UNSW Business School Research Paper No. 2016 ECON 17

Utilisation and Selection in an Ancillaries Health Insurance Market

Nathan Kettlewell

Utilisation and selection in an ancillaries health insurance market

Nathan Kettlewell*
University of New South Wales

October 2016

Abstract

I study two important aspects of the Australian private ancillaries health insurance (PAHI) market. First, I estimate the effect of PAHI on utilisation of various health services using instrumental variable methods to identify causal effects. Next I test for the presence and direction of selection effects by identifying variables not used in pricing that influence both the insurance and utilisation decision. PAHI covers a wide range of out-of-hospital health services, including many discretionary and preventative services. The most quantitatively important are dental, optometry, physiotherapy and chiropractic. I find that PAHI does increase utilisation of health services, particularly the probability of visiting a dentist, physiotherapist, chiropractor, osteopath or acupuncturist. I find evidence of selection effects in the sense that a number of different variables can predict a person's propensity to insure as well as their propensity to utilise health services. The variables that I identify generally result in adverse selection to insurers for higher frequency health services, although selection bias is more heterogeneous for lower frequency services. There is little evidence of self-selection based on the joint probability of different health services, which indicates that diversified policy menus are a possible strategy for addressing adverse selection in the PAHI market.

Keywords: health insurance, moral hazard, adverse selection, favourable selection

JEL: I13; D82

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

1 Introduction

Understanding who purchases health insurance, and how the provision of insurance affects decisions to utilise health services, is important to inform future health policy. There now exists a large volume of research on these topics across different institutional environments and for different types of health services and health behaviours. However, no paper to date has examined these issues in detail in the case of Australian private ancillaries health insurance (PAHI). PAHI, otherwise known as general treatment or 'extras' health insurance, is a supplementary health insurance product that provides benefits for a range of out-of-hospital services not covered by Medicare, Australia's universal health insurance system. These include some preventative and discretionary treatments like massage, acupuncture and naturopathy. The most quantitatively important service is dental, followed by optometry, physiotherapy and chiropractic. Just under 56% of Australians have PAHI. Similar supplementary insurance products are available in other developed countries such as Switzerland, Germany, the Netherlands and Canada (Colombo & Tapay, 2004).

The first contribution of this paper is to estimate the causal effect PAHI has on the utilisation of the following out-of-hospital services: dental, physiotherapy, chiropractic, osteopathy, acupuncture, optometry, naturopathy, psychology and counselling. While some recent work has examined the role of PAHI on utilisation of dental services (Hopkins et al., 2013; Srivastava et al., 2015), this is the first paper that extends attention to other expense categories. A second contribution is to examine the characteristics of those who purchase PAHI and in particular, the presence and direction of selection effects. There has been no focused study on selection effects in this market to date.

Health events in the PAHI market differ from those in most insurance markets, which usually occur with great uncertainty. For example, 65% of dental visits in Australia are for preventative check-ups (Chrisopoulos & Harford, 2013). Physiotherapy, chiropractic, osteopathy, acupuncture, naturopathy, psychology and counselling are also more discretionary than other events typically associated with insurance since treatment is often non-essential.

The predictability of expenses and discretion afforded to consumers about whether to actually utilise services suggest that the traditional risk sharing model may be unable to rationalise the PAHI market. For example, without a risk reducing role, self-selection is expected to be more strongly related to expected usage (adverse selection) than other insurance markets, which could create the conditions for complete market unravelling (Akerlof, 1970). The discretionary nature of expenses should also create pressure for this market through moral hazard (consumers either taking actions that increase their probability of making a claim or reacting to the price incentives provided by coverage). Despite this, the PAHI market is actually more profitable than many other insurance markets, most notably the private hospital market (gross profit margins are 11.7% for hospital cover compared to 24.3% for PAHI (PHIAC, 2013)).

The fact that PAHI policies cover a wide variety of health services could provide an insight into their continued popularity and profitability. Ellis and McGuire (2007) show that the service provision incentives for insurers are a multiplicative function of both the predictability of a service (how certain consumers are about their expected utilisation) and the predictiveness (how well utilisation of one service predicts utilisation of other services). While PAHI health services are likely to be predictable, it is less clear that they are predictive. In fact, I show later that the utilisation correlations are low for these services. This suggests a multidimensional approach to studying selection effects. More specifically, I study selection effects across different health services and consider whether groups of consumers are likely to be adversely selected based on one health service dimension but not another. For example, a consumer may choose to purchase PAHI based on their expected utilisation of dental services but the same consumer also has low expected utilisation of other health services. If this is the case, it would suggest that diversified menus are a strategy insurers may use to overcome adverse selection when the insurance covers highly predictable and discretionary services.

Policy makers are likely to be interested in the utilisation and selection effects of ancillary health services as these services often exist at the frontier of public provision. While ancillaries are not covered by the public system currently, future policy may change this and understanding whether insurance increases utilisation is critical to evaluating any extension to current public insurance systems. In Australia, this question has particular relevance as the federal Government already spends over \$6 billion annually – almost 13 per cent of federal health expenditure in 2015-16 – on an ongoing rebate to encourage people to purchase private hospital insurance and PAHI[1]. Policy implications of this research are discussed in the conclusion.

The Australian PAHI system has many desirable features for the applied researcher. Insurance status is not tied to employment and therefore reflects personal demand only. Community rating means that all people within each state and territory are charged the same premium for a given policy, regardless of personal risk factors. Therefore any observed selection effects can confidently be attributed to consumer preferences for insurance rather than responses to price. Studies in other institutional environments have often been hampered by the fact that survey data cannot perfectly capture the price setting behaviour of insurers. Administrative data has overcome this problem in some cases, but then results are usually for a specific insurer and may not be generalisable. I use a nationally representative dataset and therefore can make observations about overall market behaviour.

A key difficulty in estimating the effect of insurance on utilisation is dealing with the endogeneity of insurance. The main way that researchers using non-experimental data have addressed this is to use instrumental variables strategies, and this approach is also taken in this paper. I follow Hopkins et al. (2013) and use an indicator variable for whether the respondent wears glasses or contact lenses. The underlying assumption is that wearing glasses or contact lenses will increase demand for PAHI (since almost all policies include coverage for optometry) but will not influence demand for non-optometry health services.

[1] The rebate, which is available for hospital insurance as well, was introduced in 1999 as a 30 per cent government contribution for all Australians with private health insurance. Since July 2012 the rebate has been means tested and more recently changes were made to link the benefit to a base premium rate, effectively reducing the government's contribution over time. Nevertheless, the rebate remains at close to 30 per cent for most people and importantly, these policy changes occurred after the data was collected for this study.

The validity of this instrument is discussed in more detail later. As a robustness check I also use self-reported reasons for purchasing health insurance to identify people who only purchased it to avoid punitive government policies. After conditioning on the variables that force these respondents into insuring (namely income and age), their insurance choice is plausibly exogenous and I can estimate a treatment effect for this group separately.

To explore selection in the PAHI market, the main approach I use is the 'unused observables' test of Finkelstein and Poterba (2014). This approach involves identifying one or more variables that are potentially observed by the insurance provider but not used in setting premiums. If any variable has a statistically significant effect on both the decision to insure and utilisation, conditional on the variables the insurer does use in setting premiums, then we have evidence of selection bias. As noted above, the PAHI market is an ideal setting for this approach because providers are prohibited from adjusting premiums based on personal risk. It is therefore straightforward to identify candidate 'unused variables'.

I find that PAHI does increase utilisation of health services. In particular, people with PAHI are 83% more likely to have visited a dentist in the last 12 months, and 149% more likely to have visited a physiotherapist, chiropractor, osteopath or acupuncturist. There is weaker evidence that PAHI increases utilisation of mental health and naturopathy services. I also identify a number of variables that are a source of adverse selection to insurers. For example, people with more education and females are more likely to insure as well as being more likely to utilise each of the health services I consider. Similarly, income (on average) increases the probability of insurance and the probability of utilisation for dental and optometry. Interestingly, I find little evidence of adverse selection due to objective or subjective measures of health. The one exception is that the number of long-term health conditions a respondent has is positively correlated with both insurance and utilisation. Consistent with other recent work, I also document evidence of favourable selection through some variables, especially for mental health and naturopathy services, highlighting the possibility that people are less likely to self-select based on these lower frequency health services. Importantly, I do not find significantly higher correlations between utilisation of different types of health services for the insured compared to the uninsured. This suggest that consumers are unlikely to be adversely selected based on their total health service use. The diverse range of health services included in PAHI policies may therefore be a market solution to adverse selection.

The paper is organised as follows. I provide a brief review of related literature in section 2. I provide important background information on the PAHI market in section 3. I discuss my data and key variables in section 4. I discuss the empirical strategy in section 5. I present my results in section 6. Section 7 concludes.

2 Previous literature

Empirically, this paper is closely related to Hopkins et al. (2013) who use older versions of the same data series as this paper to estimate the effect of PAHI on the frequency of dental service utilisation[2]. To address the endogeneity of insurance, they use information on whether the respondent wears glasses or contact lenses as an instrumental variable[3], which I follow in this paper. I expand on their research by estimating the causal effect of PAHI on health services other than dental and by undertaking a detailed analysis of selection effects. In addition, by utilising plausibly exogenous variation in the decision to insure created by policy incentives, I also provide an alternative approach to estimating causal effects for dental utilisation, which will provide a useful robustness check to current research.

The small body of research on dental utilisation supports the hypothesis that PAHI causes increased utilisation. Hopkins et al. (2013) look at the frequency of dental visits while Srivastava et al. (2015) look at whether there was a dental visit in the previous year. Srivastava et al. (2015) address the endogeneity of insurance using a simultaneous equations framework. Less is known about the effect of PAHI on utilisation of other health

[2]They use the 1995 and 2001 versions of the Australian National Health Survey whereas as I use the 2011-12 version.

[3]Their instrument is actually an indicator for whether any family member wears glasses or contact lenses. However, they note that their results are not sensitive to using an indicator only for the respondent, which is what I observe in my dataset.

services. Keeffe et al. (2002) find that insurance is positively correlated with visiting either an optometrist or ophthalmologist, although their analysis includes a limited set of controls.

It is worthwhile mentioning briefly the research on private hospital insurance in Australia since private hospital insurance is subject to virtually the same regulatory settings as PAHI and it will therefore be interesting to compare the outcomes of this research to that literature. Studies looking at the effect of hospital insurance on general hospitalisations have found positive effects (Cameron et al., 1988; Savage & Wright, 2003; Lu & Savage, 2006; Srivastava & Zhao, 2008; Cheng & Vahid, 2011; Doiron et al., 2014; Doiron & Kettlewell, 2015). Doiron et al. (2014) show that this result is driven by more discretionary hospital procedures. Since many of the services PAHI covers are discretionary, this suggests the possibility of large utilisation effects. Savage and Wright (2003) identify adverse selection in the hospital insurance market, while Doiron et al. (2008) and Buchmueller et al. (2013) uncover evidence of both both adverse and advantageous selection. The nature of hospital insurance is quite different to PAHI (e.g. low frequency, high cost events vs. high frequency low cost events) and it is unclear whether selection bias will operate symmetrically.

3 Background

There are two main types of health insurance in Australia; private hospital insurance and PAHI. Because all Australian residents can receive free hospital treatment through Medicare, private hospital insurance is predominately purchased to provide greater choice (such as choice of physician) and reduce waiting periods for procedures. Forty-seven percent of people are covered by private hospital insurance. PAHI on the other hand provides benefits for a range of out-of-hospital procedures that receive limited or no public support such as dental, optometry, physiotherapy, osteopathy, naturopathy and dietics. Insurers are generally free to insure any health services provided they are "intended to manage or prevent a disease, injury or condition" (*Private Health Insurance Act 2007* sd.121-10(1)(a)) and are not also covered by Medicare[4], including visits to a general practitioner or specialist and eye tests at an optometrist (insurance for optometry services is generally used to subsidise the cost of corrective eye-wear).

It is difficult to generalise the structure of PAHI policies because insurers exercise a large degree of flexibility in product design[5]. However, there are some broad similarities across policies that can help to generate expectations about behaviour in this market. First, almost all policies provide cover for dental, optical and physical health services and together these services account for more than 80% of claims in the market. Second, insurers cap the maximum benefits receivable for each health service a policy covers. The structure of these caps varies. For example, there may be an individual cap for each service[6] or a combined cap for a group of services (or even all services on the policy.). Third, benefits are subject to personal co-payment rates and these can be high. As an indication, benefits accounted for 53% of fees charged in the 12 months to March 2015 (PHIAC, 2015). Fourth, insurers cannot refuse to insure a customer or price discriminate regardless of pre-existing medical conditions, age, or any other characteristics that may increase utilisation risk[7]. Prices do vary by state because insurers can reduce premiums for states with lower insurance levies and costs of health care. Finally, of the almost 56% of Australians who have PAHI, the majority (85%) choose a combined hospital and PAHI policy (PHIAC, 2015).

The largest single service category for claims is for dental care. Forty-two per cent of claims made were for dental expenses in the 12 months to March 2015. The next largest claims are optometry (13%), physiotherapy (8%) and chiropractic (6%). Similar patterns occur for benefits. Throughout this paper I refer to these health services as 'major' and

[4]There are some exceptions to this. A limited number of therapeutic, oral and maxillofacial, pathology and diagnostic imagery related services can be included in PAHI policies despite also being subsidised through Medicare.

[5]A complete list of policies currently available is at `www.privatehealth.gov.au`.

[6]Typically well below the cost of the policy, so that somebody who regularly uses one health service (e.g. makes weekly visits to a physiotherapist) is unlikely to receive back more from her insurer than she paid for the policy.

[7]Waiting periods can be imposed (usually less than 12 months). These are typically longer for more expensive procedures.

naturopathy and mental health services as 'minor'. Claim and benefit statistics are detailed in Figure 1 for the health services considered in this study.

The discretionary nature of the ancillary health services I study, combined with community rating, should give rise to adverse selection. However, since PAHI covers a range of health services that are likely to be uncorrelated (or weakly correlated), selection effects may be limited to certain health services only. In particular, it is likely that selection bias would be more prominent in the case of the major health services. The potential for adverse selection may also be muted by the decision to purchase private hospital insurance and PAHI being made jointly. If PAHI is a secondary choice, people may not actively 'self-select' into it. It is worth noting that this is not a threat to identification of utilisation effects since the estimation strategy harnesses variation in demand from motivated consumers (i.e. those purchasing PAHI to subsidise corrective eye-wear). Figure 1 here

3.1 Policy incentives

In the late 1990s, a number of policy initiatives were introduced to address declining participation in private hospital insurance. These policies included a tax penalty for high income earners who did not purchase insurance (the Medicare Levy Surcharge (MLS)[8]), a mandate requiring insurers to increase the price of policies by 2% for each year a person remained uninsured since turning 31 (Lifetime Health Cover (LHC) loading), and a 30% premium rebate[9]. Only the rebate also applied to PAHI. These policies resulted in a significant increase in private hospital insurance coverage in 2000[10]. While the MLS and LHC loading do not directly apply to PAHI, the joint nature of the decision to purchase private hospital insur-

[8]In 2011-12 (the data collection period) the MLS thresholds were $80,000/$160,000 for singles/couples and the penalty was 1% of taxable income.

[9]The rebate is increased to 35% and 40% when the policy holder turns 65 and 70 years respectively. I treat these increases as inconsequential, which is supported by work in progress by the author using large sample survey data.

[10]See Palangkaraya and Yong (2005) and Ellis and Savage (2008) for empirical investigations into the separate effect of these policies.

ance and PAHI means that these incentives are likely to affect demand for PAHI as well[11]. Later, I exploit this variation in the incentive to insure to separate selection and utilisation effects.

4 Data

The data for this study are from the 2011-12 Australian National Health Survey (NHS). The NHS is a large random sample survey conducted by the Australian Bureau of Statistics periodically. Survey questions provide measures of health status, income, employment and other demographics expected to influence the demand for health insurance and health care utilisation.

The analysis sample is restricted to adults (defined by being 18 years or older) who are not studying full-time or living at home as a dependent student. These groups would generally be eligible for benefits under their parents' insurance policies and so their insurance status may not be personally determined. People who do not know their insurance status are also dropped from the sample (90 individuals) as well as a small number of individuals who did not provide information on exercise or mental health. People with missing income data are also dropped in the main part of the analysis (1,619 individuals)[12]. The final sample size is 12,261.

A set of independent variables was selected by identifying important correlates of health insurance from previous research (e.g. Hopkins and Kidd (1996) and papers cited in Section 2) and using a variety of health indicators to capture additional variation in utilisation. To capture financial means and security, I use weekly household cash income[13], dummies for different education levels and an employment indicator. Family status is captured by a

[11]This is supported by the fact that in 2000 PAHI coverage increased in parallel with private hospital insurance despite the fact that LHC has been largely attributed to the rise.

[12]Sensitivity analysis is presented in Section 6 where these people are included in the analysis.

[13]Household income is more likely to reflect capacity to purchase insurance and health services in couples and families than personal income. Utilisation results controlling for personal income instead are consistent with those reported in the paper.

coupled indicator and number of children aged 0-17 years. One of the advantages of using the NHS is the detailed information on health status and health risk. Specifically, I include indicators for core movement disability, cancer, arthritis, osteoporosis, diabetes, number of long term health conditions, categorical dummies for self-assessed health, number of days per week undertaking moderate exercise, and an indicator for whether the person smokes. An indicator for government health care card holders is also used because these cards provide some of the same benefits as PAHI (including subsidised dental care) and are therefore likely to reduce demand while also affecting utilisation. Other controls are a female indicator, age, a regional indicator and state and territory dummies. I use an indicator for whether the respondent wears glasses or contact lenses as my main instrumental variable. Details and means for these variables are in Table 1.

Note that apart from income, data are only available for the individual (although region, state and number of children are also likely to be consistent across the family unit). If the decision to purchase PAHI is at the household level, the omission of data on partners and children could introduce some omitted variable bias for the insurance decision. If I did have these data they could be incorporated into the model as additional instruments (Doiron & Kettlewell, 2015), assuming they do not affect own-utilisation, which does not seem unreasonable. Note however that the main instrument (Glasses) does not require conditioning on family characteristics to achieve exogeneity. As such, while it would be interesting to incorporate family information into the analysis this is not critical to the aims of the paper[14].

PAHI status is identified by a question that asks if the respondent is currently covered by health insurance and a follow up question that asks what type of health insurance they have (i.e. PAHI, hospital insurance or both). Fifty per cent of the sample have PAHI.

People with health insurance answer a multiple response follow up question asking them

[14]Family characteristics could be incorporated using other data sets, such as the Household Income and Labour Dynamics in Australia (HILDA) survey. However, HILDA does not include the Glasses instrument or the detailed health data included in the NHS.

to identify the main reasons for purchase. These include "*LHC or avoid age surcharge*" and "*To gain government benefits or avoid extra Medicare levy*". I identify those people who indicate one or both of these policy incentives as their only reason for purchasing health insurance. This is a relatively small number of people (316 people, 5.2% of the insured). Nevertheless, they are sufficient to provide meaningful results. Obviously people who purchase insurance solely because of policy incentives are not a random subset of the population. Critically, I can control for the characteristics that determine selection into this group, namely age and income. Conditional on these variables, participation in PAHI is plausibly exogenous and I can estimate a separate local average treatment effect for this group. This offers a useful robustness check to my main instrumental variables results[15].

Forty-six per cent of people visited a dentist in the last 12 months. This figure is much lower for the other dependent variables. Because physiotherapy, osteopathy, chiropractic and acupuncture are likely to be strong substitutes, I combine these variables in the empirical analysis into a new variable called 'Physical'. Just under 13% of people used one of these services. I also combine psychology and counselling into a single dependent variable called 'Mental'. Four per cent of people visited a psychologist or counsellor, 7% visited an optometrist and 2% of visited a naturopath.

Since policy details are not available in the NHS data, there are some limitations to the analysis that are worth noting. First, I cannot determine whether a respondent has PAHI that provides cover for each of the health services I consider. It is difficult to quantify the extent of this problem as publicly available administrative data do not provide information on policy details. However, a personal search for PAHI policies (as well as information on

[15]One potentially serious concern is that the people who purchase health insurance to avoid the LHC or MLS only need to purchase hospital cover, not ancillaries. Therefore, those who also purchase ancillaries cover may still have self-selected into PAHI. It is well known in Australia that insurers offer policies with minimal hospital coverage for those who simply want to avoid government penalties. However, these low quality hospital policies often include cover for ancillaries, possibly to retain customers who will never actually use their hospital cover. For example, at the time of writing the lowest cost policy for Australia's second largest insurer (advertised as a way to avoid penalties) included ancillaries and comparison against other insurers suggests this strategy is not isolated. Only 20% of respondents in the sample who identified policy incentives as their only reason for purchasing insurance held stand-alone hospital cover.

claims and benefits - see Table 1) leads me to believe that the health services I consider are likely to feature on the majority of policies in the market. In particular, dental, optometry and physical health services (such as physiotherapy) seem to be a standard feature of most policies. To the extent that in some cases people who I treat as insured may actually not have insurance for a particular health service, this would tend to result in an underestimation of the utilisation effect of PAHI.

A second limitation is that I do not observe the price of insurance. It is likely, for example, that utilisation effects will be greater for policies with more generous co-payment rates. While it would be interesting to estimate utilisation responses across the distribution of policies, this is not possible with the data available. The results on utilisation should therefore be interpreted as the average of this distribution, in other words the utilisation response to having some level of insurance.

Table 1 here

5 Model

The utility of insurance to individual i is given by the following linear function

$$y_i^* = \mathbf{Z}_i\beta + \epsilon_i \text{ where } y_i = 1 \text{ if } y_i^* > 0, 0 \text{ otherwise.} \tag{1}$$

where $y_i \in \{0,1\}$ is an indicator for choosing to insure and y_i^* is a latent variable. The vector $\mathbf{Z}_i$ represents the independent variables described in Table 1. By assuming that ϵ_i is normally distributed, with mean zero and constant variance σ (normalised to unity), the probability individual i will choose to insure can be estimated by binary probit regression. The coefficient vector β provides information on the direction of correlation between various personal characteristics and insurance.

A separate linear net-benefit function exists for the utilisation decision. Defining, $h_i \in \{0,1\}$ as an indicator for receiving treatment, this can be written as a function of insurance as follows.

$$h_i^* = \mathbf{X}_i\beta_{\mathbf{ins}} + \alpha y_i + \eta_i; \text{ where } h_i = 1 \text{ if } h_i^* > 0, 0 \text{ otherwise} \tag{2}$$

To identify selection bias in the PAHI market, I estimate separate probit regressions of equations (1) and (2). To estimate utilisation effects, I assume that ϵ_i and η_i are bivariate normal with zero mean, unit variance and $\text{cov}(\epsilon_i, \eta_i) = \rho$ so that the insurance and utilisation choice can be estimated jointly as a bivariate probit. For formal identification, the vectors $\mathbf{Z}_i$ and $\mathbf{X}_i$ may contain exactly the same variables because of the non-linear functional form. However, for causal inference it is undesirable to rely on non-linearity for identification and an exclusion restriction (i.e. a variable that is included in $\mathbf{Z}_i$ only) is recommended.

The main instrument I use is an indicator variable for whether or not the respondent wears glasses or contact lenses (which I call 'Glasses'). Relevance comes from the fact that people who wear glasses and contact lenses expect to visit an optometrist in the future to update prescriptions and replace their corrective eye-wear. These costs are covered under most PAHI policies and quantitatively are the second most important class of claims in the PAHI market.

Establishing instrument validity is more difficult. The Glasses instrument will be valid as long as it only influences the utilisation decision indirectly through insurance and the vector of controls $\mathbf{X}_i$. This does not seem unreasonable for the health services I consider as poor eyesight should in general be unrelated to physical, oral or mental health. Some evidence for this comes from the fact that when I include Glasses in utilisation equations for the expenses I consider, it is never significant (this is not a formal test for instrument validity however). One threat is that the instrument may be invalid because it is driven by reverse causality (i.e. insurance cover causing people to buy glasses and contact lenses who otherwise would not wear them). However, in Australia it is generally free to have your vision tested as this is subsidised through Medicare. In addition, vision tests are also carried out during car licence renewals and for those with poor vision the right to drive can be made conditional on corrective measures. Given these settings, a reverse causality interpretation would argue that people without insurance tend to not treat their vision impairment, since there is no reason to believe the uninsured should be less aware of their vision status. An additional threat is that poor eyesight could be the result of poor health. In particular, having diabetes may affect utilisation of certain health services as well as the probability of wearing glasses or contact lenses. I am able to control for this condition with my data (as well as many other measures of health) and further, can identify those people whose need to wear glasses or contacts is a direct effect of having diabetes[16]. Overall, the assumptions needed for the Glasses instrument to be valid do not seem unreasonable in the context of this paper.

There is one health service that I study for which the Glasses instrument is not valid; optometry. To identify the causal effect of PAHI on the probability of visiting an optometrist a different instrument is needed. The instrument I use is a couple indicator. Being coupled is one of the strongest predictors of purchasing PAHI but should have no direct effect on the need to visit an optometrist. It is worth noting that because I can use Glasses as a control variable for Optom, and this variable should capture a large amount of variation in the decision to visit an optometrist, there may be less pressure on the instrument in this case.

6 Results

6.1 Who buys ancillaries insurance?

In Table 2 are the results from a probit regression on the decision to purchase insurance. The insured differ from the uninsured on a number of dimensions. Age and household income are both positively correlated with insurance. However, this declines at higher levels (both turning points are in the right tails of the distribution at 99 years and \$7,167 respec-

[16]In the main analysis, I deal with this by including a indicator variable for this group. There are relatively few of these people in my data (89 individuals) and unsurprisingly, dropping them from the sample has no effect on my findings.

tively). Females are 5.5 percentage points (ppts) more likely to insure. Education is also important; those with a degree or diploma are around 16.1 and 14.4 ppts more likely to insure respectively relative to those with no tertiary education. Being in a couple increases the probability of insurance by 8.3 ppts. Variables that reduce the probability of insurance include living outside a major city (-8.4 ppts) holding a government health care card (-17.4 ppts) and number of children.

There is mixed evidence for the effect of health and risk variables on the probability of insurance. Self-assessed health is actually negatively correlated with insurance. In particular, those who self-assess their health as 'poor' are 11.6 ppts less likely to insure suggesting these variables may reflect more than just health status, for example attitudes towards risk (Doiron et al., 2008). In terms of objective measures of health, only the number of long-term health conditions and the osteoporosis indicator are statistically significant and positively related to insurance. Being in poor mental health is negatively related to insurance. Smoking is also negatively correlated with insurance and has a large effect (-14 ppts) while exercise and pregnancy have no statistically significant relationship. Importantly, Glasses is strongly correlated with insurance. Wearing glasses or contact lenses increases the probability of insurance by 8.6 ppts. Because instrument relevance is better defined for linear regression models, I also estimated equation (1) by linear OLS to obtain the F-statistic for Glasses and Couple (not shown). These F-statistics are 44.90 and 57.65 respectively, well over the commonly accepted threshold of 10.

table 2 here

6.2 Utilisation

In all instances, the dependent variable is a binary indicator equal to one if the respondent visited a relevant health care provider in the last 12 months. For each type of utilisation, four models are estimated. The first is a binary probit with no controls. In Australia, where insurers are not allowed to vary premiums based on risk factors, the unconditional correlation between insurance and utilisation provides evidence on the overall presence of asymmetric information in the market (Chiaporri & Salanie, 2000)[17]. The second model is a binary probit that controls for $\mathbf{X}_i$. The third model is the bivariate probit model described in the previous section with Glasses used as an instrument in the insurance equation for all regressions except where optometry is the dependent variable, where Couple is used as an instrument. The fourth model is the same as (2) but with people whose only reason for purchasing insurance is to avoid punitive policies treated as a separate group. Only the results for this group are reported and these can be interpreted as local average treatment effects. The main utilisation results are reported in Table 3.

Table 3 here

The unconditional regressions reveal a positive correlation between PAHI and all the health services I consider. This could indicate either adverse selection or utilisation. Only the correlation between insurance and Mental is statistically insignificant. Conditioning on observables reduces the magnitude of the relationships between insurance and utilisation for all the health services except Mental, which is also now statistically significant.

To identify the causal relationship between PAHI and utilisation, we need to also account for possible selection on unobservables. Using the bivariate probit specification, there is evidence that PAHI increases utilisation of dental services. This effect is large in magnitude – insurance is associated with a 26 ppts or 83% increase in the probability of having visited a dentist in the past 12 months. Using the alternative local average treatment effect approach for those avoiding punitive policies, I also find evidence of a causal relationship. This estimate is smaller in magnitude than the bivariate probit. However, this should be interpreted carefully. The bivariate probit model uses information on every respondent to calculate average partial effects (APEs) and therefore is estimating the average treatment effect. Model

[17]This is the so called 'positive correlation test'. Strictly speaking, we need to condition on state and territory indicators to perform this test as insurers can charge different premiums based on this observable. Including these indicators results in correlations that are only marginally different from the unconditional correlations.

(4) on the other hand only uses information on those avoiding punitive policies and therefore is estimating a local average treatment effect for this group. These estimates may differ because of heterogeneity in utilisation responses to PAHI.

For Physical, the APE for the bivariate probit is not significant, however it is of similar magnitude to models (1) and (2). Furthermore, the local average treatment effect result provides statistically significant evidence of a causal relationship. The APE for the bivariate probit implies that insurance increases the probability of having visited a physiotherapist, osteopath, chiropractor or accupuncurist by 9.8 ppts or 149% (although this is imprecise). The results for optometry are inconclusive. PAHI is positively correlated with Optom even when I condition on wearing glasses or contact lenses and the extensive list of controls. However, when using Couple as an instrument and estimating by bivariate probit, the correlation is close to zero (and actually negative). It is possible this result is due to the invalidity of Couple as an instrumental variable, although overidentification tests dicsussed below support its validity for other health services. The local average treatment effect APE is close to that obtained in model (2), however it is not statistically significant.

For the minor health services, if we assume that PAHI is exogenous to these health services, then the results for model (2) provide evidence of a causal relationship. Arguably, PAHI is more likely to be exogenous in the case of these health expenses since they are unlikely to feature heavily in the decision to insure[18]. Under the bivariate probit specification, the relationship between PAHI and Mental is positive and marginally significant [p=0.07]. Instrumenting results in a much larger point estimate. The APE implies that insurance increases the probability of visiting a counsellor or psychologist by 5.4 ppts. For Naturo, the APE is very similar to the unconditional estimate and implies that PAHI increases the probability of visiting a naturopath by 2.1 ppts. While the APE is insignificant, the coefficient is significant at the 10% level. Note the results for model (4) are omitted for the minor health services to avoid misinterpretation[19].

[18] This argument is weakened if a correlation exists between minor and major health service usage.

[19] Purchasing insurance to avoid punitive policies, visiting a mental health professional and visiting a

In sensitivity analysis presented in Table 4, I recovered those with missing income data using (i) dummies for missing income and (ii) single imputation and re-estimated the bivariate probit APEs. In both cases, the results for Dentist, Physical and Mental are similar to those in Table 3, while the APE for Naturo is closer to zero and the APE for Optom is much larger (the coefficient estimate (not shown) is actually significant at the 10% level). I also repeated the analysis using Couple as an additional instrument and obtained similar results to Table 3, with the APE on Physical becoming marginally significant[20]. Overidentification tests on the exogenity of Glasses and Couple support the validity of both instruments[21].

Overall, there is evidence that people respond to PAHI with increased health service utilisation. The evidence is particularly strong in the case of dental and physical health services. There is weaker evidence for mental health and naturopathy.

6.3 Selection

The proceeding results suggest that the net effect of selection bias in the PAHI market may be zero. Controlling for observables tends to reduce the correlation between PAHI and utilisation for most health services (Mental is the exception), while controlling for unobservables using instrumental variable methods generally leads to higher APEs close to the unconditional correlations. Because the confidence intervals for these estimates include the unconditional correlations, I cannot reject the null hypothesis of no selection bias. However, the standard errors on these estimates are also large meaning my test suffers from low statistical power.

To identify which observables contribute to selection bias in the PAHI market, I follow Finkelstein and McGarry (2006) and Finkelstein and Poterba (2014) by comparing the co-

naturopath are all low frequency events and therefore there is insufficient variation in these variables for meaningful analysis.

[20]This does not include Mental. Couple is not a suitable instrument for Mental since being in a couple could have a direct effect on visiting a counsellor or psychologist (this is evident in Table A1).

[21]Because overidentification tests are not well defined for non-linear models, I used the Sargen's statistic from linear two stage least squares regression for this purpose. The Sargen's statistic was 0.355 (p=0.354), 0.459 (p=0.498), 0.103 (p=0.748) for Dental, Physical and Naturo respectively.

efficients from probit regressions on the decision to purchase insurance and the utilisation decision. This procedure requires me to condition on observables the insurer may use in setting premiums. As explained in Section 3, insurers are not allowed to price discriminate in the PAHI market with the exception that premiums can vary by state[22]. For this reason I include a full set of state and territory dummies (7 in total) in my regressions but do not pursue these variables as sources of selection. (It would be impossible to distinguish between self-selection from consumers against price discrimination by insurers for these variables.)

Table 5 shows the direction of correlations for insurance and utilisation[23]. Because we are only interested in coefficients that are statistically significant in both the insurance equation and the utilisation equation, only variables that match this criteria for at least one health service are included in the table (judged at the 5% level). In each cell, a '+' indicates a positive correlation, '-' indicates a negative correlation and '.' indicates no statistically significant relationship. For age and income, where quadratic terms are included, the correlations are based on the APEs. This focuses attention on the overall impact of these variables on selection bias. Adverse selection is identified when the correlation for a particular variable is in the same direction for both the insurance equation and the relevant utilisation equation (i.e when a group who are more (less) likely to purchase PAHI are also more (less) likely to utilise the health service). Where these correlations are in the opposite direction, this indicates advantageous selection to the insurer.

Table 4 here

The results indicate that there are numerous sources of selection bias in the PAHI market. Groups who are adversely selected across all classes of health services are females and those

[22]While LHC is not directly applied to PAHI, the joint nature of the decision to purchase PAHI and hospital insurance means that this policy could feel like a price increase for some people. However, since LHC loading is not the outcome of an optimal pricing strategy for insurers, it is not obvious that it should be controlled for. Moreover, since LHC does not affect the revenue insurers receive for PAHI, even if it influences the relationship between PAHI and age it is still experienced as selection bias to the insurer (i.e. is not compensated by price).

[23]The APEs for each utilisation equation are reported in Appendix Table A1.

with a degree or diploma. Each of these variables is positively correlated with both insurance and utilisation. It is not clear what the mechanism is for these associations. Risk aversion in one possibility. Evidence suggests that females are more risk averse than males (e.g. Hartog et al., 2002; Eckel & Grossman, 2002; Borghans et al., 2009) and this would tend to increase demand for insurance as well as preventative health services. However, research has also found that risk aversion is negatively related to education (e.g. Hartog et al., 2002; Brunello, 2002; Belzil & Leonardi, 2007, 2013), so risk preferences cannot explain both of these results. While it would be interesting to explore the mechanisms of these associations further, the aims of this paper are simply to show that observables can be linked to selection bias in the PAHI market, so this is left for future work.

While the number of long-term health conditions does not feature in Table 4, this variable is positively correlated with every health service and is positive and only marginally insignificant ($p=0.056$) in the insurance equation. Apart from this result, there is little evidence for selection on health characteristics, which highlights the limits in restricting attention to 'classic' sources of selection bias. Smokers are less likely to purchase insurance, but also less likely to utilise each of the health services except Mental. Potential explanations for this are that the cost of smoking reduces money available for health care, or that smoking is correlated with unobserved attitudes towards health care utilisation (e.g. low risk aversion). Income is associated with adverse selection for Dentist and Optom, with people on higher income more likely to insure and more likely to use these services. Age is a source of favourable selection for Physical and Mental, with older people more likely to insure but less likely to use these services.

For the major health services, there are several instances where adversely selected groups for a particular service are not adversely selected for another service. For example, income is a form of adverse selection for Dentist and Optom but not Physical. However, there are also several groups who are adversely selected across all utilisation dimensions, namely females, the highly educated and non-smokers. This suggests that it may be similar consumers who are driving adverse selection across all major health services, in which case diversified policy menus may not address selection effects. This conclusion is premature however. For example, the probabilities for PAHI, Dentist and Physical are all higher for a highly educated female non-smoker than the inverse. However, this does not imply that the joint probability of Dentist and Physical is higher for a highly educated female non-smoker. In other words, there may be low correlation between different types of health service utilisation regardless of whether consumers with the same features are responsible for adverse selection.

To explore this, it is useful to examine the correlation matrix for the major health services. To begin, I report below the overall correlation matrix (**C**), which shows that the strongest relationship between major health services is between Physical and Optom (r=0.28). This indicates that 8% of the variation in Optom can be explained by Physical, which does not seem particularly high. The next largest correlation is between Dentist and Physical and it suggests that only 2% of the variation in the Physical can be explained by Dentist. In the context of Ellis and McGuire (2007), the major ancillaries health services have low 'predictiveness', which is expected to help to overcome adverse selection.

$$
\mathbf{C} = \begin{array}{c} \\ \textit{Dentist} \\ \textit{Physical} \\ \textit{Optom} \end{array}
\begin{array}{c}
\begin{array}{ccc} \textit{Dentist} & \textit{Physical} & \textit{Optom} \end{array} \\
\begin{pmatrix} 0 & & \\ 0.145 & 0 & \\ 0.117 & 0.284 & 0 \end{pmatrix}
\end{array}
$$

Next I consider the correlations conditional on PAHI status. Let **C1** be to correlation matrix for the insured and **C2** be the correlation matrix for the uninsured. If consumers are not self-selecting based on expected probability of using multiple health services (i.e. are primarily choosing insurance based on expected utilisation of a single type of service), then the off-diagonal elements in $\mathbf{C1} - \mathbf{C2}$ should be close to zero.

In the results below, it is clear that differences in the propensity to utilise multiple health services are not significantly driven by insurance status. In fact, the largest element in $\mathbf{C1} - \mathbf{C2}$ is only r=0.08 (the difference in correlation between Physical and Optom). This estimate is statistically significant but is not economically large - the difference in r^2 values is 0.04, which implies that variation in Physical explains only 4.1 ppts more of the variation in Optom for the insured. The differences in correlations for the other elements are statistically insignificant and close to zero.

$$\mathbf{C1} - \mathbf{C2} = \begin{array}{l} \\ Dentist \\ Physical \\ Optom \end{array} \begin{array}{c} \begin{array}{ccc} Dentist & Physical & Optom \end{array} \\ \begin{pmatrix} 0 & & \\ -0.002 & 0 & \\ 0.021 & 0.079 & 0 \end{pmatrix} \end{array}$$

A more formal approach is to conduct the so called positive correlation test (Chiaporri & Salanie, 2000) on the change in correlation between health service j and k when insurance status changes. Because we are dealing with joint decisions it is reasonable to work with bivariate probabilities. Let P_j and P_k be the bivariate probit probability of utilising health service j and k respectively. I first jointly estimate $(P_j|PAHI = 1)$ and $(P_k|PAHI = 1)$[24]. The covariance between error terms, $\rho_{i=1}$, captures correlation between j and k unrelated to insurance status. Next I jointly estimate $(P_j|PAHI = 0)$ and $(P_k|PAHI = 0)$ and obtain $\rho_{p=0}$. The difference in covariance, $\rho_{i=1} - \rho_{i=0} = \tilde{\rho}_{jk}$, should equal zero if there is no self-selection on the joint probability of utilisation, since it reflects the change in covariance unrelated to insurance. The estimates and p-values (in parentheses) for the major health services are as follow; $\tilde{\rho}_{DP} = -0.044$ (0.259), $\tilde{\rho}_{DO} = 0.026$ (0.569) and $\tilde{\rho}_{PO} = 0.070$ (0.203) where D, P, and O denote Dentist, Physical and Optom respectively. These values are small and statistically insignificant. Altogether, these results imply that consumers are likely to self-select in a narrow way (i.e. based on a single health service dimension). This suggests that diversified policy menus may indeed help to explain how insurers deal with adverse

[24] This simply involves estimating a bivariate probit regression with only a constant term on the sub-sample of respondents with PAHI.

selection in the PAHI market.

For the minor health services, I find more evidence of heterogeneity in selection effects. Couples are both more likely to buy insurance and less likely to visit a counsellor or psychologist. People who score high on the Kessler 10 measure of mental distress are less likely to insure, but have expectedly higher utilisation of mental care. This is also the case for Naturo. In addition, people living outside major cities are less likely to insure but more likely to visit a naturopath. The fact that many consumers are likely to be favourably selected in regards to the minor health services indicates their possible use as a further way of reducing the risk borne by insurers.

7 Conclusion

This paper contributes to the research on selection and utilisation in insurance markets by studying these phenomenon in the Australian PAHI market. I find evidence that PAHI does increase utilisation of various health services. In particular, it increases visits to the dentist and physical health specialists. I find weaker evidence that it increases utilisation of mental health and naturopathy services. To assess whether there are selection effects in the PAHI market I identify variables that are correlated with both the insurance and utilisation decision. I find numerous sources of selection bias that often vary across utilisation types. I identify a number of groups that are adversely and advantageously selected from the perspective of insurers. The data suggest that consumers who self-select are likely to do so based on a single health service, which implies that diversified policy menus could be a strategy for addressing adverse selection in the PAHI market.

While this paper does not directly evaluate the policy environment for PAHI, it is worth noting that the strict premium setting and non-refusal restrictions placed on insurers are not crippling the market with adversely selected consumers. In fact, it is likely that insurers are managing this issue through policy design. This brings into question the rationale for current subsidies for purchasing PAHI, which are one of the largest outlays of the Federal health budget.

References

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*, 488–500.

Belzil, C., & Leonardi, M. (2007). Can risk aversion explain schooling attainments? Evidence from Italy. *Labour Economics*, *14*, 957–970.

Belzil, C., & Leonardi, M. (2013). Risk Aversion and Schooling Decisions. *Annals of Economics and Statistics*, *111/112*, 35–70.

Borghans, L., Golsteyn, B. H. H., Heckman, J. J., & Meijers, H. (2009). Gender Differences in Risk Aversion and Ambiguity Aversion. *Journal of the European Economic Association*, *7*, 649–658.

Brunello, G. (2002). Absolute risk aversion and the returns to education. *Economics of Education Review*, *21*, 635–640.

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.

Cameron, A. C., Trivedi, P. K., Milne, F., & Piggot, J. (1988). A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia. *Review of Economic Studies*, *55*, 85–106.

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.

Chiaporri, P., & Salanie, B. (2000). Testing for Asymmetric Information in Insurance Markets. *Journal of Political Economy*, *108*, 56–78.

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.

Colombo, F., & Tapay, N. (2004). *Private Health Insurance in OECD Countries: The*

Benefits and Costs for Individuals and Health Systems. OECD Health Working Papers No. 15.

Doiron, D., Fiebig, D. G., & Suziedelyte, A. (2014). Hips and hearts: The variation in incentive effects of insurance across hospital procedures. *Journal of Health Economics*, *37*, 81–97.

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.

Doiron, D., & Kettlewell, N. (2015). *Using partner's information to identify the causal effects of insurance on utilization.* (unpublished manuscript)

Eckel, C. C., & Grossman, P. J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior*, *23*, 281–295.

Ellis, R. P., & McGuire, T. G. (2007). Predictability and predictiveness in health care spending. *Journal of Health Economics*, *26*, 25–48.

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.

Finkelstein, A., & McGarry, K. (2006). Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market. *American Economic Review*, *96*, 938–958.

Finkelstein, A., & Poterba, J. (2014). Testing for Asymmetric Information Using "Unused Observables" in Insurance Markets: Evidence From the U.K. Annuity Market. *The Journal of Risk and Insurance*, *81*, 709–734.

Hartog, J., Ferrer-i Carbonell, A., & Jonker, N. (2002). Linking Measured Risk Aversion to Individual Characteristics. *Kyklos*, *55*, 3–26.

Hopkins, S., & Kidd, M. P. (1996). The determinants of the demand for private health insurance under Medicare. *Applied Economics*, *28*, 1623–1632.

Hopkins, S., Kidd, M. P., & Ulker, A. (2013). Private Health Insurance Status and Utilisation of Dental Services in Australia. *The Economic Record*, *89*, 194–206.

Keeffe, J. E., Weih, L. M., McCarty, C. A., & Taylor, H. R. (2002). Utulisation of eye care services by urban and rural Australians. *British Journal of Opthalmology*, *86*, 24–27.

Lu, M., & Savage, E. (2006). *Do financial incentives for supplementary private health insurance reduce pressure on the public system? Evidence from Australia.* CHERE working paper 2006/11.

Palangkaraya, A., & Yong, J. (2005). Effects of Recent Carrot-and-Stick Policy Initiatives on Private Health Insurance Coverage in Australia. *The Economic Record*, *81*, 262–272.

PHIAC. (2013). *Competition in the Australian Private Health Insurance Market.* Private Health Insurance Administration Council: Research Paper 1.

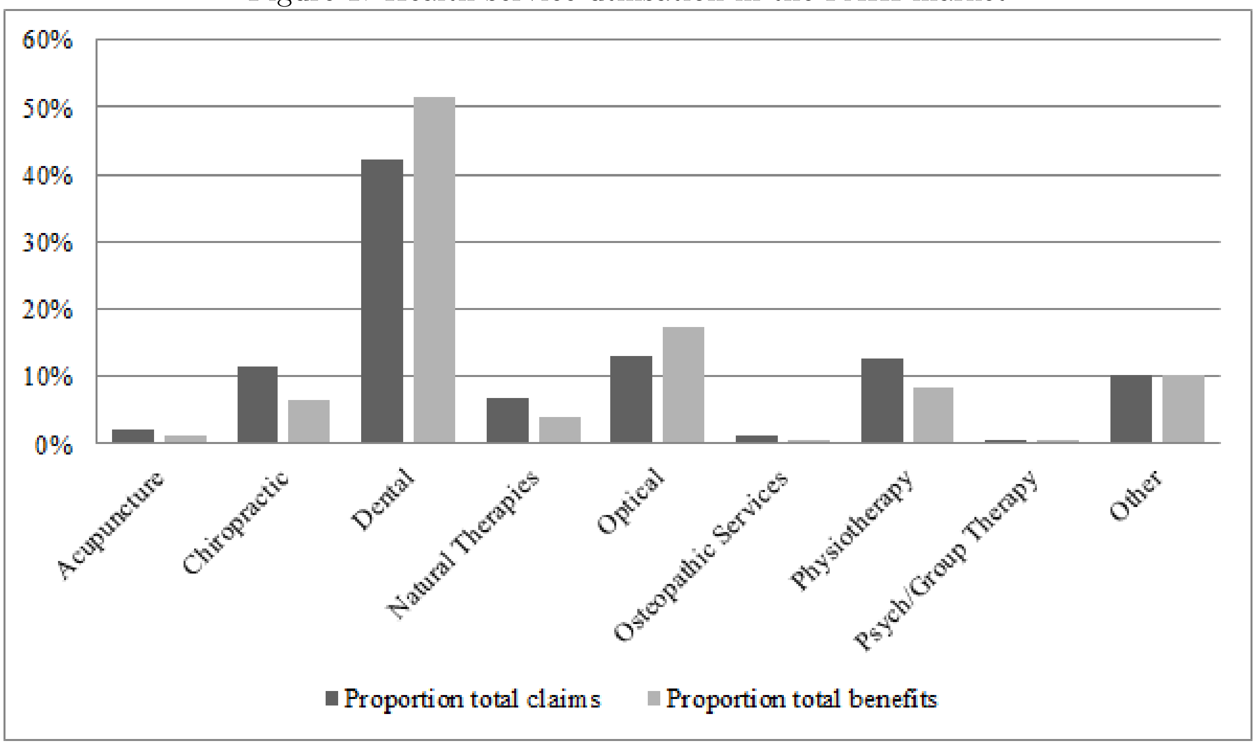
PHIAC. (2015). *PHIAC A Report.* Private Health Insurance Administration Council: `http://www.apra.gov.au/PHI/Publications/Pages/Industry-Statistics.aspx`. (accessed 25-07-2015)

Savage, E., & Wright, D. (2003). Moral hazard and adverse selection in Australian private hospitals: 1989-90. *Journal of Health Economics*, *22*, 331–359.

Srivastava, P., Chen, G., & Harris, A. (2015). Oral Health, Dental Insurance and the Demand for Dental Service Use in Australia. *Health Economics*, *Forthcoming*.

Srivastava, P., & Zhao, X. (2008). *Impact of Private Health Insurance on the Choice of Public versus Private Hospital Services.* HEDG Working Paper 08/17.

Figure 1: Health service utilisation in the PAHI market



Source: (PHIAC, 2015). Claim/benefit statistics are for the 12 months to March 2015. Acupuncture also includes claims for acupressure.

Table 1: Main variables

Variable	Mean (s.d.)
PAHI - has private ancillaries health insurance	0.500
PAHI_exog - purchased health insurance to avoid LHC/MLS	0.017
Dentist - visited dentist in last 12 months	0.463
Accup - visited accupuncturist in last 12 months	0.017
Chiro - visited chiropractor in last 12 months	0.052
Naturo - visited naturopath in last 12 months	0.020
Optom - visited optometrist in last 12 months	0.074
Osteo - visited osteopath in last 12 months	0.014
Physio - visited physiotherapist in last 12 months	0.074
Couns - visited counsellor in last 12 months	0.015
Psych - visited psychologist in last 12 months	0.029
Age - continuous (years)	49.44 (16.87)
Female - female dummy	0.521
Regional - lives outside major city	0.378
Income - weekly cash household income	1699 (1457)
Employed - employment dummy	0.649
Degree - degree highest level of education	0.252
Diploma - diploma highest level of education	0.107
Cert - certificate highest level of education	0.242
Nonenglish - English not main language at home	0.072
Overseas - born overseas	0.278
Couple - coupled (married or de facto)	0.550
Nochild - number of children aged 0-17	0.584 (0.992)
H_excell - Self-assessed health (SAH) excellent	0.187
H_vgood - SAH very good	0.348
H_fair - SAH fair	0.121
H_poor - SAH poor	0.046
Ltcond - number of long-term health conditions	3.678 (3.145)
Disabled - has core activity limitation	0.213
K10_high - high category for Kessler 10 mental distress	0.111
Diabetes - has diabetes	0.120
Eye_diabetes - vision impairment directly due to diabetes	0.007
Arthritis - has arthritis	0.220
Osteopor - has osteoporosis	0.054
Cancer - currently has some form of cancer	0.159
Preg - currently pregnant (only females)	0.012
Gov_HCC - has government health care card	0.349
Glasses - wears glasses or contact lenses	0.642
Smokes - current daily smoker	0.184
Days_exer - times did moderate exercise in last week	0.852 (1.827)

Note: Sample size is 12,261. Means for state and territory dummies are excluded for brevity.

Table 2: Probability of purchasing PAHI

Variable	Coefficient	SE	APE	SE
Age	0.023	(0.005)***	0.004	(0.001)***
Age_sqr[a]	-0.119	(0.046)**		
Female	0.138	(0.026)***	0.055	(0.010)***
Regional	-0.217	(0.033)***	-0.084	(0.012)***
Income[a]	0.405	(0.026)***	0.102	(0.007)***
Income_sqr[b]	-0.028	(0.003)***		
Employed	0.003	(0.039)	0.001	(0.015)
Degree	0.406	(0.035)***	0.161	(0.014)***
Diploma	0.363	(0.043)***	0.144	(0.017)***
Cert	0.025	(0.032)	0.010	(0.013)
Nonenglish	-0.345	(0.055)***	-0.131	(0.020)***
Overseas	-0.209	(0.031)***	-0.081	(0.012)***
Couple	0.209	(0.028)***	0.083	(0.011)***
Nochild	-0.049	(0.014)***	-0.016	(0.005)***
H_excell	0.011	(0.038)	0.004	(0.015)
H_vgood	0.045	(0.031)	0.018	(0.012)
H_fair	-0.059	(0.043)	-0.023	(0.017)
H_poor	-0.306	(0.072)***	-0.116	(0.026)***
Ltcond	0.011	(0.006)*	0.004	(0.002)*
Disabled	-0.042	(0.039)	-0.016	(0.015)
K10_high	-0.107	(0.044)**	-0.042	(0.017)**
Diabetes	-0.054	(0.040)	-0.021	(0.016)
Eye_diabetes	-0.259	(0.153)*	-0.099	(0.056)
Arthritis	-0.001	(0.036)	-0.000	(0.014)
Osteopor	0.123	(0.057)**	0.049	(0.023)**
Cancer	0.031	(0.036)	0.012	(0.014)
Preg	-0.134	(0.115)	-0.052	(0.044)
Gov_HCC	-0.475	(0.034)***	-0.174	(0.015)***
Smokes	-0.372	(0.034)***	-0.140	(0.013)***
Glasses	0.217	(0.032)***	0.086	(0.013)***
Days_exer	0.002	(0.007)	0.001	(0.002)

Note: Sample size is 12,261. Robust standard errors in parentheses. APE is average partial effect. For binary variables, this is the change in the predicted probability when switching values from 0 to 1. For continuous variables, this is the mean of the vector of individual marginal effects for sample observations. Standard errors on APEs are calculated using the delta method. Coefficients on state and territory controls are omitted for brevity. *,** and *** is significance at the 10%, 5% and 1% level respectively.
[a] Row values have been multiplied by 1000.
[b] Row values have been multiplied by 1000^2.

Table 3: Impact of PAHI on health care utilisation

	(1) Probit - no controls	(2) Probit - with controls	(3) Bivariate probit	(4) Probit - exog. PAHI
	Dependent variable = Dentist			
APE	0.259 (0.008)***	0.207 (0.010)***	0.258 (0.103)**	0.132 (0.029)***
Pseudo R^2	0.0496	0.081		0.081
$\hat{\rho}$			-0.096	
	Dependent variable = Physical			
APE	0.104 (0.006)***	0.076 (0.006)***	0.098 (0.069)	0.049 (0.022)**
Pseudo R^2	0.032	0.097		0.097
$\hat{\rho}$			-0.062	
	Dependent variable = Optom			
APE	0.048 (0.005)***	0.029 (0.004)***	-0.021 (0.080)	0.027 (0.017)
Pseudo R^2	0.017	0.148		0.148
$\hat{\rho}$			0.286	
	Dependent variable = Mental			
APE	0.004 (0.003)	0.006 (0.002)***	0.054 (0.030)*	
Pseudo R^2	0.0004	0.178		
$\hat{\rho}$			-0.345*	
	Dependent variable = Naturo			
APE	0.019 (0.002)***	0.009 (0.002)***	0.021 (0.014)	
Pseudo R^2	0.023	0.119		
$\hat{\rho}$			-0.112	

Note: Sample size is 12,261. Robust standard errors in parentheses. APE is average partial effect, which is the change in the predicted probability when switching PAHI values from 0 to 1. Standard errors on APEs are calculated using the delta method. In the bivariate probit (3), Glasses used as an instrument except for Optom where Couple is used as an instrument. *,** and *** is significance at the 10%, 5% and 1% level respectively.

Table 4: Sensitivity of bivariate probit APE estimates

Dep. Var.	(1) Dummy for missing income	(2) Single imputation for missing income	(3) Glasses and Couple as instruments
Dentist	0.343 (0.079)***	0.355 (0.076)***	0.197 (0.092)**
Physical	0.084 (0.065)	0.091 (0.067)	0.105 (0.056)*
Optom	0.094 (0.062)	0.101 (0.063)	
Mental	0.041 (0.023)*	0.042 (0.024)*	
Naturo	0.002 (0.014)	0.003 (0.014)	0.017 (0.014)

Note: Sample size in (1) and (2) is 14,612 and in (3) is 12,261. Income imputation in (2) uses the variables in Table 1. Robust standard errors in parentheses. APE is average partial effect, which is the change in the predicted probability when switching PAHI values from 0 to 1. Standard errors on APEs are calculated using the delta method. *,** and *** is significance at the 10%, 5% and 1% level respectively.

Table 5: Sources and direction of selection effects

Variable	PAHI	Dentist	Physical	Optom	Mental	Naturo
Age	+	.	-	.	-	.
Female	+	+	+	+	+	+
Regional	-	-	.	.	.	+
Income	+	+	.	+	.	.
Degree	+	+	+	+	+	+
Diploma	+	+	+	+	+	+
Nonenglish	-	.	-	.	-	.
Couple	+	.	.	.	-	.
Nochild	-	-	.	.	.	.
K10_high	-	.	.	.	+	+
Gov_hcc	-	.	.	.	+	.
Smokes	-	-	-	-	.	-
Glasses	+	.	.	+	.	.

Note: '+', '-' and '.' indicates a positive, negative and statistically insignificant correlation between independent (column) and dependent variable (row) judged at the 5% level. Correlations are obtained via separate binary probit regressions. See Table 1 for additional control variables.

Appendix

Table A1: Impact of controls on health care utilisation

	Dentist		Physical		Optom		Mental		Naturo	
Variable	APE	SE	APE	SE	APE	SE	APE	SE	APE	SE
Age	0.001	(0.001)*	-0.001	(0.000)***	0.000	(0.000)	-0.001	(0.000)***	-0.000	(0.000)
Female	0.079	(0.010)***	0.023	(0.006)***	0.009	(0.004)***	0.009	(0.002)***	0.012	(0.002)***
Regional	-0.051	(0.012)***	-0.003	(0.007)	0.002	(0.005)	-0.001	(0.003)	0.007	(0.002)***
Income	0.027	(0.006)***	0.003	(0.003)	0.004	(0.003)	0.000	(0.002)	0.001	(0.002)
Employed	-0.010	(0.015)	0.030	(0.008)***	-0.002	(0.005)	0.004	(0.003)	0.001	(0.003)
Degree	0.120	(0.013)***	0.064	(0.009)***	0.032	(0.006)***	0.026	(0.005)***	0.007	(0.003)**
Diploma	0.101	(0.016)***	0.058	(0.012)***	0.029	(0.008)***	0.016	(0.006)***	0.014	(0.005)***
Cert	0.047	(0.012)***	0.012	(0.008)	0.006	(0.005)	0.009	(0.004)***	0.008	(0.003)***
Nonenglish	-0.015	(0.020)	-0.047	(0.010)***	-0.010	(0.007)	-0.015	(0.003)***	-0.001	(0.004)
Overseas	-0.005	(0.012)	-0.012	(0.007)*	-0.007	(0.004)*	0.000	(0.003)	-0.003	(0.002)
Couple	0.010	(0.011)	0.003	(0.006)	-0.003	(0.004)	-0.008	(0.003)***	-0.001	(0.002)
Nochild	-0.017	(0.005)***	-0.004	(0.003)	-0.004	(0.002)	-0.001	(0.002)	-0.002	(0.001)
H_excell	0.056	(0.014)**	-0.027	(0.008)***	-0.011	(0.005)**	-0.009	(0.003)***	-0.001	(0.002)
H_vgood	0.023	(0.012)**	-0.005	(0.007)	-0.002	(0.004)	-0.006	(0.003)**	0.000	(0.002)
H_fair	-0.008	(0.017)	-0.013	(0.009)	0.000	(0.006)	0.005	(0.004)	-0.001	(0.003)
H_poor	-0.011	(0.026)	0.000	(0.015)	-0.000	(0.008)	0.008	(0.007)	0.006	(0.006)
Ltcond	0.011	(0.002)***	0.010	(0.002)***	0.007	(0.001)***	0.005	(0.001)**	0.002	(0.001)***
Disabled	-0.000	(0.014)	0.058	(0.010)***	0.013	(0.005)**	0.008	(0.004)**	-0.002	(0.002)
K10_high	0.013	(0.017)	0.009	(0.010)	0.012	(0.007)*	0.043	(0.007)***	0.001	(0.003)
Diabetes	-0.010	(0.015)	-0.010	(0.009)	0.022	(0.006)***	0.001	(0.004)	-0.002	(0.002)
Eye_diabetes	-0.034	(0.056)	-0.021	(0.030)	0.008	(0.018)	-0.016	(0.004)***	0.002	(0.011)
Arthritis	-0.010	(0.013)	0.007	(0.008)	-0.004	(0.004)	-0.006	(0.003)**	0.002	(0.002)
Osteopor	0.019	(0.022)	0.026	(0.014)*	-0.003	(0.006)	-0.002	(0.005)	-0.002	(0.003)
Cancer	0.038	(0.014)***	0.003	(0.008)	0.010	(0.005)**	0.005	(0.004)	0.006	(0.003)**
Preg	-0.085	(0.042)**	0.024	(0.028)	-0.017	(0.013)	0.003	(0.010)	0.009	(0.009)
Gov_HCC	0.018	(0.016)	-0.009	(0.009)	-0.010	(0.005)*	0.007	(0.004)**	-0.000	(0.003)
Smokes	-0.029	(0.013)**	-0.041	(0.007)***	-0.016	(0.005)***	-0.002	(0.003)	-0.007	(0.002)***
Days_exer	0.011	(0.002)***	0.004	(0.001)***	0.001	(0.001)	-0.000	(0.001)	0.001	(0.001)**
Glasses					0.059	(0.004)***				
Pseudo R^2	0.081		0.097		0.148		0.178		0.119	

Note: Sample size is 12,261. Results shown are for binary probit with the full set of controls (Model 2). APE is average partial effect. For binary variables, this is the change in the predicted probability when switching values from 0 to 1. For continuous variables, this is the mean of the vector of individual marginal effects for sample observations. Standard errors on APEs are calculated using the delta method. Marginal effects on state and territory controls are omitted for brevity. *,** and *** is significance at the 10%, 5% and 1% level respectively.