare) and parallel private insurance. Medicare covers admissions to and procedures at public hospitals and most primary care costs.

The private health insurance market includes two types of cover – hospital cover and ancillaries cover. Private hospital cover is largely duplicate to Medicare but can also include expenses excluded from Medicare (e.g cosmetic procedures). It is generally used to cover procedures in private hospitals. The benefits of private treatment include choice of physician and potentially avoiding long waiting times for elective procedures in the public system. People without private cover can still access the private hospital system at their own expense.

Ancillaries cover is supplementary health insurance for health services excluded from Medicare. Insurers have flexibility over what services they cover provided the service can be used to prevent or treat a sickness or injury and is excluded from Medicare⁷. The main services covered under ancillaries policies are dental, optical, chiropractic, physiotherapy and

⁷A limited number of therapeutic, oral and maxillofacial, pathology and diagnostic imagery related services can be covered despite also being subsidised through Medicare.

natural therapies.

The private health insurance market is subject to community rating. This means that insurers are prohibited from price discriminating on any basis or denying insurance to any consumer⁸. All premium increases are subject to ministerial approval.

Some incentives have been implemented to encourage purchasing private cover, including penalties based on age and income for the uninsured and a premium rebate. In 2015-16, the rebate for most singles (families) with income below \$90,000 (\$180,000) is 27.82%. Later I consider the effect of the rebate on choice quality.

Demand for health insurance in Australia has been linked to higher income, education, age, being coupled and being female (Hopkins & Kidd, 1996; Barret & Conlon, 2003; Cheng & Vahid, 2011). Interestingly, a positive correlation between health and demand has been observed for hospital insurance leading to so called advantageous selection, which has been explained by income, risk preferences and optimism for the future (Doiron et al., 2008; Johar & Savage, 2012; Buchmueller et al., 2013).

The main way the paper advances the Australian literature is by evaluating choice quality. There are several reasons to believe that consumers may be making low quality choices. The strict price regulation for insurers has created incentives to compete on product design (Deloitte Access Economics, 2012). As a result, consumers are faced with increasingly complex insurance contracts and extensive policy menus. Consumers' current choice set includes between 48 to 2050 products, depending on the state, for policies that include ancillaries cover. These policies distinguish themselves on numerous dimensions including premium, service coverage, excess, co-payments, waiting periods, exclusions and restrictions, annual limits on benefits, network service providers, loyalty bonuses and more. Several reviews into the private health insurance market have raised concerns that consumers are inhibited by the complexity of these products (Deloitte Access Economics, 2012; PHIAC, 2013; ACCC, 2015).

⁸Waiting periods can be imposed to combat the risk that consumers purchase insurance as a response to an adverse health shock.

The paper focuses particularly on policy choice within the ancillaries market and how choices are affected when consumers choose combined ancillaries and hospital cover. Combined policies include numerous additional features and therefore should complicate the choice environment. Currently 56% of Australians (approximately 13.3 million individuals) have some type of ancillaries cover, of which around 85% are combined policies.

There are several features of ancillaries policies that make them an ideal focal point for evaluating choice quality. First, the benefits receivable for each service are typically subject to annual caps. This provides an upper bound for rational consumers' WTP for additional coverage of these services, which the experimental design exploits.

Second, ancillaries health insurance is supplementary (not duplicate) and therefore expected benefits from insurance are simply the expected expenditure on health services minus benefits paid by the insurer. In the case of hospital insurance, consumers always have the option of treatment in a public hospital and consequently the expected benefits are harder to characterise (they depend on perceived quality differences between the public/private sector).

Third, although ancillaries services are generally low cost, high frequency and subject to less uncertainty than hospitalisations, ancillaries policies are highly profitable. Across the industry, the gross profit margin for ancillary services is 24.3% whereas it is only 11.7% for hospital benefits (PHIAC, 2013). This provides a further incentive for understanding whether consumers make consistent choices in this market.

4 Data

4.1 Choice task

The choice task involves respondents from an online panel choosing between two hypothetical health insurance policies eight times. Respondents are randomly assigned to one of two treatments. In treatment one (T1), the health insurance products are all ancillaries only

policies. In treatment two (T2), the policies are all combined hospital/ancillaries policies.

The attribute levels and policy features were chosen based on an extensive search of policies classified as either basic or medium⁹ on the Private Health Insurance Ombudsman's (PHIO's) information website www.privatehealth.gov.au and were subjected to pilot testing¹⁰. The PHIO website presents summary details for every policy offered in Australia in so called Standard Information Statements (SIS). The information for each policy is strictly regulated by PHIO and intended to provide consumers with a simple platform for comparing policies across insurers. While the PHIO website is not the only way that consumers choose health insurance, it is growing in popularity. There were 1,054,858 unique visits in 2014-15, an increase of 17% on the previous year and 70% on the previous two years (PHIO, 2015).

The hypothetical insurance policies are presented in a similar way to the PHIO website. Specifically, the ordering of information and terminology of the SIS is closely adopted. Respondents could access information about more ambiguous policy features by hovering their mouse over the feature (see Appendix A for definitions). This information was drawn from the PHIO glossary. Because the PHIO website is specifically designed to assist consumers to make better informed choices, studying choices within this context provides a useful evaluation of its efficacy.

The ancillaries health services common to both treatments are general dental, optical, physical health services (physiotherapy, chiropractic, osteopathy and acupuncture), natural therapies and remedial massage¹¹. Dental, optical (predominately glasses) and physical health services are the largest claim and benefits categories in the market and feature on almost all policies. Natural therapies also features on many policies and are one of the fastest growing categories¹². Remedial massage was included because it has high utilisation rates.

⁹Top level policies were omitted because they are necessarily comprehensive and to have included all of the necessary features in the experimental design would have been infeasible.

¹⁰Details of the pilot study are available on request.

¹¹Because of feasibility constraints, only a subset of information contained in the SIS could be included in the hypothetical choice task. The hypothetical choice tasks are therefore simpler than the real world choice task and any deviations from rational behaviour could be understated. Details on the discrepancies between the hypothetical choices and the SIS are in Appendix A.

¹²Between 2010-11 and 2013-14 benefits for natural therapies increased by 88% compared to 33% for

It is also interesting to include a highly discretionary health service for comparison. For example, it is possible that consumers feel more confident choosing coverage for discretionary services (where there is less uncertainty) and therefore focus heavily on these when choosing insurance. The levels for each health service are determined by the annual benefits cap (the maximum the insurer will return in benefits). Policies also varied with regard to the insurers co-payment rate.

For hospital services, three levels of inclusions were chosen, which sequentially broaden appeal to people with different health risks. The most generous policies provide some cover for palliative care, psychiatric, rehabilitation, cataract and eye lens procedures, gastric banding and related services, sterilisation, cardiac and cardiac related services and hospital treatment for which Medicare pays no benefit (e.g. cosmetic surgery). The least generous policies only cover the first three. The excess (deductible) was also varied as well as a policy coverage indicator stating how many hospital services out of 10 the policy would provide some benefits towards. Some additional information was included (following the SIS) but did not vary across policies.

Table 1 shows the attributes and levels for T1 and T2. An example of the choice task in T2 is in Figure 1 (the T1 choice task is nested).

4.2 Experimental design

To generate the choice sets, a main effects D-efficient fractional factorial design was implemented using SAS (see Kuhfeld, 2010)¹³. Each treatment includes two blocks of eight choice scenarios. Respondents are instructed to choose a preferred policy for the next 12 months based on their personal circumstances only. They are specifically advised that the policy will not cover the health costs of their spouse or children (see Appendix A). The order of ancillaries overall (PHIAC, 2014).

¹³The covariance matrix of a multinomial logit model with zeros for all coefficients was used. The choice set in T2 was updated based on pilot estimates. This was because the initial choice set included two highly undesirable options, one of which was never chosen. There was no such issue with T1. Further details on the pilot study are available on request.

choice tasks is randomised.

Note there is no opt-out option and respondents are therefore forced to choose a preferred policy. While an opt-out would improve realism, it was omitted since the estimation is primarily concerned with evaluating choice quality rather than modelling demand and take-up. Further, the uninsured (around 50% of the sample) would be expected to almost always opt-out. Omitting this group from the analysis would discard an opportunity to compare the choices made by current (experienced) consumers and possible future consumers.

While the experimental design prevented purely dominated policies, it did not prevent financially dominated policies in the following sense. Say Policy A is worse or no better than Policy B on every attribute except that the cap for remedial massage is \$100 higher. If the price of Policy A is \geq \$101 higher than the price of Policy B, then no respondent should choose Policy A. In the choice sets 2/16 choice tasks included a financially dominated policy (with one financially dominated policy appearing in each block so that each respondent faces a choice task with a financially dominated option once).

4.3 Online panel

The survey company Qualtrics supplied a sample of online respondents for this study. In addition to the DCE questions, the survey collected information on respondents' basic demographics (e.g. age, sex, education, relationship status, employment), self-assessed health, health care utilisation, risk aversion, health insurance status, health insurance comprehension and household income. The full survey is available on request.

1,528 people from Qualtrics' online panels completed the survey and were evenly assigned to T1 and T2¹⁴. People aged 25-64 years were invited to complete the survey and quotas for age, sex and education were used to improve representativeness¹⁵. Responses were collected between 10-21 December 2015 (just under 70% of these responses were collected between

¹⁴The choice task was unintentionally altered for six respondents in T2, so they are omitted in the main analysis.

¹⁵The 2011 Census was used to create quotas for: people aged 25-34 years; 35-44 years; 45-54 years; 55-64 years; sex; and people with university degrees or higher.

15-16 December). The average response time was 16 minutes. Of those who started the survey, 44% completed. 263 people were ejected from the survey because they failed an attention filter¹⁶. The project was approved by the University of New South Wales Human Research Advisory Panel F (reference HC15702).

4.4 Descriptive statistics

Table 2 compares sample means for key variables across treatments and against population benchmarks. Overall, the treatments are well balanced and the distribution of characteristics closely matches the Australian population, even on non-quota variables like couple status, income and insurance coverage. Just under half the sample is male and 26.5% have a university degree. Sixty two per cent are employed, which is slightly lower than the population benchmark. Around 50% have some type of health insurance. Dental is the most commonly utilised health service (57% visited a dentist in the previous 12 months). Conversely, only 5% of respondents visited a natural therapist in the previous 12 months. The sample age distribution is slightly more right skewed than the population benchmark due to a sampling error that resulted in extra 55-64 year olds being sampled¹⁷.

To gauge health insurance comprehension, respondents are asked to answer 3 questions. The first 2 questions are multiple choice and ask respondents to identify an excess (deductible) and a co-payment (with “I don’t know” as the fifth option). These questions measure understanding of insurance specific features. The third question is a simple probability question, commonly used in measures of financial literacy and numeracy. The question asks “If the chance of getting a disease is 10 percent, how many people out of 1,000 would be expected to get the disease?”. All three questions are detailed in Appendix A.

The results on health insurance comprehension are in Table 3. Only 34.6% and 38.1% of

¹⁶The attention filter was a 5-point question matrix asking respondents how satisfied they were with their life, neighbourhood, financial situation and then the filter question that instructed them to choose ‘dissatisfied’.

¹⁷No weighting adjustment is made in the analysis since the oversampling was relatively modest and critically was not different across treatments. Age is controlled for in the main regressions.

respondents can correctly identify an excess or co-payment respectively and 17.5% answered all three questions correctly. Even among the insured and university educated, less than half answer the definition questions correctly. Given this low level of comprehension, there is reason to doubt the quality of consumers' choices. I am not aware of any comparable Australian estimates of health insurance comprehension and this quantification may be a new finding. Loewenstein et al. (2013) measure understanding of an excess and co-payment using a representative sample of insured Americans and find that 81% and 72% of their sample correctly identify these definitions in a multiple choice question. The apparently higher rate of comprehension in the US may reflect the relative importance of health insurance there compared to Australia (since Medicare provides a safety net to consumers in Australia). The probability question was correctly answered by 83.6% of respondents, which is similar to Lipkus et al. (2001) and Lusardi and Mitchell (2007) who use samples of highly educated adults and older American respectively.

5 Modelling

The modelling strategy involves estimating preferences using the random utility framework (McFadden, 1973), which assumes that respondents are expected utility decision makers subject to random error. This framework delivers predictions for the model parameter estimates. Choice consistency with expected utility theory can then be evaluated by testing whether the estimated parameters obey these predictions (which is similar to the approach in Abaluck and Gruber (2011)). The model is also used to predict policy choices, which are compared to the choices that would minimise out-of-pocket costs.

Utility individual i receives from policy choice j in scenario s is expressed as

$$U_{ijs} = x'_{ijs}\beta_i + \epsilon_{ijs} \quad i = 1, \dots, N; j = 1, \dots, J; s = 1, \dots, S \quad (1)$$

where utility depends on a vector of policy attributes x'_{ijs} with preference weights β_i that

may vary across individuals. ϵ_{ijs} is a random error term discussed below.

With rational agents, x'_{ijs} can be reduced to a vector of the premium, expected out-of-pocket (OOP) costs and the variance associated with policy j for individual i , since agents should only care about insurance to the extent that it reduces OOP costs and their variance (Abaluck & Gruber, 2011). However, this would require high quality data on expected utilisation and costs. While data on utilisation was collected (and is used later), available cost data are not precise enough to estimate this model. Instead, the model controls for policy features only under the constraint that the coefficients cannot distinguish between consistent preferences (utility reflecting the reduction in OOP costs and variance) or inconsistent preferences (utility exceeding the impact of policy features on OOP costs and variance).

I overcome this interpretive constraint by exploiting the caps for ancillaries health insurance. For these services caps on the maximum benefits receivable in a 12 month period apply. These caps provide boundaries on rational consumers' WTP. Respondents whose choices imply WTP for policy features exceeding these caps are assumed to follow a decision strategy inconsistent with expected utility theory. This method for identifying choice inconsistency has the advantage of being independent of risk preferences as well as expected health costs (which are often difficult to estimate).

Assuming homogeneous preferences and ϵ_{ijs} is identically and independently distributed (IID) as type 1 extreme value, equation (1) can be estimated by multinomial logit (MNL). However, this implies highly restrictive substitution patterns, specifically independence from irrelevant alternatives. Moreover, preference homogeneity is inconsistent with heterogeneous expected health service costs. To allow for preference heterogeneity and flexible substitution patterns, this paper estimates latent class logit (LCL) models for policy choice.

LCL models treat preference weights as discrete random variables and identify unique respondent classes. In this way, LCL has the advantage over continuous mixture models of not restricting the distribution of preference heterogeneity to a known distribution (typically multivariate normal). LCL can also uncover interesting patterns of preference heterogeneity

by incorporating observables into the class allocation model (Hess et al., 2011). A further advantage is that LCL can help to identify respondents who are non-compliant with the choice task without imposing restrictions on the form of non-compliance (Burton & Rigby, 2009). This is important since non-compliance could result in high WTP estimates and bias upwards estimates of choice inconsistency.

The LCL probabilities are expressed in equation (2). c indicates the class of preferences and y_{ijs} is an indicator for whether respondent i chose policy j in scenario s .

$$P_{ij} = \prod_{s=1}^S \prod_{j=1}^J \left(\frac{e^{x'_{ijs}\beta_c}}{\sum_{j=1}^J e^{x'_{ijs}\beta_c}} \right)^{y_{ijs}} \quad (2)$$

Class allocation is modelled as a fractional multinomial logit with π_{ci} used to represent the class allocation probability. These are incorporated into the log-likelihood function as follows.

$$LL = \sum_{i=1}^n \ln \sum_{c=1}^C \pi_{ci} P_{ij} \quad (3)$$

One drawback of LCL is that the correct number of latent classes is unknown and specifying additional classes can reduce efficiency and result in classes with very low predicted shares. This is often addressed by estimating the model with various numbers of latent classes and comparing a measure of model fit such as Bayesian Information Criterion (BIC), which is the approach taken in this paper.

Finally, one might worry that within each class preferences are themselves heterogeneous. A hybrid model is more likely to capture the behaviour of respondents with near lexicographic preferences (Keane & Wasi, 2013), such as heuristic decision makers who focus narrowly on a single attribute.

To accommodate this, random parameter latent class logit (RPLCL) models are also estimated and compared to standard LCL results. This model, described in Greene and Hensher (2013), is essentially the same as standard LCL except that the choice probabilities

in (2) are replaced by (4)

$$P_{ij} = \int \prod_{s=1}^S \prod_{j=1}^J \left(\frac{e^{x'_{ijs}\beta_{ic}}}{\sum_{j=1}^J e^{x'_{ijs}\beta_{ic}}} \right)^{y_{ijs}} \phi(\beta_{ic} | \beta_c, \Sigma_c) d\beta_{ic} \quad (4)$$

where $\beta_{ic} = \beta_c + \eta_{ic}$, ϕ is the normal density function and η_{ic} is a multivariate normal random variable with $E(\eta_{ic}) = 0$ and variance Σ_c .

Estimation of (2) is straightforward but the log likelihood can be difficult to maximise when evaluated directly. Train (2008) suggests the expectation maximisation (EM) algorithm as a way of overcoming the computational difficulties of standard gradient based maximum likelihood estimation. Estimation of (4) is more difficult because the probabilities are calculated by integrating over the unknown vector β_{ic} . EM is used to estimate LCL and a hybrid EM and maximum simulated likelihood (MSL) approach is adopted for RPLCL, with MSL estimates for the parameters within each class used to update the EM algorithm¹⁸.

5.1 Model predictions

The modelling strategy offers two tests for whether choices are consistent with expected utility theory.

Hypothesis 1: WTP for attribute K will be $\leq \Delta CAP_k$ for capped benefits

This is the main prediction. The table below shows the thresholds for rational consumers' WTP for each capped health service.

Attribute	Dental	Optical	Phys. 2	Phys. 3	Naturo	Massage	Excess
WTP threshold	\$350	\$150	\$150	\$300	\$100	\$100	\$250

Hypothesis 2: WTP will be equal for common attributes in T1 and T2

This hypothesis is more restrictive than hypothesis 1 because it requires that $cov(\beta_a, \beta_h) = 0 \quad \forall a, h$, where ancillaries and hospital attributes are denoted by a and h respectively. This

¹⁸The estimation for (2) was implemented in Stata using the llogit command (Pacífico & Yoo, 2013). The estimation in (4) used a modified version of llogit with the within class random parameter model estimated with the mixlogit command (Hole, 2007).

would be violated if, for example, hospital services included in T2 are substitutes for ancillaries in both treatments (in which case WTP might be lower in T2). Given the ancillaries health services are exclusively administered out-of-hospital the zero correlation assumption seems reasonable. MNL estimation imposes $cov(\beta_a, \beta_h) = 0$ and I calculate WTP using this model as a baseline.

6 Results

6.1 Dominated choices

As a precursor to the main analysis I will briefly discuss respondents choosing financially dominated policies (i.e. direct violations of hypothesis 1). 16.0% and 23.6% of respondents chose a dominated option in T1 and T2 respectively¹⁹. I show later that these choices are largely due to non-compliance (i.e. respondents not engaging with the choice tasks). However, I also find that heuristic choice strategies are likely to be important. In T1 the dominated policy only offers an improvement on remedial massage and natural therapies cover in block 1 and 2 respectively. For people who visited a massage therapist or natural therapist in the previous 12 months, the probability of choosing a financially dominated policy increases from 15.7% to 24.6% and from 13.0% to 45.0% in these blocks²⁰. The choices may therefore be partly due to a narrow focus heuristic triggered by expected utilisation. I provide more evidence of this later.

¹⁹These figures are not directly comparable. In T1, there was a financially dominated option in block 1 that involved a \$140 premium increase for a maximum gain of \$100 in the cap for remedial massage. In block 2, there was a similar trade-off but for a gain in the cap for natural therapies. In T2 (both blocks) the inferior policy only improved on the insurer's co-payment rate (from 60% to 70%), which was insufficient to offset the premium increase.

²⁰These figures (significant at the 10% and 1% level respectively) were estimated from a logistic regression controlling for age, income, sex, region, education, risk aversion, health and health insurance status. The full results are omitted for brevity.

6.2 Attribute preferences

The main LCL results are presented in Table 5. The BIC minimising number of latent classes was three in both treatments²¹. Two variables are used in the class allocation model to capture attention to the choice task; health insurance comprehension (indicator for answering all three insurance questions in Appendix A correctly) and (log) response time. People who rush through the survey necessarily have shorter response times and are also less likely to make the effort to correctly answer questions on insurance. These variables should help to identify non-compliers. Additional variables in the class allocation models are age, sex, education, health insurance status and self-assessed health, while a number of additional variables were dropped due to statistical insignificance²².

The three latent classes correspond closely across treatments, suggesting distinct types of consumers for health insurance. Class 1 respondents are relatively price sensitive and generally respond in sensible and expected ways to attributes. Class 2 respondents pay little or no attention to price but have generally expected signs for other attributes. Class 3 respondents have imprecise and often inconsistent or unexpected preferences and seem most likely to be non-compliant. The probability of being allocated to this class is strongly related to shorter response time and lower health insurance comprehension²³. Interestingly, the estimated class size is relatively similar between T1 (28.3%) and T2 (22.8%) and is in fact smaller in T2. This suggests that the effect of additional complexity in T2 on compliance may be modest (or nil).

To facilitate a discussion about choice quality and preferences, WTP estimates are reported in Figure 2²⁴. The first panel corresponds to a simple MNL model. MNL imposes

²¹Class allocation covariates were excluded when choosing the number of latent classes.

²²These included income, region, self-assessed risk aversion, couple status and various measures of prior utilisation.

²³For example, the average partial effect (APE) for health insurance comprehension is 28.3 percentage points (ppts) in T1 and 17.7 ppts in T2. The APE to go from 25th to 75th response time is 10.1 ppts and 14.6 ppts respectively.

²⁴WTP is measured as the ratio of the (prior) class weighted attribute coefficient and the class weighted price coefficient in the case of LCL models (multiplied by 12 to express as an annual figure).

zero correlation between the coefficients on attributes, which was a condition for hypothesis 2 (equal preference weights across treatments). Panel 2 is the LCL model with average covariate values imposed on respondents in the class allocation model. Panel 3 is the WTP assuming 75th percentile response time, correct answers to all health insurance questions in Appendix A and average values for other covariates. Respondents with these characteristics are expected to provide higher quality responses and are therefore labelled as ‘thoughtful respondents’ (conditional on these characteristics, the probability of being in class 3 is 2.46% and 3.84% in T1 and T2 respectively)²⁵. Panel 4 is the WTP for hospital features in T2 for average and thoughtful respondents.

Hypothesis 1: The first way I assess choice quality is to test for violations of expected utility theory. Hypothesis 1 states that WTP should not exceed the increase in the benefits cap for any ancillaries service (these thresholds are outlined in Subsection 5.1). Beginning with the MNL estimates in panel 1, for T1 all WTP estimates are well within these thresholds. However, in T2 hypothesis 1 is not satisfied at the mean for optical (WTP = \$193 > \$150) and is only marginally satisfied for general dental (WTP = \$340 compared to the threshold of \$350).

Next I allow for preference heterogeneity through the LCL model in panel 2. All WTP estimates in T1 continue to satisfy hypothesis 1. For T2, the high WTP for optical is not robust to this more flexible model. However, WTP for general dental and remedial massage now violate hypothesis 1 at the mean. WTP to increase the cap for general dental and massage is \$405 and \$128 respectively. WTP for high physical health services is also near its threshold (\$271 compared to \$300). Focusing on ‘thoughtful respondents’ does not reverse these findings. In fact, WTP in panel 3 is higher for general dental, high physical health services and remedial massage. This suggests that non-compliance cannot account for these choice inconsistencies.

²⁵An alternative approach would be to identify respondents most likely to be in class 3 and drop them from the analysis. However, Lancsar and Louviere (2006) caution against this since it may remove potentially valid preferences and introduce selectivity bias.

Finally, note that in panel 4, WTP to reduce the excess for hospital admission by \$250 is \$456, which also violates hypothesis 1. This is consistent with other evidence that consumers overvalue excess reductions (e.g Shapira & Venezia, 2008; Sydnor, 2010). Overall, the results are indicative of choice quality being lower in T2.

Hypothesis 2: Hypothesis 2 states that WTP should be equal for common policy features across treatments. A perusal of panels 1-3 in Figure 2 suggest this is not satisfied. While the confidence intervals are often large for T2, there are several examples where WTP is statistically significantly higher in this treatment²⁶. Focusing on the LCL model in panel 2, WTP for general dental, high physical health services and remedial massage is \$87, \$67 and \$11 respectively in T1. In T2 these figures are \$405, \$271 and \$128. While the WTP estimates for other attributes sometimes also differ by large magnitudes at the mean, the differences are generally not statistically significant.

An important condition for hypothesis 2 to hold is that there is no correlation between preference weights for ancillary and hospital features in T2. The fact that WTP for general dental and high physical health services is higher in T2 even under MNL estimation in panel 1 (which imposes zero correlation between preference weights) suggests that violations of hypothesis 2 for these health services are not driven by preference dependencies. WTP for massage is not higher in T2 under the MNL baseline, although it is not clear why utility from massage would be dependent on coverage for hospital services.

*Preferences*²⁷: Overall there is considerable heterogeneity in preferences with many coefficients experiencing large changes in relative magnitudes across classes. Every attribute is significant in at least one class. Although only 46% of respondents in T2 value price, the average price elasticities are similar across treatments, namely 0.65 in T1 and 0.75 in T2²⁸.

²⁶One reason that the confidence intervals are larger in T2 is that T2 involves more attributes, while the number of choice scenarios was fixed across treatments, and it is therefore necessarily less efficient. Increasing the number of scenarios in T2 would have added additional complexity to this treatment, which would have made it difficult to isolate the effect of product bundling.

²⁷To maintain focus on choice quality and for brevity, analysis of the impact of covariates on preferences is reserved for future work.

²⁸This is the class weighted sample average elasticity i.e. $\frac{1}{n} \sum_{i=1}^n \sum_{c=1}^C \frac{\pi_{ic} P_i(y|\beta_c)}{P_i(y|\beta)} [\beta_{prem} \cdot prem (1 - P_i(y|\beta_c))]$ where $P_i(y|\beta)$ is the marginal choice probability and $P_i(y|\beta_c)$ is the choice probability for class c .

This sits at the upper end of existing estimates for Australia of between 0.2-0.6 (Butler, 1999, 2001; Ellis & Savage, 2008; Cheng, 2014)²⁹.

In both treatments dental is the most valued ancillaries attribute. However, the cap on dental also varied by the largest amount so it is difficult to compare its WTP to other attributes. The fact that coefficients for dental are always positive and significant in the ‘complying’ classes speaks to its importance to consumers (this is unsurprising given utilisation patterns in Table 2). There is little evidence that the co-payment rate is important to respondents, which suggests they care more about caps and coverage. Natural therapies also appears relatively unimportant on average (again, this corresponds well with utilisation patterns). Interestingly, in T1 there seems to be very little difference in WTP between a \$150 cap in physical health services and a \$300 cap. Conversely, in T2 WTP for the \$150 cap is low and insignificant while WTP for the \$300 cap is high. This may indicate something about the choice process. For example, the additional complexity in T2 may have diverted respondents’ attention towards high values for particular attributes.

Looking at WTP for hospital features, inclusions is the most important attribute for average consumers. WTP to go from low to high inclusions is \$840. As discussed above, WTP for reducing the excess is also high. The least important hospital feature is whether 9/10 (versus 8/10) hospital services are covered (WTP = \$131). Respondents may have favoured information on inclusions and the excess because it is more concrete.

Finally, note that the posterior probability of being in class 3 is 89.7% and 59.6% for respondents who chose a dominated policy in T1 and T2 respectively. While the previous section argued such choices were partly due to heuristic decision strategies, non-compliance also seems to contribute to a large degree³⁰.

²⁹These estimates should be compared cautiously since the DCE only includes the intensive margin.

³⁰This can also explain the positive coefficients on natural therapies and massage in T1 and ancillaries co-payment in T2. These features are the only ‘gains’ respondents receive when choosing the dominated policies.

6.2.1 Allowing for random parameters within classes

As discussed above, RPLCL is an attractive extension in this application. A drawback is the large number of parameters to estimate and increased complexity in interpreting the results. I will focus on the main insights gained by RPLCL relative to standard LCL.

RPLCL estimates are reported in Table 6. The statistical significance of standard deviations for several parameters reveals more heterogeneity than captured previously. For example, the previous estimates restrict WTP for general dental in class 2 T1 to a fixed value of \$312, implying that nobody has WTP that violates hypothesis 1 (i.e. $WTP > \$350$). However, with RPLCL we find that preferences within this class are themselves quite heterogeneous and $WTP > \$350$ is possible for a non-trivial share of respondents. This is more consistent with results below where I show a number of respondents use a narrow focus heuristic and choose policies to maximise a single attribute.

To utilise the additional information on the distribution of WTP values, the RPLCL coefficients are used to estimate the proportions of respondents with WTP exceeding the rational threshold for each capped attribute. This is achieved by calculating the proportion of people with WTP exceeding the threshold within each class, ρ_c , and weighting by the class share i.e. $\sum_{c=1}^C \pi_c \rho_c$. Since class 3 respondents are potentially non-compliant, the analysis is done separately on the class 1 and 2 respondents only and including all classes.

The results in Table 7 indicate that in both treatments many respondents make choices that violate hypothesis 1. Focusing on the results with all classes, between 15% and 46% of respondents have WTP in excess of the change in cap in T1. Decision quality is lowest for physical health services. In T2 the proportions of ‘irrational’ respondents are all higher, often by large amounts. In particular, 63% of respondents have WTP that is irrational for a reduction in the excess for hospital services and 53% for an increase in the cap for general dental. Interestingly, between 15-21% of respondents have WTP that is irrational for general dental in T1 compared to 53-56% in T2. This large increase could reflect a narrower focus on specific services in the more complicated (bundled) treatment. This would also be consistent

with the high WTP for high physical health services in the previous subsection.

The conclusions are not sensitive to restricting attention to class 1 and 2 only³¹. There continues to be a large proportion of respondents willing to pay irrational amounts (up to 32% in T1 and 60% in T2) and the proportions are generally higher for each health service in T2. This suggests that non-compliance is not driving the results.

6.3 Heuristic choice strategies

Previous experimental work has shown that respondents are likely to engage in heuristic decision making strategies when choosing insurance. A simple choice strategy would be to maximise a single attribute, a type of lexicographic decision making process. It is the extreme of what I will refer to as a narrow focus heuristic. One way this behaviour could be rationalised is through expected utilisation, similarly to Bhargava et al. (2015) who link irrational WTP for a low deductible to poor health. Narrow focus would also be consistent with recent research on selection effects in the ancillaries health insurance market. Kettlewell (2015) finds a positive correlation between utilisation and insurance status for a range of health services but little correlation between the joint probability of utilising multiple health services and insurance status.

Table 7 shows that 32% of respondents in T1 and 46% in T2 choose policies to maximise a single attribute. While it is tempting to conclude that bundling increases the probability of heuristic decision making, these figures are not directly comparable because there is less variation in attributes in T2³². Nevertheless, the focus of respondents is clearly affected by bundling. The most frequent strategy in T1 is to minimise the premium (16%) whereas only 6% of respondents follow this strategy in T2. The most frequent strategy in T2 is to maximise the level of hospital inclusions (11%). Product bundling seems to increase the

³¹Although the percentage of respondents with irrational WTP for natural therapies and massage in T1 does decrease to almost zero. This is related to the fact that respondents who chose dominated policy options are predominately allocated to class 3 and those choices drive up WTP for these attributes in T1.

³²In a D-efficient choice set, some attributes are occasionally held constant across scenarios and the frequency of this is necessarily higher in T2 than T1 due to the additional hospital features.

likelihood of deviating from a price focus towards an attribute focus.

To test whether a narrow focus heuristic can be rationalised by expected utilisation, Table 8 reports linear probability model estimates for the probability of following an attribute maximisation strategy for each ancillary service. For brevity only the coefficients on prior utilisation are reported. For T1, each visit to a physical health service provider or massage therapist increases the probability of maximising these attributes by 1.9 and 2.3 percentage points respectively. These effects are large when interpreted against the low shares of people using these strategies (see Table 7). The fact we can link these health services to narrow focus could reflect their regularity and predictability relative to other services. Correlations for dental and optical are also positive but statistically insignificant. In T2 we see the same pattern except that prior dental utilisation is now statistically significant at the 10% level.

Overall, there is evidence that expected utilisation does influence some respondents to focus narrowly on particular health services. It is worth noting that while the number of respondents following a strict attribute maximisation strategy is low for most health services, other respondents may be following a less extreme version of this strategy whereby they still overweight particular services.

6.4 Policy choice simulation

Violations of expected utility theory do not necessarily imply poor real world choices. For example, heuristic choice strategies that result in irrationally high WTP in the DCE could perform well in practice, where competitive forces mean that prices do not permit the kind of irrational trade-offs the models imply consumers would be willing to make. For this reason it is important to consider how implied preferences might affect expected out-of-pocket (OOP) costs.

In the literature it has been common to use the OOP cost minimising distribution of policies as a benchmark for choice quality. I follow this by simulating the choice distribution for a two policy menu comprising ‘basic’ and ‘comprehensive’ ancillaries cover, with premi-

ums reflecting the real market. The basic policy features the minimum survey levels for each service (i.e. \$350 general dental, \$150 optical), while the comprehensive policy features maximum values. For T2 respondents, each policy also includes basic hospital cover. Premiums are set at 35% of the sum of ancillary health service caps (+\$80/month in the case of T2)³³. See Table 9 for further details.

To calculate the OOP cost minimising distribution I use utilisation data collected in the survey and state level price data from the Australian Prudential Regulation Authority for most variables. Price information for dental expenses is from Teusner et al. (2013) while expected expenditure on optical comes from self-assessed probability of replacing glasses next year multiplied by expected replacement cost. Further details on how expected health expenses are calculated are in Appendix B.

Predictions are generated using the LCL estimates in Table 4³⁴. Although previous results indicate that the random utility model may not be the best behavioural framework for estimating preferences, the flexibility of the LCL model suggests good forecasting ability. Model validation tests demonstrate this. In T1 (T2) the model correctly predicts 70% (64%) of choices conditional on covariates³⁵. This is notably better than the naive model with 50% accuracy (given there are always two choices). An alternative approach, that matches more closely the results to follow, is to estimate the model without the last choice scenario and predict the average probability of choosing Policy A in the omitted scenario. This results in mean predicted probabilities of 20% and 62% in T1 and T2 respectively, which is close to the observed frequencies of 24% and 60%.

The predictions in Table 9 are reported for the full sample as well as the uninsured, the university educated and their inverses. Differences in choice quality based on insurance status could reflect experience effects (although risk preferences may also differ across insurance

³³In developing the DCE I found that ancillaries policies sold by major insurers were generally priced at 30-40% of the sum of health service caps.

³⁴Predictions using RPLCL were similar.

³⁵Successful prediction is where the predicted probability is $> 50\%$ ($< 50\%$) and the respondent chooses (does not choose) the policy. The predictions based on the unconditional (prior) class membership probabilities perform even better, with 80% success and 77% success in T1 and T2 respectively.

status, so this is not necessarily identified). Education may proxy for cognitive ability and suggest whether cognition affects choice quality. The predictions are for the mean probability of choosing the comprehensive policy and the level of over-insurance is the percentage of people predicted to choose this policy minus the percentage that would choose this policy if they were cost minimising³⁶.

One issue around actual insurance status is that OOP cost estimates include possible moral hazard effects. If moral hazard exists, this would mean that for the insured OOP costs are potentially over-estimated for the ‘basic’ policy because they do not account for the expected reduction in health service usage³⁷. The opposite bias occurs in the case of the uninsured. I follow other papers that compare OOP costs (e.g. Abaluck & Gruber, 2011; Ketchum et al., 2012; Heiss et al., 2013), and do not make any correction for moral hazard, but acknowledge its potential influence. One reason for not correcting for moral hazard is that it is difficult to separately identify from adverse selection. The sample is almost evenly divided among the insured and uninsured, so it is possible that any bias is largely offset overall.

The cost minimising percentage of people choosing the comprehensive policy is 12% (9%) in T1 (T2). However, 42% (64%) are actually predicted to make this choice, which implies over-insurance rates of 31 percentage points (ppts) and 55 ppts in T1 and T2 respectively. This suggests that violations of expected utility theory uncovered in the previous subsections are associated with potentially large consequences for consumers in terms of OOP costs. The results also continue to highlight that product bundling reduces choice quality, in this case

³⁶This is an inherently conservative measure of choice quality. It does not take into account the fact that at the individual level respondents may be making sub-optimal choices even if the aggregate choice distributions match closely the cost minimising distribution. I choose to focus on the aggregate choices rather than individual choices because the cost data are based on average health service costs.

³⁷I am implicitly assuming that for those respondents with insurance, policies on average are more generous than the hypothetical basic policy (if for example a policy held was less generous than the hypothetical basic policy, the influence of moral hazard on observed utilisation would actually be less than what would be expected under the hypothetical policy). This is fairly non-controversial – the basic policy is equivalent to the least generous ancillaries policy I observed when reviewing products in the market. One caveat is that in Table 7 I treat as insured anyone with private health insurance, regardless of whether it covers ancillaries. Since most policies include ancillaries (see Table 2) this is not particularly important.

causing the degree of over-insurance to be 24 ppts higher on average.

Unsurprisingly, the cost minimising share is higher for the insured relative to the uninsured. In T1 the rate of over-insurance differs by less than 1 ppt across insurance status. Note however, that this rate is potentially biased downward for the insured and upward for the uninsured. In T2 there is a larger difference between the insured and uninsured (9 ppts). Given that the choice probabilities are almost identical for these groups (so that the difference is driven almost entirely by OOP costs), it is possible that this can be explained by moral hazard. Overall, there is no clear evidence of an experience effect from having insurance in terms of choice quality.

There is also only weak evidence that university education is associated with better choices. The rate of over-insurance is 9 ppts less for the university educated (compared to non-university educated) in T1. Since university status is positively correlated with insurance status this estimate may be biased upwards by moral hazard. This gap is also smaller (4 ppts) in T2. The rate of over-insurance is also much higher in T2 (52% compared 24% in T1). Altogether, the results suggest the returns to cognition are small at best and become smaller as the choice task becomes more complex.

Finally, Table 9 also presents predictions with the private health insurance rebate deducted from the insurance premium³⁸. Overall, the rebate does not change choice quality as measured by the rate of over-insurance in T1, while over-insurance modestly improves in T2. The rebate increases probabilities for the comprehensive policy and simultaneously reduces the cost minimising share by a similar magnitude. The small probability shifts reflect the low price elasticity for extra cover. One interpretation is that removing the rebate from ancillaries could result in a modest downgrading of policies with OOP costs remaining virtually unchanged on average.

³⁸As a simplification, the premium is assumed to be 27% for all respondents. In practice, this rate is slightly higher for the elderly and lower for those on high incomes.

7 Summary and discussion

The principle findings are that respondents frequently make low quality choices and product bundling exacerbates this. Up to 46% of respondents in T1 and 63% in T2 make choices that imply they are willing to pay more for increased coverage than they could possibly get back in benefits. Although health insurance comprehension is low in the sample, this cannot explain the results because even among better informed respondents preferences still violate expected utility theory³⁹. These preferences are partly driven by narrow focus choice strategies, which themselves can be linked to expected utilisation. Additionally, product bundling seems to cause respondents to deviate from a premium focus to an attribute focus. This is evident in the preference estimates, where only 46% of respondents in T2 have a negative and statistically significant premium coefficient, as well as the choice strategies with 16% of respondents minimising price in T1 compared to 6% in T2. When comparing predicted choices to cost minimising choices, respondents significantly over-insure on average, even among the highly educated. Again, the rate of over-insurance is much higher for those purchasing combined hospital/ancillaries cover.

These results are consistent with other recent studies that bring into question the traditional models of insurance choice. Models that assume consumers choose insurance based on expected value and risk preferences only are likely to mischaracterise these choices for a non-trivial share of consumers. Handel and Kolstad (2015) focus on this assumption and show that ignoring behavioural frictions can seriously bias welfare estimates when evaluating reforms. This paper shows that such bias is likely to be greater when the health insurance policies are more complicated (in particular, when they include additional health service categories). One stark example from this paper is that, in the context of combined hospital/ancillaries insurance, increasing the cap on coverage for general dental in the population from \$350 to \$700 is estimated to make the average person \$405 better off (as indicated

³⁹Also note that respondents were provided with all necessary definitions of attributes to make informed decisions.

by WTP estimates). This is clearly not true. When faced with ancillaries only cover this estimate is more reasonable (\$87). However, this figure itself may be unreliable for welfare analysis since 15-21% of respondents in the treatment are willing to pay > \$350. Researchers and policy makers should be cautious when interpreting preference estimates derived from models with standard behavioural assumptions. This echoes Baicker et al. (2015) who argue that estimating demand and moral hazard based on copays is unreliable in the presence of behavioural hazard (i.e psychological biases that shift demand beyond its effects on personal welfare).

One important finding is that breaking the choice task up into smaller tasks may improve choice quality. The fact that choice quality was frequently low in both treatments suggests there are limits to unbundling policies in terms of choice quality. Explicit unbundling may also have important supply side implications (e.g. Lavetti & Simon, 2016). Nevertheless, since 85% of policies in Australia are combined hospital/ancillaries, this could be an important focus for regulators. Another important finding is the propensity to over-insure. This raises questions about the generous subsidies currently provided for health insurance in Australia. A comprehensive welfare evaluation of the incentives for health insurance are beyond the scope of this paper, but the results do indicate that incorporating behavioural elements into such an evaluation would be important. The results also suggest that consumers are unlikely to utilise the information on the PHIO health insurance website correctly when choosing insurance. This suggests that simply presenting information in a straightforward way is not be enough to ensure sensible policy choices.

An interesting finding is that low quality choices tend to involve overly comprehensive insurance (of course, this does not preclude some choices leading to underinsurance). This pattern has also been observed with US data in the case of Medicare Part D (e.g. Abaluck & Gruber, 2011; Zhou & Zhang, 2012; Ketchum et al., 2012; Heiss et al., 2013) and with employer provided insurance (e.g. Handel & Kolstad, 2015; Bhargava et al., 2015) but there

appear to be few similar studies in countries with universal healthcare⁴⁰. One reason we might expect a difference between countries with universal health care and the US is that the stakes are likely to be higher in the US. In the case of ancillaries health insurance, even though these services are not covered by Medicare the stakes for consumers are still low because the benefits are capped at low values. It is also worth noting the very high estimated WTP for excess reductions, since this only applied to hospital services and is one of the major cost sharing instruments used by insurers.

These results indicate a common propensity to overinsure across different institutional settings. More evidence is needed on this, but it could help to identify the behavioural tendencies that distort consumer decisions and improve models of health insurance choice. For example, affect (feelings) experienced at the moment of decision making (Loewenstein et al., 2001) in the case of health insurance may systematically drive consumers towards additional coverage. At the same time, it is important to acknowledge the significant literature that finds consumers are often reluctant to take-up insurance, resulting in underinsurance on the extensive margin. Baicker et al. (2012) have proposed a myriad of different behavioural explanations for this such as choice overload, lack of understanding, misperception about risk, present-bias, transaction costs, reference dependence, framing and social comparisons. Understanding the different drivers of behaviour on the intensive and extensive margins and incorporating these into better models of consumer choice constitutes an important agenda for future research.

⁴⁰Buckley et al. (2012) find somewhat related experimental evidence in the case of WTP for parallel private health insurance in Canada.

References

- Abaluck, J., & Gruber, J. (2011). Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. *American Economic Review*, *101*, 1180–1210.
- ACCC. (2015). *Information and informed decision making: A report to the Australian Senate on anti-competitive and other practices by health insurers and providers in relation to private health insurance*. Australian Competition and Consumer Commission: Canberra.
- Arrow, K. J. (1963). Uncertainty and the Welfare Economics of Medical Care. *American Economic Review*, *53*, 941–973.
- Baicker, K., Congdon, W. J., & Mullainathan. (2012). Health Insurance Coverage and Take-Up: Lessons from Behavioral Economics. *The Milbank Quarterly*, *90*, 107–134.
- Baicker, K., Mullainathan, S., & Schwartzstein, J. (2015). Behavioral Hazard in Health Insurance. *Quarterly Journal of Economics*, *130*, 1623–1667.
- Barnes, A. J., Hanoch, Y., & Rice, T. (2014). Determinants of Coverage Decisions in Health Insurance Marketplaces: Consumers’ Decision-Making Abilities and the Amount of Information in Their Choice Environment. *Health Services Research*, *50*, 58–80.
- Barret, G. F., & Conlon, R. (2003). Adverse Selection and the Decline in Private Health Insurance Coverage in Australia. *The Economic Record*, *79*, 279–296.
- Besedes, T., Deck, C., Sarangi, S., & Shor, M. (2012a). Age Effects and Heuristics in Decision Making. *The Review of Economics and Statistics*, *94*, 580–595.
- Besedes, T., Deck, C., Sarangi, S., & Shor, M. (2012b). Decision-making strategies and performance among seniors. *Journal of Economic Behavior and Organization*, *81*, 524–533.
- Bhargava, S., Loewenstein, G., & Sydnor, J. (2015). *Do Individuals Make Sensible Health Insurance Decisions? Evidence from a Menu with Dominated Options*. NBER working paper 21160.

- Buchmueller, T. C., Fiebig, D. G., Jones, G., & Savage, E. (2013). Preference heterogeneity and selection in private health insurance: The case of Australia. *Journal of Health Economics*, *32*, 757–767.
- Buckley, N. J., Cuff, K., Hurley, J., McLeod, L., Mestelman, S., & Cameron, D. (2012). An experimental investigation of mixed systems of public and private health care finance. *Journal of Economic Behaviour & Organization*, *84*, 713–729.
- Burton, M., & Rigby, D. (2009). Hurdle and Latent Class Approaches to Serial Non-Participation in Choice Models. *Environmental and Resource Economics*, *42*, 211–226.
- Butler, J. R. G. (1999). *Estimating elasticities of demand for private health insurance in Australia*. NCEPH working paper No. 43.
- Butler, J. R. G. (2001). *Policy change and private health insurance: Did the cheapest policy do the trick?* NCEPH working paper No. 44.
- Cheng, T. C. (2014). Measuring the effect of reducing subsidies for private health insurance on public expenditure for health care. *Journal of Health Economics*, *33*, 159–179.
- Cheng, T. C., & Vahid, F. (2011). *Demand for Hospital Care and Private Health Insurance in a Mixed Public–Private System: Empirical Evidence Using a Simultaneous Equation Modeling Approach*. Melbourne Institute Working Paper Series, Working Paper No. 22/11.
- Chrisopoulos, S., & Harford, J. (2013). *Oral health and dental care in Australia: Key facts and figures 2012*. Australian Institute of Health and Welfare Cat. no. DEN 224.
- Deloitte Access Economics. (2012). *The Future of Private Health Insurance Premium-Setting: Seeking Integrative Solutions*. Research Report commissioned by Medibank Private Limited.
- Doiron, D., Jones, G., & Savage, E. (2008). Healthy, Wealthy and Insured? The Role of Self-assessed Health in the Demand for Private Health Insurance. *Health Economics*, *17*, 317–334.

- Ellis, R. P., & Savage, E. (2008). Run for Cover Now or Later? The Impact of Premiums, Threats and Deadlines on Private Health Insurance in Australia. *International Journal of Health Care Finance and Economics*, 18, 257–277.
- Ericson, K. M., & Starc, A. (2012). Hueristics and Heterogeneity in Health Insurance Exchanges: Evidence from the Massachusetts Connector. *American Economic Review: Papers & Proceedings*, 102, 493–497.
- Frank, R. G., & Lamiraud, K. (2009). Choice, price competition and complexity in markets for health insurance. *Journal of Economic Behavior and Organization*, 71, 550–562.
- Greene, W. H., & Hensher, D. A. (2013). Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. *Applied Economics*, 45, 1897–1902.
- Handel, B. R. (2013). Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. *American Economic Review*, 103, 2643–2682.
- Handel, B. R., & Kolstad, J. T. (2015). Health Insurance for “Humans”: Information Frictions, Plan Choice, and Consumer Welfare. *American Economic Review*, 105, 2449–2500.
- Hanoch, Y., Rice, T., Cummings, J., & Wood, S. (2009). How Much Choice Is Too Much? The Case of the Medicare Prescription Drug Benefit. *Health Services Research*, 44, 1157–1168.
- Harris, K. M., & Keane, M. P. (1999). A model of health plan choice: Inferring preferences and perceptions from a combination of revealed preference and attitudinal data. *Journal of Econometrics*, 89, 131–157.
- Heiss, F., Leive, A., McFadden, D., & Winter, J. (2013). Plan selection in Medicare Part D: Evidence from administrative data. *Journal of Health Economics*, 32, 1325–1344.
- Hess, S., Ben-Akiva, M., Gopinath, D., & Walker, J. (2011). *Advantages of latent class over continuous mixture of Logit models*. mimeo.
- Hibbard, J. H., Jewett, J. J., Engelmann, S., & Tusler, M. (1998). Can Medicare beneficiaries

- make informed choices? *Health Affairs*, 17, 181–193.
- Ho, K., Hogan, J., & Morton, F. S. (2015). *The Impact of Consumer Inattention on Insurer Pricing in the Medicare Part D Program*. NBER working paper 21028.
- Hole, A. R. (2007). Fitting mixed logit models using maximum simulated likelihood. *The Stata Journal*, 7, 388–401.
- Hopkins, S., & Kidd, M. P. (1996). The determinants of the demand for private health insurance under Medicare. *Applied Economics*, 28, 1623–1632.
- Johar, M., & Savage, E. (2012). Sources of advantageous selection: Evidence using actual health expenditure risk. *Economics Letters*, 116, 579–582.
- Kairies-Schwarz, N., Kokot, J., Vomhof, M., & Wessling, J. (2014). *How Do Consumers Choose Health Insurance? An Experiment on Heterogeneity in Attribute Tastes and Risk Preferences*. RUHR Economic Paper 537.
- Keane, M., & Wasi, N. (2013). Comparing Alternative Models of Heterogeneity in Consumer Choice Behavior. *Journal of Applied Econometrics*, 28, 1018–1045.
- Kesternich, I., Heiss, F., McFadden, D., & Winter, J. (2013). Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. *Journal of Health Economics*, 32, 1313–1324.
- Ketchum, J. D., Lucarelli, C., Miravete, E. J., & Roebuck, M. C. (2012). Sinking, Swimming, or Learning to Swim in Medicare Part D. *American Economic Review*, 102, 2639–2676.
- Kettlewell, N. (2015). *Utilisation and selection in an ancillaries health insurance market*. Paper presented at the 3rd Australian Health Economics Doctoral Workshop, Brisbane.
- Kling, J. R., Mullainathan, S., Shafir, E., Vermeulen, L. C., & Wrobel, M. V. (2012). Comparison Friction: Experimental Evidence from Medicare Drug Plans. *Quarterly Journal of Economics*, 127, 199–235.
- Kuhfeld, W. F. (2010). *Marketing Research Methods in SAS*. SAS Institute Inc.
- Lancsar, E., & Louviere, J. (2006). Deleting ‘irrational’ responses from discrete choice

- experiments: a case of investigating or imposing preferences? *Health Economics*, 15, 797–811.
- Lavetti, K., & Simon, K. (2016). *Strategic Formulary Design in Medicare Part D Plans*. NBER working paper 22338.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General Performance on a Numeracy Scale among Highly Educated Samples. *Medical Decision Making*, 21, 37–44.
- Loewenstein, G., Friedman, J. Y., McGill, B., Ahmad, S., Linck, S., Sinkula, S., . . . Volpp, K. G. (2013). Consumer’s misunderstanding of health insurance. *Journal of Health Economics*, 32, 850–862.
- Loewenstein, G., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as Feelings. *Psychological Bulletin*, 127, 267–286.
- Lusardi, A., & Mitchell, O. S. (2007). Baby Boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, 54, 205–224.
- McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). Academic Press: New York.
- Pacifico, D., & Yoo, H. (2013). lcglogit: A Stata module for estimating latent class conditional logit models via the Expectation-Maximisation algorithm. *The Stata Journal*, 13, 625–639.
- PHIAC. (2013). *Competition in the Australian Private Health Insurance Market*. Private Health Insurance Administration Council: Research Paper 1.
- PHIAC. (2014). *Operations of Private Health Insurers: Annual Report 2013-24*. Private Health Insurance Administration Council: Canberra.
- PHIO. (2015). *Private Health Insurance Ombudsman: Annual Report 2014-15*. Private Health Insurance Ombudsman: Canberra.
- Polyakova, M. (2016). Regulation of Insurance with Adverse Selection and Switching Costs: Evidence from Medicare Part D. *American Economic Journal: Applied Economics*, 8,

165–195.

- Rothschild, M., & Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics*, *90*, 629–649.
- Schmitz, H., & Ziebarth, N. R. (2016). Does price framing affect the consumer price sensitivity of health plan choice? *Journal of Human Resources*, *Forthcoming*.
- Schram, A., & Sonnemans, J. (2011). How individuals choose health insurance: An experimental analysis. *European Economic Review*, *55*, 799–819.
- Shapira, Z., & Venezia, I. (2008). On the preference for full-coverage policies: Why do people buy too much insurance? *Journal of Economic Psychology*, *29*, 747–761.
- Sinaiko, A. D., & Hirth, R. A. (2011). Consumers, health insurance and dominated choices. *Journal of Health Economics*, *30*, 450–457.
- Sydnor, J. (2010). (Over)insuring Modest Risks. *American Economic Journal: Applied Economics*, *2*, 177–199.
- Teusner, D. N., Brennan, D. S., & Gnanamanickam, E. S. (2013). Individual dental expenditure by Australian adults. *Australian Dental Journal*, *58*, 498–506.
- Train, K. E. (2008). EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, *1*, 40–69.
- Zhou, C., & Zhang, Y. (2012). The Vast Majority Of Medicare Part D Beneficiaries Still Don't Choose The Cheapest Plans That Meet Their Medication Needs. *Health Affairs*, *31*, 2259–2265.

Appendix A

This section discusses the structure and content of information presented to respondents in the DCE. The choice task was inspired by the Private Health Insurance Ombudsman (PHIO) website www.privatehealth.gov.au, where consumers can compare all private health insurance policies available in Australia. The details of policies are presented in Standard Information Statements (SIS), which are regulated by the *Private Health Insurance (Complying Product) Rules*. The main similarities and discrepancies between the DCE and SIS are described below.

Comparison of DCE to SIS

Notable similarities

(1) The order and terminology of rows. (2) Expressing premiums as monthly figures. (3) Expressing ancillaries benefits as annual caps. (4) The definitions of ambiguous health services (see below) were adapted from the PHIO website’s glossary page.

Discrepancies

(1) The SIS does not include a single insurer’s co-payment figure because most policies have more complicated co-payment structures. To accommodate this, the SIS has an extra column where insurers report the average expected benefits per typical claim for each health service. This structure was too complicated to incorporate into the DCE so a single co-payment figure was used instead. (2) There are 16 ancillaries health services (including ambulance) listed on the SIS for stand-alone ancillaries policies. To reduce the dimensionality of the choice task, the DCE only focused on a subset of services⁴¹. (3) No information on waiting periods. Waiting periods are to combat adverse selection from prospective members and are generally similar across policies.

⁴¹Note that some hospital features that were not varied in the DCE were still included in the policy description to avoid respondents making their own assumptions. For example, the row “What is covered if I have to go to hospital” has the same value for every policy.

Overall, the discrepancies between the SIS and the DCE resulted in the DCE being a simpler choice task. This is an important point given the research is concerned with choice quality and choice quality is expected to decrease with complexity.

Attribute levels

The levels for attributes were chosen based on searches of policies offered by major insurers in August 2015. Search was limited to basic and medium level policies. This was for two reasons. First, the sample includes uninsured respondents and higher level policies are unlikely to be attractive to this group. Second, higher level policies necessarily cover more health services. Incorporating additional health services would increase the complexity of the choice task and was not feasible given constraints on consumer engagement and survey length. The levels were tested in a small pilot and the level of trading-off observed indicated they were reasonable.

Note that physical health services (physiotherapy, chiropractic, osteopathy and acupuncture) were amalgamated into a single cap. The search of insurance policies found that insurers often combined these services or a subset of these services, which are often substitutes.

Preambles

Treatment 1

“The following questions are designed to understand what features of health insurance policies are important to you. You will be asked to choose between two different general treatment (‘extras’) private health insurance plans a total of eight times. These plans provide cover for you only (i.e. cannot be used to cover health services received by your child or partner). Please indicate your preferred plan, taking into account all features and your personal circumstances.

Important information before you start

For each health service, the amount displayed is the annual cap, which is the maximum amount the insurer will cover each year. For example, a \$350 cap for General Dental means that the most you can get back from the insurer on General Dental services is \$350 each year. For definitions and further information on policy features, you can hover your cursor over the feature you would like more information about.”

Treatment 2

“The following questions are designed to understand what features of health insurance policies are important to you. You will be asked to choose between two different combined hospital and general treatment (‘extras’) private health insurance plans a total of eight times. These plans provide cover for you only (i.e. cannot be used to cover health services received by your child or partner). Please indicate your preferred plan, taking into account all features and your personal circumstances.

Important information before you start

Each policy described below provides full exemption from the Medicare levy surcharge and Lifetime Health Cover loading.

For the ancillaries health services, the amount displayed is the annual cap, which is the maximum amount the insurer will cover each year. For example, a \$350 cap for General Dental means that the most you can get back from the insurer on General Dental services is \$350 each year. For definitions and further information on policy features, you can hover your cursor over the feature you would like more information about.

Note that you may find it easier to compare policies by reducing the text size. On most browsers you can reduce the text size by pressing ‘ctrl’ and ‘-’ (minus sign) together (or ‘command’ and ‘-’ together on Safari).”

Definitions available to respondents

Excess: Also known as a front-end deductible, an excess is an amount you must pay to the insurer if you are admitted to hospital. The excess is only payable for the first admission in a given year.

Co-payment (hospital): This is the amount you must contribute for each hospital admission or night in hospital.

Insurer's co-payment (ancillaries): This is the percentage of the service fee that you can claim back from your insurer. The remaining service fee must be paid by you.

General dental: Includes minor dental services, such as annual checkups, cleaning and fluoride treatment. Does not include endodontic services, orthodontic services or significant dental services, such as complex fillings, tooth extractions, crowns and bridges.

Optical: Includes prescription lenses, spectacle frames, and contact lenses.

Physiotherapy; chiropractic; osteopathy; acupuncture: Includes visits to a physiotherapist, osteopath, chiropractor or acupuncturist.

Naturopathy: Naturopathy uses a range of alternative approaches to medical treatments. Naturopathy can include nutrition, dietetics, herbal medicine and homoeopathy.

Health insurance comprehension questions

- 1) *Which of the following correctly defines a co-payment in the context of insurance?*
 - a) The amount that must be paid by the claimant before the insurer begins to cover any costs.
 - b) A contribution that the claimant pays for each service claimed (answer).
 - c) A refund paid back to the insurance holder when no claims are made during a specified period.
 - d) A claim threshold above which the insurer stops paying benefits to the claimant.
 - e) I don't know.

2) *Which of the following correctly defines a deductible (excess) in the context of insurance?*

a) The amount that must be paid by the claimant before the insurer begins to cover any costs (answer).

b) A contribution that the claimant pays for each service claimed.

c) A refund paid back to the insurance holder when no claims are made during a specified period.

d) A claim threshold above which the insurer stops paying benefits to the claimant.

e) I don't know.

3) *If the chance of getting a disease is 10 percent, how many people out of 1,000 would be expected to get the disease?*

a) I don't know.

b) If you think you know the answer, please write it below.

Appendix B

This section describes how ancillaries health care costs were estimated for the purpose of Table 9. I also provide some summary statistics on costs.

Data on utilisation for each health service was collected in the survey. Specifically, respondents were asked how many visits they had with each type of health care provider in the previous 12 months up to a maximum of five⁴².

Expected costs for general dental, physical health services, natural therapies and remedial massage are simply the number of visits multiplied by the average cost for the service.

General dental: Average costs are from Teusner et al. (2013), who estimate dental expenses for Australian adults. They find that most demographics are not statistically significantly correlated with average cost per visit. One exception is whether the person experiences toothaches (either sometimes, often or very often). I use this as a source of variation in average costs. Visits to the dentist for respondents who rarely or never experience toothaches are assumed to cost \$292.50 and visits for those who experience toothaches are assumed to cost \$328.

One issue with costs for dental is that they include both general dental (e.g. check-ups, cleaning etc.) and major dental (e.g. surgical tooth removal). Unfortunately Teusner et al. (2013) do not separate general and major dental costs, although fillings and scale and cleans (which are general) account for 82% of services for dentate adults aged 25-64 and 65% of visits are for check-ups (Chrisopoulos & Harford, 2013), so it is likely that the costs exclusive major dental would not be significantly lower. Furthermore, if we are overestimating costs then we are also underestimating over-insurance in Table 9, so this cannot reverse the conclusions in the paper.

Physical health services: Utilisation data was collected for each type of physical health service (i.e. physiotherapy, chiropractic, osteopathy and acupuncture). Average costs

⁴²While this top-coding does bias downwards estimates for total health costs, it is not a problem for analysing policy choice because five visits to any health care provider ensures costs are above any of the caps used in this study. Moreover, very few respondents had this many visits to any provider.

were calculated using September 2015 quarter Australian Prudential Regulation Authority (APRA) data on the value and number of claims made for each of these health services. The APRA data include the entire universe of claims made by Australians with private health insurance. Different prices were estimated for each state and territory. Average costs for each service ranged from \$46.97 for acupuncture visits in New South Wales to \$106.92 for osteopathy in the Australian Capital Territory.

Natural therapies and remedial massage: These costs were calculated identically to physical health service costs. The average cost to visit a natural therapist ranged from \$61.28 in New South Wales to \$79.94 in the Northern Territory. Remedial massage is included in the natural therapies category in the APRA data so was assumed to have the same costs.

Optical: Cover for optical primarily subsidises the cost of new glasses and lenses (consultations with an optometrist are generally covered by Medicare). While APRA data are available on optical claims and benefits, since the choice of glasses is highly discretionary I decided against using these data. Instead, respondents (conditional on requiring corrective eye-wear) were asked their expected probability (on a sliding scale from 0% to 100%) of purchasing new corrective eye-wear in the next 12 months and how much it would cost to replace their corrective eye-wear. Although respondents were told to ignore any rebates they might get from their private health insurer or otherwise, when reporting replacement costs a small number of respondents (5%) reported \$0. The average reported replacement cost was imposed on these respondents.

The average reported replacement cost is \$461, which is higher than expected. Inspection of the data found that this was being driven up by a small number of very high costs. The mean for respondents with replacement costs < \$1000 is \$269, which is more reasonable. Outliers are not a serious issue in terms of Table 9 because the maximum cap on optical of \$300 means there is effectively no difference between someone who reports a replacement cost of \$500 versus \$5000.

Finally, to estimate expected costs the self-reported probability of buying new glasses

was multiplied by the self-reported replacement cost. This takes a prospective view on costs. I also considered a retrospective view (more in line with the approach for other services, where the cost distribution relates to the previous 12 months) in which costs were calculated as self-reported replacement cost times a dummy variable equal to one if the respondent visited an optometrist in the previous 12 months. This effectively assumes that any visit to an optometrist for a respondent who wears glasses resulted in new glasses. The implied OOP cost minimising distribution using this approach for optical costs was not materially different from the one reported in Table 9.

Below are the means, medians and 75th and 25th percentiles for the costs of each ancillaries health service. For every health service there is a mass of respondents who do not have any costs. For only dental and optical are costs > 0 for the median respondent. However, more than 75% of respondents incur some cost.

	Dental	Optical	Physical	Naturo	Massage	Total
Mean	\$336	\$213	\$79	\$7	\$33	\$669
Median	\$328	\$41	\$0	\$0	\$0	\$461
75 th Perc.	\$656	\$198	\$69	\$0	\$0	\$875
25 th Perc.	\$0	\$0	\$0	\$0	\$0	\$102

Table 1: Attributes and levels

Attribute	T1 Levels	T2 Levels
Premium	\$14.17, \$20, \$25.83, \$31.67	\$100, \$112.5, \$125, \$137.5
<i>Ancillaries features</i>		
Insurer's co-payment rate	60%, 70%	60%, 70%
General dental	\$350, \$700	\$350, \$700
Optical	\$150, \$300	\$150, \$300
Physical health services	\$0, \$150, \$300	\$0, \$150, \$300
Natural therapies	\$0, \$100	\$0, \$100
Remedial massage	\$0, \$100	\$0, \$100
<i>Hospital features</i>		
Inclusions		Low, Medium, High
Excess		\$500, \$250
Services coverage		8/10 services receive benefits, 9/10 services receive benefits

Note: Low inclusions are palliative care, psychiatric and rehabilitation. Medium inclusions are low inclusions plus cataract and eye lens procedures, gastric banding and related services and sterilisation. High inclusions are medium inclusions plus cardiac and cardiac related services and hospital treatment for which Medicare pays no benefit (e.g. cosmetic surgery).

Figure 1: Example of treatment 2 choice task

	Policy A	Policy B
Monthly premium	\$137.50	\$100.00
Hospital features		
What is covered if I have to go to hospital?	<ul style="list-style-type: none"> • Doctor's bills in hospital • Ambulance fees 	<ul style="list-style-type: none"> • Doctor's bills in hospital • Ambulance fees
What services are not covered at all (exclusions)?	<ul style="list-style-type: none"> x Pregnancy and birth related services x Assisted reproductive services x Joint replacements ie shoulder, knee, hip and elbow, including revisions x Dialysis for chronic renal failure x Cataract and eye lens procedures x Gastric banding and related services x Sterilisation x Cardiac and cardiac related services x Hospital treatment for which Medicare pays no benefit eg most cosmetic surgery 	<ul style="list-style-type: none"> x Pregnancy and birth related services x Assisted reproductive services x Joint replacements ie shoulder, knee, hip and elbow, including revisions x Dialysis for chronic renal failure
What services are only covered to a limited extent (restrictions, benefit limitations periods)?	<ul style="list-style-type: none"> ✓ Palliative care ✓ Psychiatric services ✓ Rehabilitation 	<ul style="list-style-type: none"> ✓ Palliative care ✓ Psychiatric services ✓ Rehabilitation ✓ Cataract and eye lens procedures ✓ Gastric banding and related services ✓ Sterilisation ✓ Cardiac and cardiac related services ✓ Hospital treatment for which Medicare pays no benefit eg most cosmetic surgery
Doctors and hospital bills	9/10 medical services paid by this health insurance policy have no out-of-pocket expenses	8/10 medical services paid by this health insurance policy have no out-of-pocket expenses
Excess	\$500	\$500
Co-payment	None	None
Ancillaries features		
Insurer's co-payment rate	70%	60%
General dental	\$700	\$350
Optical	\$300	\$300
Physiotherapy; chiropractic; osteopathy; acupuncture	\$300	\$0
Naturopathy	\$100	\$0
Remedial massage	\$0	\$0

Note: The T1 choice task looks the same as the T2 with the rows for hospital features removed.

Table 2: Descriptive statistics (mean values)

Variable	T1	T2	All	Population
Age 25-34	0.240	0.216	0.228	0.258
Age 35-44	0.217	0.251	0.234	0.267
Age 45-54	0.226	0.228	0.227	0.257
Age 55-64	0.317	0.305	0.311	0.218
Male	0.454	0.530	0.492	0.493
University	0.284	0.246	0.265	0.265
Couple	0.594	0.644	0.619	0.635
Employed	0.628	0.605	0.616	0.713
PHI(hospital)	0.490	0.493	0.491	0.51
PHI (ancillaries)	0.505	0.497	0.501	0.54
<i>HH income</i>				
<\$60K	0.445	0.427	0.435	0.430
\$60K-<\$125K	0.386	0.397	0.392	0.336
\$125K+	0.169	0.176	0.173	0.235
<i>Visited health service provider in previous 12 months</i>				
Dental	0.564	0.582	0.573	0.549
Optical	0.474	0.480	0.477	-
Physical	0.272	0.267	0.270	0.245
Naturopathy	0.046	0.050	0.048	0.050
Massage	0.191	0.161	0.176	-
Hospital	0.317	0.318	0.317	-

Note: Sample size is 1,528 (764 in each treatment). Population values for age, sex, university, couple and employed are from the 2011 Census. PHIA/APRA figures (December 2011) and the 2011 ABS Census have been used for PHI coverage. Population figures for income are weighted values from the 2011 wave of the Household Income and Labour Dynamics in Australia Survey (HILDA). Figures for previous health service use are from the 2013 wave of HILDA. Note that for Physical, acupuncture visits are not included in the population value. Figures for optometry and hospital are omitted because these are defined differently in HILDA than the survey. Sample quotas were used for age, education and sex.

Table 3: Health insurance comprehension

Variable	All	Has PHI	University
Excess correct	0.346	0.420	0.447
Co-payment correct	0.381	0.419	0.435
Probability correct	0.836	0.854	0.842
All correct	0.175	0.219	0.244

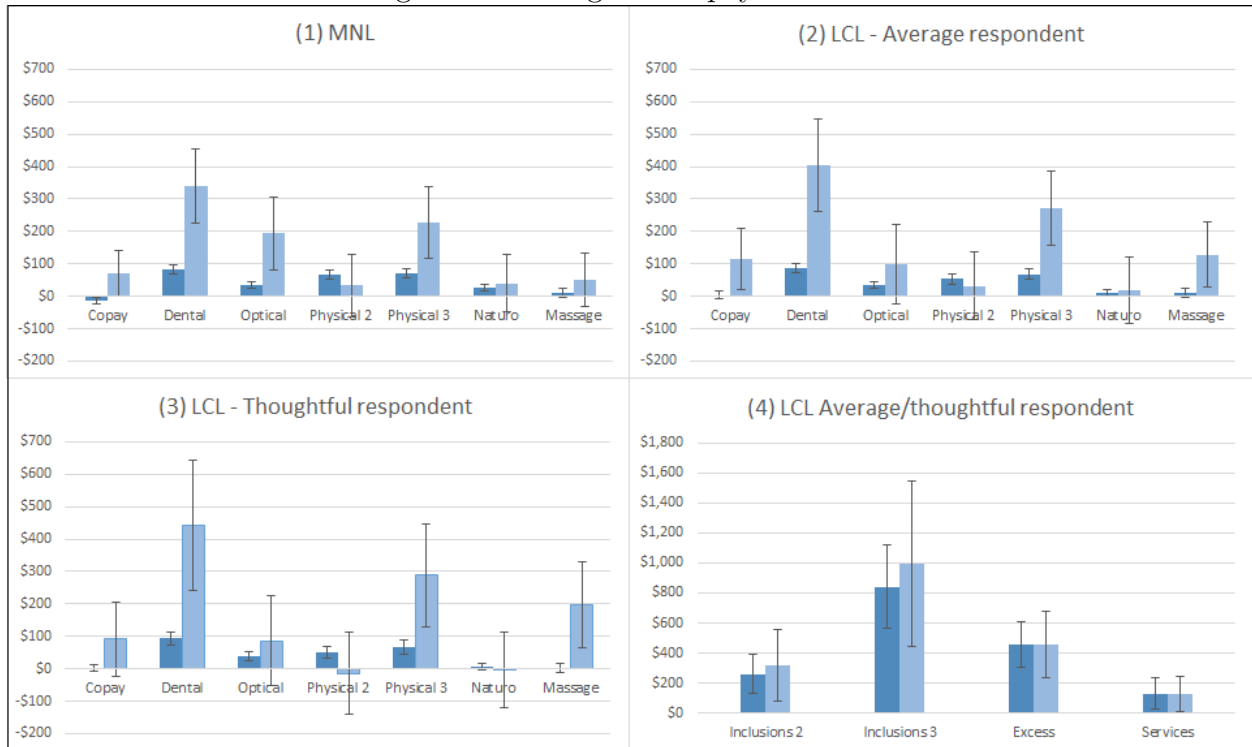
Note: Sample size is 1,528. The questions on on health insurance literacy were multiple choice with 4 options (see Appendix A).

Table 4: Latent class logit results

Variable	T1			T2		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
Premium	-0.251 (0.021)	-0.056 (0.009)	-0.033 (0.006)	-0.031 (0.004)	0.004 (0.005)	0.007 (0.004)
<i>Ancillaries features</i>						
Co-payment	-0.011 (0.138)	0.058 (0.079)	0.091 (0.063)	0.024 (0.080)	0.222 (0.162)	0.294 (0.110)
Dental	0.959 (0.152)	1.456 (0.138)	-0.244 (0.072)	0.540 (0.083)	0.671 (0.191)	-0.061 (0.123)
Optical	0.286 (0.127)	0.756 (0.087)	-0.202 (0.067)	0.287 (0.096)	-0.099 (0.184)	0.013 (0.131)
Physical 2	0.460 (0.181)	0.825 (0.103)	0.177 (0.088)	-0.262 (0.128)	0.234 (0.179)	0.542 (0.152)
Physical 3	0.538 (0.199)	1.158 (0.157)	0.276 (0.093)	0.276 (0.110)	0.525 (0.169)	0.071 (0.134)
Naturopathy	0.336 (0.115)	-0.158 (0.080)	0.206 (0.062)	-0.110 (0.088)	-0.152 (0.153)	0.085 (0.117)
Massage	0.547 (0.185)	-0.499 (0.105)	0.378 (0.080)	0.244 (0.080)	0.341 (0.155)	-0.468 (0.126)
<i>Hospital features</i>						
Inclusions 2	-	-	-	-0.210 (0.095)	1.218 (0.178)	0.124 (0.122)
Inclusions 3	-	-	-	0.100 (0.096)	2.877 (0.268)	0.105 (0.114)
Excess	-	-	-	0.436 (0.083)	0.817 (0.231)	0.317 (0.123)
Services	-	-	-	0.226 (0.072)	0.222 (0.164)	0.048 (0.093)
<i>Class allocation</i>						
Age	-	-0.006 (0.009)	-0.023 (0.008)	-	0.013 (0.008)	0.037 (0.011)
Male	-	-0.411 (0.216)	-0.259 (0.226)	-	-0.405 (0.215)	0.342 (0.292)
University	-	-0.637 (0.257)	-0.648 (0.267)	-	-0.475 (0.275)	0.227 (0.346)
Uninsured	-	-1.050 (0.230)	-0.740 (0.234)	-	-0.001 (0.223)	-0.137 (0.302)
SAH Excell	-	-0.251 (0.611)	0.034 (0.487)	-	0.099 (0.514)	1.498 (0.627)
SAH Vgood	-	0.272 (0.425)	-0.739 (0.387)	-	0.362 (0.374)	0.888 (0.526)
SAH Good	-	0.467 (0.415)	-0.429 (0.372)	-	0.417 (0.367)	1.246 (0.503)
SAH Fair	-	0.442 (0.444)	-0.042 (0.409)	-	0.594 (0.403)	0.720 (0.578)
Log rtime	-	0.198 (0.149)	-0.824 (0.171)	-	0.516 (0.139)	-1.196 (0.328)
PHI literate	-	-0.037 (0.265)	-2.638 (1.046)	-	0.098 (0.273)	-1.482 (0.575)
Constant	-	0.287 (0.662)	4.084 (0.587)	-	-2.394 (0.579)	-0.688 (0.721)
Class share	0.362	0.355	0.283	0.456	0.316	0.228
LL	-3290.33			-3663.02		
BIC	6886.03			7710.63		
n	12224			12128		

Note: Clustered standard errors in parentheses. Log rtime is the natural logarithm of response time and PHI literate is an indicator for answering the three PHI comprehension questions correctly (See Appendix A). Coefficients in bold are statistically significant at the 5% level.

Figure 2: Willingness to pay estimates



Note: The left columns in panels 1-3 are for T1 and the right columns are for T2. WTP is calculated as $12 \cdot \sum_{c=1}^C \hat{\pi}_c \hat{\beta}_{k,c} / \sum_{c=1}^C \hat{\pi}_c \hat{\beta}_{prem,c}$ using estimates from Table 4, where $\hat{\pi}_c$ is the estimated class share according to covariates in the class allocation model. In panel 2, the columns are average WTP for a respondent with average values for the covariates in the class allocation model. In panel 3, average values are used for all covariates except that PHI literate = 1 and Log rtime = the value for the 75th percentile respondent. In panel 4 are WTP estimates for average (left column) and thoughtful (right column) respondents. Error bars are 95% confidence intervals calculated using the delta method.

Table 5: Random parameters latent class logit results

Variable	T1			T2		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
Premium	-0.281 (0.021)	-0.073 (0.014)	-0.003 (0.006)	-0.042 (0.004)	0.006 (0.010)	0.010 (0.004)
<i>Ancillaries features</i>						
Co-payment	-0.015 (0.126)	0.093 (0.107)	0.096 (0.064)	0.068 (0.093)	0.214 (0.375)	0.345 (0.109)
Dental	1.038 (0.121)	1.851 (0.223)	-0.274 (0.069)	0.685 (0.104)	0.923 (0.359)	-0.142 (0.119)
Optical	0.308 (0.118)	0.978 (0.161)	-0.218 (0.066)	0.349 (0.115)	-0.073 (0.421)	-0.026 (0.122)
Physical 2	0.490 (0.157)	1.055 (0.160)	0.167 (0.088)	-0.211 (0.093)	0.233 (0.370)	0.647 (0.128)
Physical 3	0.572 (0.156)	1.483 (0.228)	0.101 (0.087)	0.386 (0.096)	0.635 (0.262)	0.056 (0.110)
Naturopathy	0.361 (0.101)	-0.187 (0.108)	0.231 (0.061)	0.132 (0.096)	-0.092 (0.332)	0.113 (0.109)
Massage	0.624 (0.123)	-0.595 (0.138)	0.419 (0.078)	0.285 (0.082)	0.356 (0.285)	-0.544 (0.103)
<i>Hospital features</i>						
Inclusions 2	-	-	-	-0.197 (0.102)	1.482 (0.289)	0.172 (0.117)
Inclusions 3	-	-	-	0.163 (0.163)	3.631 (0.545)	0.121 (0.113)
Excess	-	-	-	0.563 (0.115)	1.146 (0.373)	0.320 (0.128)
Services	-	-	-	0.308 (0.081)	-0.026 (0.284)	0.049 (0.099)
<i>Standard deviations</i>						
Premium	0.036 (0.037)	0.033 (0.026)	0.011 (0.029)	0.032 (0.005)	0.018 (0.028)	0.002 (0.025)
Co-payment	0.148 (0.825)	0.401 (0.275)	0.357 (0.180)	0.033 (0.328)	0.010 (1.001)	0.039 (0.486)
Dental	0.449 (0.290)	0.852 (0.281)	0.020 (0.593)	0.597 (0.143)	0.123 (0.877)	0.551 (0.148)
Optical	0.031 (1.050)	0.653 (0.233)	0.062 (0.510)	0.319 (0.262)	0.034 (1.030)	0.356 (0.308)
Physical 2	0.295 (0.417)	0.006 (0.754)	0.005 (0.475)	0.059 (0.314)	0.014 (0.716)	0.383 (0.262)
Physical 3	0.035 (0.920)	0.732 (0.269)	0.583 (0.121)	0.576 (0.159)	0.730 (0.437)	0.297 (0.248)
Naturopathy	0.523 (0.259)	0.022 (0.591)	0.106 (0.292)	0.399 (0.165)	0.041 (0.692)	0.532 (0.182)
Massage	0.085 (0.641)	0.445 (0.383)	0.524 (0.133)	0.429 (0.154)	0.277 (0.503)	0.293 (0.171)
Inclusions 2	-	-	-	0.293 (0.271)	0.509 (0.406)	0.061 (0.740)
Inclusions 3	-	-	-	0.007 (0.229)	0.618 (0.693)	0.414 (0.869)
Excess	-	-	-	0.379 (0.185)	0.669 (0.896)	0.705 (0.274)
Services	-	-	-	0.107 (0.261)	0.015 (1.143)	0.133 (0.565)
Class share	0.357	0.361	0.282	0.487	0.305	0.208
LL	-3292.95			-3665.59		
BIC	7050.59			7954.46		
n	12224			12128		

Note: Clustered outer product gradient standard errors in parentheses. 500 Halton draws were used when updating the MSL estimates. Coefficients in bold are statistically significant at the 5% level.

Table 6: Proportions of respondents with WTP > benefits cap

Attribute	Classes 1-2 only		All classes	
	T1	T2	T1	T2
Dental	20.75%	56.05%	14.89%	52.70%
Optical	26.65%	22.19%	19.14%	27.22%
Physical 2	31.86%	35.39%	46.16%	47.95%
Physical 3	19.04%	43.22%	26.84%	46.40%
Natural therapies	0.03%	19.53%	26.11%	27.69%
Massage	0.42%	60.22%	21.77%	48.38%
Excess	-	58.46%	-	62.99%

Note: Figures are the percentage of respondents whose preferences imply WTP that exceeds the increase in the relevant health service annual cap (or the reduction in the excess in the case of excess). These figures are simulated by multiplying the simulated proportion of respondents with irrational WTP in each class and taking the class-share weighted average. WTP is simulations use 10000 pseudo random draws.

Table 7: Attribute maximisation choice strategies

Attribute maximised	T1	T2
Price (lowest)	16.63%	6.20%
Ancillaries co-payment	0.91%	1.58%
Dental	3.93%	3.17%
Optical	2.49%	6.86%
Physical	2.36%	3.30%
Natural therapies	3.01%	6.73%
Massage	2.89%	0.66%
Inclusions		11.48%
Excess		5.01%
Services		2.64%
Any	32.20%	46.17%

Note: Figures are the percentage of respondents who consistently chose the policy that maximised a given attribute (minimised in the case of premium and excess). Not all attributes varied in all scenarios and for this reason results in columns 2 and 3 are not directly comparable. In particular, there is less variation in T2 which would tend to drive up figures. For massage in T2 only respondents in the second block are included because this attribute only varied in two scenarios in block 1, giving an unrealistically high figure for the percentage of respondents maximising it.

Table 8: Attribute maximisation choice strategies

Variable	Dental	Optical	Physical	Naturo	Massage
Prior utilisation (T1)	0.009 (0.008)	0.017 (0.012)	0.019 (0.006)	-0.008 (0.007)	0.023 (0.009)
Prior utilisation (T2)	<i>0.012</i> (0.007)	0.014 (0.023)	0.011 (0.004)	-0.013 (0.016)	<i>0.013</i> (0.007)

Note: Coefficient estimates are for a linear probability model with a dummy equal to one if the respondent maximised the relevant attribute as the dependent variable. Sample size is 764 in T1 and 758 in T2. Prior utilisation is the number of visits to the relevant health service provider in the previous 12 months except for Optical where it is an indicator for any visits. Other controls are prior utilisation for the other ancillaries, age, sex, university, couple status, region, income, self-assessed risk aversion, health insurance status and health insurance comprehension. Robust standard errors in parentheses. Coefficients in bold are statistically significant at the 5% level and coefficients in italics are significant at the 10% level.

Table 9: Policy choice probabilities (best expected value)

	T1		T2	
	Pr(Comp. policy)	Over-insurance	Pr(Comp. policy)	Over-insurance
All	42.37% (11.65%)	30.72	64.36% (9.04%)	55.32
Uninsured	35.82% (4.57%)	31.25	64.56% (5.56%)	60.00
Insured	47.30% (16.97%)	30.33	64.19% (12.98%)	51.21
University	39.05% (14.75%)	24.30	62.66% (10.27%)	51.94
No University	43.69% (10.42%)	33.27	64.91% (9.42%)	55.49
<i>With health insurance rebate applied</i>				
All	47.25% (16.75%)	30.5	66.51% (15.96%)	50.55
Uninsured	40.63% (6.71%)	33.92	66.75% (9.06%)	57.69
Insured	52.24% (24.31%)	27.93	66.31% (21.63%)	44.68
University	44.08% (22.12%)	21.96	65.07% (17.84%)	47.23
No University	48.51% (14.63%)	33.88	66.98% (15.36%)	51.62

Note: The probabilities in columns 2 and 4 are the sample average probability of choosing the comprehensive health insurance policy according to the LCL estimates. The comprehensive policy includes the following features; \$42.86/month premium, \$700 general dental, \$300 optical, \$300 physical, \$100 natural therapies, \$100 remedial massage and 60% insurer's co-payment. The basic policy costs \$14.28/month and only includes \$350 dental, \$150 optical and 60% insurer's co-payment. In T2 both policies also include low inclusions, \$500 excess and 8/10 service coverage. With the rebate applied, the premiums are reduced by 27%. The figures in parentheses are the percentage of respondents who should choose the comprehensive policy if they are minimising OOP costs. Over-insurance is the percentage points difference between the predicted share choosing the comprehensive policy and the cost minimising share.