Family formation and demand for health insurance*

Denise Doiron
University of New South Wales

Nathan Kettlewell†
University of Technology Sydney

December 2019

Abstract

We study how demand for health insurance responds to family formation using a unique panel of young Australian women. Our data allow us to simultaneously control for the influence of state dependence and unobserved heterogeneity as well as detailed information on children and child aspirations. We find evidence that women purchase insurance in preparation for pregnancy but then transition out of insurance once they have finished family building. Children have a large, negative impact on demand for insurance, although this effect is smaller for those on higher incomes. We also find that state dependence has a large impact on insurance demand. Our results are robust to a variety of alternative modelling strategies.

Keywords: Health insurance; Family formation; State dependence

*Although substantially different, this paper stems in part from early research conducted with Vineta Salale, Denzil Fiebig and Elizabeth Savage. We are grateful to them for their input. We also are thankful to two anonymous referees, the Editor Sally Stearns, and the participants of the 2015 Australian Health Economics Society Conference. The research on which this paper is based was conducted as part of the Australian Longitudinal Study on Women's Health by the University of Queensland and the University of Newcastle. We are grateful to the Australian Government Department of Health for funding and to the women who provided the survey data.

†Corresponding author: University of Technology Sydney, Economics Department, PO Box 123, Broadway, Sydney, NSW, 2007, Australia. Email: Nathan.Kettlewell@uts.edu.au

1 Introduction

Child aspirations and children are likely to have complicated effects on incentives to purchase private health insurance (PHI). However, there is limited empirical evidence on the effect of family formation on demand for PHI. A better understanding of this is valuable since most individuals will have children at some stage. Changes in fertility patterns could therefore have large effects on the overall market. In this paper, we use a longitudinal sample of young Australian women for the period 1996-2006. A key feature of our data are the detailed records on the age and number of children and child aspirations, which allow us to closely examine patterns in insurance around different stages of family formation while controlling for state dependence and individual specific heterogeneity.

The treatment of family formation in the PHI literature to date has taken the form of various controls for family composition (Kiil, 2012, reviews studies for countries with universal health care). Typically, researchers use a single indicator for the presence of dependent children. Depending on the data and model specification, within the Australian context studies find this variable to have no effect, (Hopkins & Kidd, 1996) significant positive (Srivastava & Zhao, 2008) or negative (Cheng & Vahid, 2011) effects on health insurance cover. Similarly controls for the number of children has yielded conflicting results (Doiron et al., 2008; Johar et al., 2011). This sensitivity is consistent with children having complex and conflicting effects on insurance demand. Barret and Conlon (2003) find that being married with children increases the probability of being insured, additional children reduce the probability, and having a child under five increases the probability. This suggests a dynamic response to the presence and age of children. Doiron and Kettlewell (2018) provide evidence that desire for children increases demand while Salale (2006) finds no significant effect in a dynamic model.[1]

As detailed later, strict regulations around underwriting, pricing and the separation of insurance from employment makes Australia an ideal case study for research on PHI. In ad-

[1]Salale (2006) uses the same data as this paper and focuses on establishing the presence of state dependence and testing the correlates of PHI more generally. Our paper extends her analysis by using more years of data and more precise controls for family formation.

dition, the institutional settings are broadly consistent with most other developed economies with mixed public/private insurance systems (Colombo & Tapay, 2004).

Our main empirical model of insurance demand is a dynamic probit with a one period lag on insurance status. We use the Wooldridge (2005) conditional maximum likelihood approach to control for unobserved time invariant heterogeneity and the endogeneity of initial insurance status. As a sensitivity check, we also estimate linear dependent variable models that involve weaker distributional assumptions but ignore the discrete nature of insurance cover. Additional sensitivity analysis is undertaken where we control for attrition bias using inverse probability weighting procedures (Wooldridge, 2002).

We find that family formation affects demand for PHI among young women. Our preferred estimates imply that women who desire additional children are 2.6 percentage points (ppts) (6%) more likely to be insured. Overall, the presence of children reduces demand, however responses are smaller for higher income earners. On average additional children reduce the probability of insurance cover by 5.4 ppts (13%). This is consistent with expenditure on children crowding out expenditure on insurance. Women retain insurance in the first year after childbirth and some women enter the market at this stage. However, after one year there is a strong negative relationship between the age of the youngest child and insurance, which suggests women transition out of insurance when they are not family building. Overall, our results support a pattern of insurance whereby women purchase in preparation of pregnancy and then leave insurance once they finish family building. We also estimate a large role for state dependence. Being insured in the previous period (three years prior) increases the probability of insurance by 25.4 ppts (62%) today.

The paper is organised as follows. In Section 2 we provide background information on the PHI market in Australia. In Section 3 we provide a theoretical framework for our empirical analysis. In Section 4 we discuss our data and variables for family formation. In Section 5 we present some descriptive results. In Section 6 we present our empirical models. Our main empirical results are presented in Section 7. Section 8 concludes.

2 Regulatory settings

All Australians can receive free hospital treatment in public hospitals through a system of universal health care (Medicare). Privately insured patients can use their cover for the cost of treatment in either a public or private hospital. Co-pays often apply and vary by plan. Private clinicians are also free to set fees which can create gaps between costs and insurer reimbursements. Public hospitals are a high quality alternative to private care; the main benefits of going private are greater choice over physicians (public patients must take the first available doctor) and shorter waiting times for elective procedures. Additionally, some patients may prefer treatment in a private hospital because of greater privacy and comfort.[2] During our sample period around 70% of births were in public hospitals, with most other births in private hospitals (Laws & Sullivan, 2004).

The regulatory settings for PHI in Australia mean that decisions to purchase are predominately due to consumers' cost/benefit analysis, making its demand easier to estimate and interpret. PHI is not tied to employment, hence accounting for selection into employment and employer-provided insurance is not needed. All consumers have access to the same products – insurers cannot price discriminate (within the same state) or refuse to insure a person based on utilisation risk. Waiting periods of up to one year are imposed for pre-existing conditions and, relevant to our study, obstetrics. Consequently, women who wish to be insured during their pregnancy must purchase insurance before becoming pregnant.

Between 1996-2006 PHI coverage went from around 34% to 46%, largely due to policy incentives to insure (see Palangkaraya & Yong, 2005; Ellis & Savage, 2008). These incentives do not affect our modelling strategy and are discussed in more detail in the Appendix.

[2]Many people also purchase PHI that includes cover for out-of-hospital services that receive no or limited public support, such as dental, optical and allied health. A small number of women in our sample (around 6% each wave) purchase ancillaries health insurance only. These women are treated as uninsured for the purposes of our study.

3 Theoretical framework

To motivate our empirical strategy, we follow Costa and García (2003) in framing the insurance decision as a choice between maximising the expected utility from purchasing PHI (V_{phi}) or using the public system (V_{pub}).[3] We assume decision making from the perspective of an individual (equivalently, this could be expressed as household utility where the household is a single agent). Insurance is purchased for some fixed period (noting that in Australian, PHI can usually be cancelled at any time). An individual will insure if

$$V_{phi} - V_{pub} > 0. \tag{1}$$

V_i depends on income y, the perceived quality of private care Q^1, public care Q^0 (with $Q^1 \geq Q^0$), the probability of hospitalisation η and the cost of private insurance π. Utility is assumed to be increasing in y and Q^i but at a decreasing rate. We further assume that the enjoyment of goods depends on health, with U denoting the good health state and u denoting the bad health state ($U_y \geq u_y$) and that quality care is a normal good i.e. $u_{yQ} > 0$ (found empirically in numerous Australian studies e.g. Hopkins & Kidd, 1996; Doiron et al., 2008; Johar et al., 2011; Buchmueller et al., 2013; Doiron et al., 2014). The individual utility expectation functions are as follows.

$$V_{phi} = \eta u(Q^1, y - \pi) + (1 - \eta)U(y - \pi) \tag{2}$$

$$V_{pub} = \eta u(Q^0, y) + (1 - \eta)U(y) \tag{3}$$

Using this framework, we can make *ceteris paribus* predictions about the role of family formation on demand for PHI. First, consider the desire for additional children. This can be expressed as an increase in η (higher risk of hospitalisation), which will increase the

[3]Costa and García (2003) also allow for patients to cover their own expenses in a private hospital. While this is possible in the Australian system, it is uncommon for most procedures (including childbirth) due to the prohibitive costs of treatment and access to high quality public treatment. We therefore omit this option for simplicity.

likelihood for (1) since $u(Q^1, y-\pi) - U(y-\pi) > u(Q^0, y) - U(y)$.

Hypothesis 1: Demand for insurance will be higher for women who desire more children.

Now consider what happens when children are present. We assume that parents continue to receive utility from income y. However, now income can be spent on the child as well as the parent. Parents receive utility from the child's welfare and the marginal utility from spending on the child initially offers the highest marginal returns so that the utility curve in parenthood U^{par} always sits above the utility curve for non-parenthood.[4] Because the utility curve for parents is steeper, purchasing insurance at fixed price π lowers utility for parents more than for non-parents (spending on children crowds out other spending, including on insurance). On the other hand, if η represents the risk of hospitalisation for either parent or child – i.e. parents move to the low utility state if their child is sick – then it is unclear whether children will increase or decrease the probability of insuring (it is also possible that parenthood changes preferences for Q). For sufficiently large increases in η, the additional risk of hospitalisation effect will exceed any marginal utility of income effect, and children will be associated with greater probability of insuring.

Two empirical regularities help us to make firmer predictions by focusing on risk and the age of children. First, hospitalisation risk is particularly high in the first year of life.[5] Second, hospitalisation risk is decreasing with age through ages 1-15 (Australian Institute of Health and Welfare, 2019).[6] Together, these facts imply that the risk effect due to η is likely to dominate the crowding effect immediately after birth, but will be less likely to dominate as the child gets older. We hypothesise that:

Hypothesis 2: Women will be more likely to be insured in the first year after childbirth.

[4]Note that children are freely included on family PHI policies, so are unlikely to affect π directly (although single parent policies are more costly than single adult policies).

[5]in Australia, 84% of deaths for children under 5 occur in the first year (National Health Performance Agency, 2014) and a study of New South Wales hospitals found that 16.5/100 babies were readmitted in the first year in 2009 (Lain et al., 2014). η may also increase due to the risk of maternal readmission, although this effect is likely to be weaker since only around 3% of women are readmitted within six weeks of childbirth in Australia (Ford et al., 2012).

[6]Our data only included young parents (< age 35); children older than 15 are therefore virtually absent.

Hypothesis 3: The probability of insurance cover will fall as the age of the youngest child rises.

Next, we consider the role of income. Since quality care is a normal good by assumption, higher income will lead to greater probability of insuring, and there will be some threshold y^* above which all women insure. This means that incentive effects from children will only be binding for the subset of women with sufficiently low income. Therefore:

Hypothesis 4: High income women will be less responsive to the negative income effect from the presence of children.

Finally, it is worth discussing state dependence, which has been found to affect demand in health insurance markets (e.g. Bolhaar et al., 2012; Handel, 2013). State dependence is easily incorporated into the above framework by assuming that individuals face transaction costs when switching into or out of insurance (this can be justified by e.g. procrastination, hassle, default bias). This is a potentially important control since state dependence may confound our estimates for the above hypotheses (which are based on dynamics in incentives to insure). Modelling state dependence also matters for understanding how demand shifting events (e.g. parenthood) affect insurance take-up into the future. We therefore control for state dependence in our empirical work.

4 Data

Our data are from the Australian Longitudinal Survey of Women's Health (ALSWH) Young Cohort. The ALSWH Young Cohort began in 1996 with a representative sample of women aged 18-23 years. Self-completion questionnaires are submitted approximately once every three years.[7]

[7]Women living in rural and remote areas were intentionally sampled at twice the rate of other women in order to capture the heterogeneity of health service for this group (Lee et al., 2005). We control for urbanisation in our empirical analysis.

There were 14247 respondents in wave 1. Attrition is non-trivial, mainly due to high levels of mobility, changes in surname, not having telephone listings and low voter registration (Lee et al., 2005). For reasons discussed below, we use a balanced panel of respondents for waves 1–4 (survey years 1996, 2000, 2003, 2006). After balancing our sample across the first four waves and dropping respondents who did not provide information on PHI we are left with 6624 women. (We discuss attrition further in the Appendix.)

Table 1 shows the proportion of women with PHI across the waves and transitions into and out of insurance. Overall, the probability that a woman has insurance in any period conditional on having had insurance in the previous period is 0.81, indicating significant persistence in our raw data.[8]

Table 1 also shows that transition rates to and from insurance generally increase over time. The one exception is wave 2, where the rate of coverage drops slightly and several women leave insurance. This may partially reflect women dropping off their parents' policies when they finish studying or become independent. Later we conduct sensitivity tests by restricting our sample to the later waves where dependence is unlikely.

To measure child aspirations we use the following question: "*by age 35, would you like to have: no children, 1 child, 2 children or 3 or more children*". Data on actual children was only available from wave 3, when mothers were asked to record birth dates for each child. Because we do not observe child age until wave 3, and use only four waves, we use a balanced panel. The variable $D > A$ is an indicator for whether desired children by age 35 exceeds actual children today. We treat women with 3 or more children as having completed their families.[9]

It is worth noting that our definition for child aspirations is not based on current intensity, but instead reflects longer term goals. If the preferences that women form in wave 2 (most

[8]For comparison, we also calculate this statistic using data from the 2004 and 2009 Household Income and Labour Dynamics in Australia Surveys (HILDA). This is a large, nationally representative panel dataset. The corresponding figure for women aged 22-32 is 0.80, almost identical to the figure for our sample.

[9]14.25% of women with children have 3 or more in our sample. The risk that some of these women actually desire more children would result in downwardly biased estimates for the effect of child aspirations.

aged 22-27) are stable, then all the variation in $D > A$ will come from women no longer desiring additional children once they reach their desired family size.[10] Indeed, for the 33% of women whose desire for additional children changes over the sample period, 74% of this variation comes from women no longer desiring more children (see Appendix Table A1). The remaining variation comes from women who previously did not desire additional children changing their preferences. Since we expect demand to spike when women are actively trying to become pregnant, our estimates may be considered a lower bound of the effect of short-term child aspirations.

To measure the 'shifting' effect having a child has on PHI demand, we include a dummy for presence of any children. This variable is also interacted with an indicator for high income since we expect high income earners to be less vulnerable to dropping insurance when they have children. To capture the 'intensity' effect of children we include a continuous variable for the number of children. Because health risks are particularly high in the first year of life, we expect that women may be inclined to maintain insurance during this period. Consequently, we include a dummy for if the woman has a child less than one year old. To allow for a gradual exit from PHI between children, we control for the age of the youngest child using a continuous variable (measured in days) starting at zero one year after birth for the most recent child. This modelling approach is supported by descriptive evidence presented later.[11]

Other variables that we use are typical of research on PHI in Australia and include age, work/study status, highest educational qualifications, country of birth, relationship status, State and Territory dummies, rurality and perceived access to hospital care (see e.g. Hopkins & Kidd, 1996; Barret & Conlon, 2003; Doiron et al., 2008; Buchmueller et al., 2013). Health and risk preference related controls include the SF-36 general health score (see Ware et al.,

[10]Infertility may also affect our source of identifying variation; howver, we do not view this as a serious concern. As reported in Boivin et al. (2007), around 10% of couples in developed countries fail to conceive after 12 months. This is similar to self-reported fertility problems in our sample (9.4%). Also, of the women in our sample who report fertility problems, 58% are parents by wave 4.

[11]We emphasize that despite the time lag between interviews, the fact that we know the exact age of children allows us to capture much of the dynamics in insurance. Thus, we control for the birth and the aging of children born between waves. The lagged insurance variable (state dependence) will also measure the insurance status of mothers before the birth of children born between interviews.

2000), body mass index, risky alcohol intake, smoking and drug use. Each of these variables may influence access to or preference for PHI. We have somewhat restrictive information on income and so use a variety of variables to capture financial resources: a high income dummy (household average gross weekly income exceeds $1000), categorical dummies for money stress, the Index of Economic Resources (IER) score for the respondent's region, a government Health Care Card dummy[12] and a dummy for household income equalling personal income. Finally, in all models we include a set of time dummies.[13]

Since insurance and family formation decisions are likely to be at the household level, our lack of spousal data is a potential source of omitted variable bias. However, we expect this bias to be small since most of the controls we expect to be strong predictors of PHI demand are consistent across family members (e.g. controls for finances, region, access, family composition and time period). Indeed, we find later that conditional on these, person-specific characteristics are generally insignificant predictors of PHI demand, which can reasonably be expected to carry over to omitted spouse characteristics.

Table A2 lists the variables included in our empirical analysis along with definitions. In Table A3, we provide means for each variable across the four waves. In wave 2, 87% of women desire more children. Predictably, this rate decreases across waves. However, even in wave 4 (where most women are 28-33 years old) 67% of women still desire more children. The question on child aspirations was not asked after this wave and consequently our analysis sample comprises the first four waves only. By wave 4 almost half the women in our sample have at least one child.

[12]While Health Care Cards do not subsidise private hospital care, they may affect demand for insurance by subsidising dental and at least add to our financial controls since they are means tested.

[13]To deal with nonresponse while maintaining a reasonable sample size, we retain those observations with missing information (other than for PHI) and use dummy variables as controls. Definitions for these variables are provided in Table A2.

5 Descriptive evidence

In Figure 1 we plot the relationship between age and coverage for women who desire and do not desire additional children, grouped by whether they already have a child. Women with children are always less likely to be insured than childless women, conditional on child aspirations. This gap is particularly noticeable for women aged 22–25 years, while for older women the impact of having a child is small. Women who desire more children consistently insure at higher rates than those who do not, particularly from age 25 (where you would expect child aspirations to become pressing for many women). This gap is of a similar magnitude whether or not the woman already has children. It is consistent with women retaining insurance after the first child is born if they desire additional children after this.

Next we consider the impact of time since most recent child (Figures 2 and 3). Three patterns emerge. First, in the first year after childbirth, insurance rates are fairly stable (they actually increase slightly). The stability may be due to inertia and/or because children this age are particularly vulnerable to hospitalisation. Second, after 1-year there is a clear negative trend in the probability of insurance, which is noisy but approximately linear. Third, when we condition on the desire to have additional children (Figure 3) we see that while the probability of insurance remains stable for those who desire additional children, it falls for those who have finished family building, particularly in the second year.[14] This pattern indicates an attachment to insurance for those still family building but gradual exit from insurance for those who finish family building.

While these descriptive results are striking, they may be driven by omitted factors. In the next section we present our empirical strategy for identifying the role of family formation conditional on observed and unobserved factors.

[14]To keep sample sizes large, we focus on the 24 months since the last child was born in Figure 3.

6 Model

Our main specification models demand for insurance as a dynamic probit regression with a single lag for PHI.[15] Specifically, the demand for PHI can be written as:

$$PHI^*_{it} = X_{it}\beta + \rho PHI_{i,t-1} + \alpha_i + u_{it} \quad i = 1, \ldots, n \quad t = 2, 3, 4 \tag{4}$$

$$PHI_{it} = \begin{cases} 1, & \text{if } PHI^*_{it} > 0 \\ 0, & \text{otherwise} \end{cases}$$

where PHI^*_{it} is the net benefit of PHI, X_{it} is a vector of conditionally exogenous variables including our markers for family formation (X_{it} does not include a constant term), α_i is an individual specific fixed effect and u_{it} is a normally distributed error term with unit variance.

Controlling for α_i is challenging with binary outcome variables since techniques available for the linear case generally result in biased and inconsistent estimates. Furthermore, when lagged dependent variables are included coefficients will be inconsistently estimated unless the initial insurance status is exogenous or the dynamic process is in long-term equilibrium with time-invariant distributional properties (Heckman, 1981).

Based on Mundlak (1978) and Chamberlain (1984)'s correlated random effects approach, we model α_i as a linear function of the exogenous covariates. Following Wooldridge (2005), we include the initial state as a conditioning variable to overcome the initial conditions problem while controlling for unobserved fixed effects. We specify α_i in the following way:

$$\alpha_i = \alpha_0 + \alpha_1 PHI_{i,1} + X_{i,2}\beta_0 + \bar{X}_{i,t>2}\beta_1 + \eta_i \tag{5}$$

$$\text{where } \eta_i | PHI_{i,1}, X_{i,2}, \bar{X}_{i,t>2} \sim N(0, \sigma^2_\alpha)$$

In equation (5), $X_{i,2}$ is the vector of exogenous variables measured at first period values

[15] The restriction to one lag is common in this literature and is compelling in this case since we have a small number of waves and the waves are 3 years apart.

and $\bar{X}_{i,t>2}$ is the within means for periods $t > 2$.[16] By substituting equation (5) into equation (4), this model can be estimated as a standard random effects probit. Average partial effects (APEs) can be estimated in the regular way after multiplying coefficients by $(1 + \hat{\sigma}_{\alpha}^{2})^{-1/2}$.

We also follow Bolhaar et al. (2012) and estimate linear fixed effects models as a sensitivity check. We deal with dependence between lagged PHI and the within-mean of the random error term by using the Arellano and Bond (1991) GMM estimator.[17]

To assess sensitivity to possible non-random attrition, we use inverse probability weighting (IPW) as detailed in Wooldridge (2002). Our IPW approach involves estimating the probability of non-attrition in wave 1 conditional on observables (including some variables not included in X_i) and then using these probabilities as weights for our main equation. This approach is not valid when there are random effects so we apply it using pooled probit regression and compare to pooled probit estimates without weighting. Further details are in the Appendix.

Finally, a discussion about endogeneity is in order. Our empirical strategy identifies the effect of family formation on demand for PHI conditional on unobserved time invariant heterogeneity, a large set of time varying observables and state dependence. The main threats to identification are reverse causality and time varying unobservables correlated with the error term. We are not overly concerned about reverse causality in our setting since the lack of private insurance is unlikely to affect family formation in an environment with high-quality, universal public insurance. We are also able to control for reverse causality to a large extent by conditioning on past PHI status.

[16] We exclude the first period from the within-means based on evidence in Rabe-Hesketh and Skrondal (2013) that doing so reduces bias and results in estimates comparable to the more computationally intensive Heckman (1981) estimator. The intuition for this is that the exogenous variables measured closer to the initial period have more explanatory power and imposing a time-constant relationship between these variables and the unobserved effect may throw away valuable information. In developing our empirical results we also specified α_i as a function of the exogenous variables measured at every period. This resulted in coefficient estimates for our main variables that were very similar to those reported in the paper.

[17] The Arellano and Bond (1991) approach involves taking first differences of all variables to eliminate individual fixed effects and using lagged values of the insurance status as instruments for the change in lagged insurance status to deal with the correlation with the error term e.g. if $t = 4$ then $\Delta PHI_{i,t-1} = PHI_{i,3} - PHI_{i,2}$ can be instrumented by $PHI_{i,2}$ and $PHI_{i,1}$.

The remaining identification issue is due to potential omitted factors that vary with time. One such example would be variations in partner's income and employment which could affect eligibility for family welfare benefits and in turn preferences for children. In this example, we would expect our controls for household finances to absorb much of the effect. Any remaining bias should be negative (i.e. against H1) for the variable $D > A$; increases in family income will decrease benefits which should reduce preferences for children, while income is positively correlated with PHI. Of course a different example could lead to different conclusions.

We believe the most obvious threats due to omitted factors come from family formation affecting unobserved factors (e.g. preferences) that then affect demand (i.e. mechanisms), or from events that reliably co-occur with changes in family formation. In both cases, while a strictly causal (*ceteris paribus*) interpretation of our estimates may not hold, our estimates remain informative for real-world application, where factors are not held constant.

With observational data, we cannot rule out omitted time varying factors. We believe our results are nevertheless likely to be informative for policy making. However, we caution against a strict causal interpretation.

7 Results

Our main results for family formation and state dependence are summarised in Table 2. We focus our discussion on the results for the Wooldridge (2005) dynamic probit model and treat our other estimates as robustness checks.

7.1 Main findings

Overall, our results support the hypotheses posed in Section 3. Women who desire more children are more likely to be insured (H1) and the presence and age of children affects insurance demand. Desire for more children increases the probability of insurance by 2.6 ppts

(6% increase relative to pooled sample mean). There is mixed evidence for a 'shifting' effect for having any children. The coefficient for presence of children is negative but insignificant while the interaction term between high income and presence of children is positive and significant. This indicates heterogeneity in response to children – high income earners are less likely to decrease demand in response to parenthood, consistent with H4. The APE for presence of children, which takes into account the interaction term and measures the overall effect of parenthood on insurance, is close to zero. The APE for the interaction term, which measures the shift in probability of insurance conditional on being high income both before and after childbirth, is positive and larger in magnitude but is not statistically significant.

Additional children have a large, negative effect on insurance demand. An additional child reduces the probability of insurance by 5.4 ppts (13%) – more than twice the absolute effect size of desire for additional children. As expected, women do not drop their cover in the first year after childbirth (H2) and in fact are 3.8 ppts more likely to be insured during this period. This may reflect heterogeneous responses to children, specifically that for some women having a child incentivises them to insure. However, after the child turns 1 year, women gradually transition out of insurance (H3). Between 1-2 years after the birth of the last child, the probability of insurance decreases by 1.7 ppts (4%). The complicated relationship between PHI demand and children can explain why previous studies using basic controls have delivered inconsistent results.

The data also support the presence of state dependence. Being insured in the previous period increases the probability of insurance today by 25.4 ppts (62%). The initial period status is also significant, supporting its inclusion.[18] To demonstrate the interaction between state dependence and family dynamics, we calculate the probability of insurance in each wave for a hypothetical scenario in which all women are induced to insure in wave 2 (due to

[18]Results for other significant coefficient estimates are in Table A4. Focusing on those results that are consistent across specifications, we find that age, marriage, living in an urban area and low money stress increase the probability of being insured. These results are as expected and consistent with previous Australian research on the general population. Health is not a significant predictor, which is unsurprising given our sample comprises young and generally healthy women.

desire for children) and reach their desired family size of two children by wave 3 (youngest aged 180 days). The sample average probability of insurance increases from 33% to 57% in wave 3, primarily due to state dependence. Having a child under one year also increases demand, while presence of children partially offsets these effects. If women retain insurance in wave 3, the probability in wave 4 increases again to 61% while if women drop insurance in wave 3, the probability in wave 4 falls to 35% – this difference reflects the role of state dependence.[19] Another way of stating the role of state dependence is to note that it would take around 7.5 years from birth of the last child for the probability of insurance to fall to the level immediately before that birth for a woman maintaining insurance.[20]

7.2 Sensitivity

Our overall findings regarding family formation are robust to different specifications. Compared to our preferred estimates, pooled probit differ quantitatively (though not qualitatively) for some variables (e.g. number of children) indicating that controlling for unobserved heterogeneity matters. Correcting for attrition however has almost no effect quantitatively or qualitatively suggesting that selectivity is well accounted for with our set of controls. Results using linear fixed effects are similar to our main models despite the fact they do not control for state dependence. This again suggests that controlling for unobserved time invariant heterogeneity is important; however, controlling for state dependence may be relatively unimportant for obtaining unbiased estimates of the family formation variables.

Finally, in the last column of Table 2 we provide estimates for the Arellano-Bond fixed effects model. Because this model works on first differences, we lose the first wave of data and therefore have less variation in our family formation variables. Nevertheless, our main conclusions are robust. Desire for additional children is positive although no longer signif-

[19]The reason that the probability increases in the retention scenario compared to the previous wave is because changes in demographic characteristics that increase insurance demand in our sample (e.g. income, marriage) offset the effect of age of youngest child.

[20]This estimate is a crude approximation using the APEs in Table 2 and assumes the woman is not a high income earner.

icant. Number of children and age of the youngest child remain negatively correlated and highly significant while the presence of children interacted with high income remains positive and significant.[21]

8 Conclusion

In this paper we estimate the demand for PHI for young women focusing on the role of family formation. Specifically, we consider the role of desire for additional children (i.e. expected future pregnancy), childbirth, age of the youngest child and total number of children. Overall, our results support a pattern of insurance whereby women purchase in preparation of pregnancy and then leave insurance once they finish family building. At the same time, some women are likely to become 'locked-in' to insurance since we estimate a large role for state dependence.

Our results are important in the context of declining fertility rates, smaller families, and later first pregnancy in Australia and other developed countries. They imply less women entering the PHI pool and entering at an older age. On the other hand, the probability of insuring is decreasing in family size, so there may be greater willingness to retain insurance as families become smaller. Our theoretical framework suggests that the reason women drop PHI is because expenditure on the child crowds out expenditure on PHI. It is therefore possible that policies that affect these costs (e.g. publicly funded childcare) have second order effects on PHI take-up. Policy makers need to also be aware that changes in the quality gap for obstetrics care between public/private hospitals is likely to influence demand for PHI. Self-reports suggest quality factors like choice of physician and continuity of care are key considerations for women's preferences for birth settings (Stevens et al., 2016).

[21] We also re-estimated our models treating wave 2 as the initial period to address concerns that some of the transitions out of insurance between waves 1 and 2 are due to young women exiting their parents' policies (see Appendix Table A5). Such exists might bias downward estimates of state dependence, with ensuing bias to other parameter estimates. Our results for family formation are generally not sensitive to omitting wave 1, although some estimates are predictably less precise. Notably, the estimate for state dependence is much higher. For the Wooldridge dynamic probit, the APE is 0.38.

Finally, our findings suggest interesting avenues for future research. For example, one path would be to analyse the overall impact of family formation on selection bias in health insurance markets. Handel (2013) finds that inertia in health insurance actually reduced the extent of adverse selection for a large US employer and resulted in positive welfare effects. While purchasing insurance for childbirth is an example of classic adverse selection, in combination with state dependence it may result in younger, lower risk individuals entering the insurance market.

References

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application the employment equations. *Review of Economic Studies*, *58*, 277–297.

Australian Institute of Health and Welfare. (2019). *Admitted patient care 2017–18: Australian hospital statistics.* Health services series no. 90. Cat. no. HSE 225. Canberra: AIHW.

Barret, G. F., & Conlon, R. (2003). Adverse selection and the decline in private health insurance coverage in Australia. *The Economic Record*, *79*, 279–296.

Boivin, J., Bunting, L., Collins, J. A., & Nygren, K. G. (2007). International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Human Reproduction*, *22*, 1506–1512.

Bolhaar, J., Lindeboom, M., & van der Klaw, B. (2012). A dynamic analysis of the demand for health insurance and health care. *European Economic Review*, *56*, 669–690.

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.

Chamberlain, G. (1984). Panel Data. In S. Griliches & M. Intriligator (Eds.), *Handbook of Econometrics.* Elsevier.

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.

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.

Costa, J., & García. (2003). Demand for private health insurance: how important is the quality gap? *Health Economics*, *12*, 587–599.

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. (2018). The effect of health insurance on the substitution between public and private hospital care. *The Economic Record*, *94*, 135–154.

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.

Ford, J. B., Algert, C. S., Morris, J. M., & Roberts, C. L. (2012). Decreasing length of maternal hospital stay is not associated with increased readmission rates. *Australian and New Zealand Journal of Public Health*, *36*, 430–434.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, *103*, 2643–2682.

Heckman, J. J. (1981). The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process. In C. F. Manski & D. McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications.* MIT Press.

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

Johar, M., Jones, G., Keane, M., Savage, E., & Stavrunova, O. (2011). Waiting times for elective surgery and the decision to buy private health insurance. *Health Economics*, *20*, 68–86.

Kiil, A. (2012). What characterises the insured in universal health care systems? A review of the empirical evidence. *Health Policy*, *106*, 60–75.

Lain, S. J., Roberts, C. L., Bowen, J. R., & Nassar, N. (2014). Trends in New South Wales infant hospital readmission rates in the first year of life: a population-based study. *Medical Journal of Australia*, *201*, 40–43.

Laws, P. L., & Sullivan, E. A. (2004). *Australia's mothers and babies 2001.* Australian Institute for Health and Welfare: Perinatal Statistical Series No. 13.

Lee, C., Dobson, A. J., Brown, W. J., Bryson, L., Byles, J., Warner-Smith, P., & Young, A. F. (2005). Cohort profile: The Australian Longitudinal Study on Women's Health. *International Journal of Epidemiology*, *34*, 987–991.

Mundlak, Y. (1978). On the pooling of time series and cross sectional data. *Econometrica*, *46*, 69–85.

National Health Performance Agency. (2014). *Healthy Communities: Child and maternal health in 2009–2012.*

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.

Rabe-Hesketh, S., & Skrondal, A. (2013). Avoiding biased versions of Wooldridge's simple solution to the initial conditions problem. *Economics Letters*, *120*, 346–349.

Salale, V. (2006). *Modelling Dynamic Choice: Private Health Insurance in Australia* (Honours thesis). School of Economics: University of New South Wales.

Srivastava, P., & Zhao, X. (2008). *Impact of private health insurance on the choice of public versus private hospital services.* HEDG Working Paper 08/17.

Stevens, G., Miller, Y. D., Watson, B., & Thompson, R. (2016). Choosing a model of maternity care: Decision support needs of Australian women. *Birth Issues in Perinatal Care*, *43*, 167–175.

Ware, J. E., Snow, K. K., & Kosinski, M. (2000). *SF-36 Health Survey: Manual and Interpretation Guide.* Lincoln, RI: Quality Metric Incorporated.

Wooldridge, J. M. (2002). Inverse probability weighted M-estimators for sample selection, attrition, and stratification. *Portuguese Economic Journal*, *1*, 117–139.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, non-linear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*, *20*, 39–54.

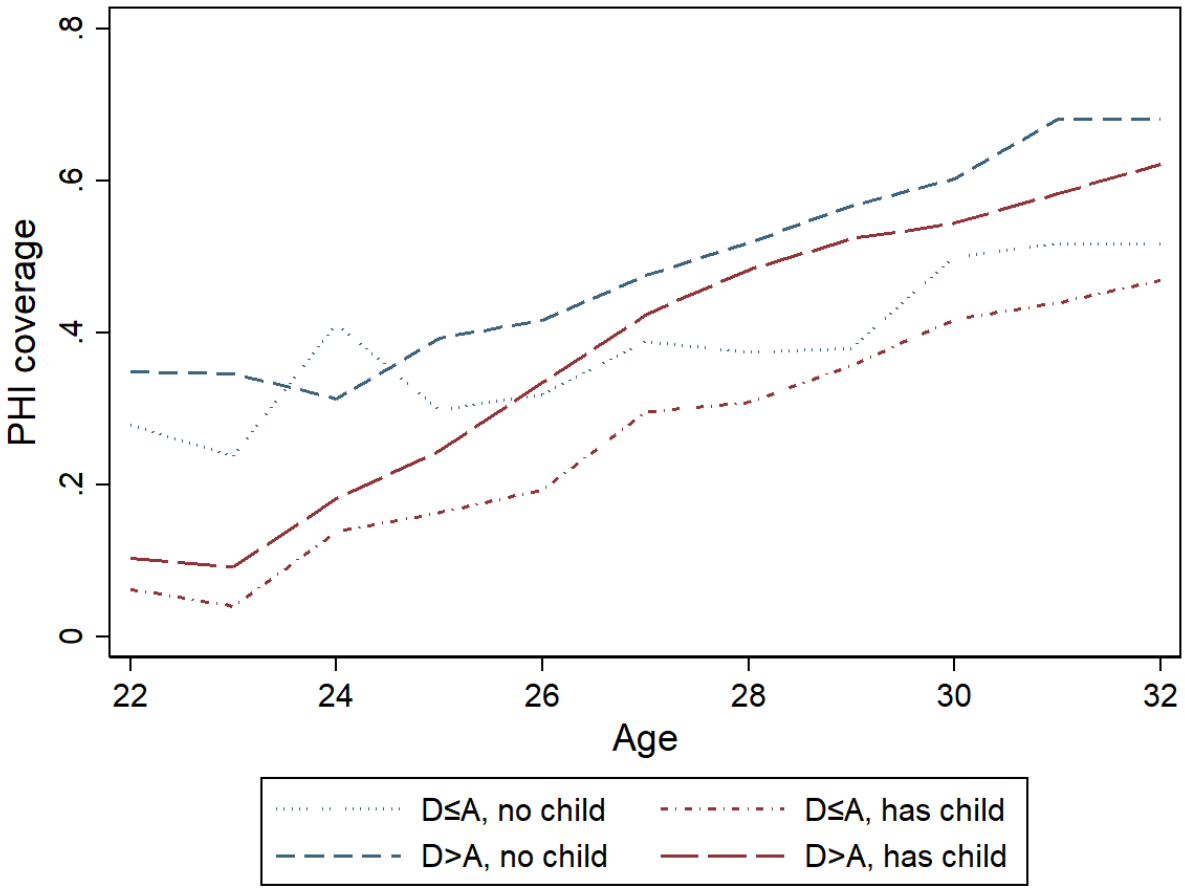
Tables and Figures

Table 1: PHI coverage and transitions in and out of coverage

	Wave 1	Wave 2	Wave 3	Wave 4
PHI	0.35	0.32	0.45	0.54
Joiners		676	1107	862
Leavers		888	244	250

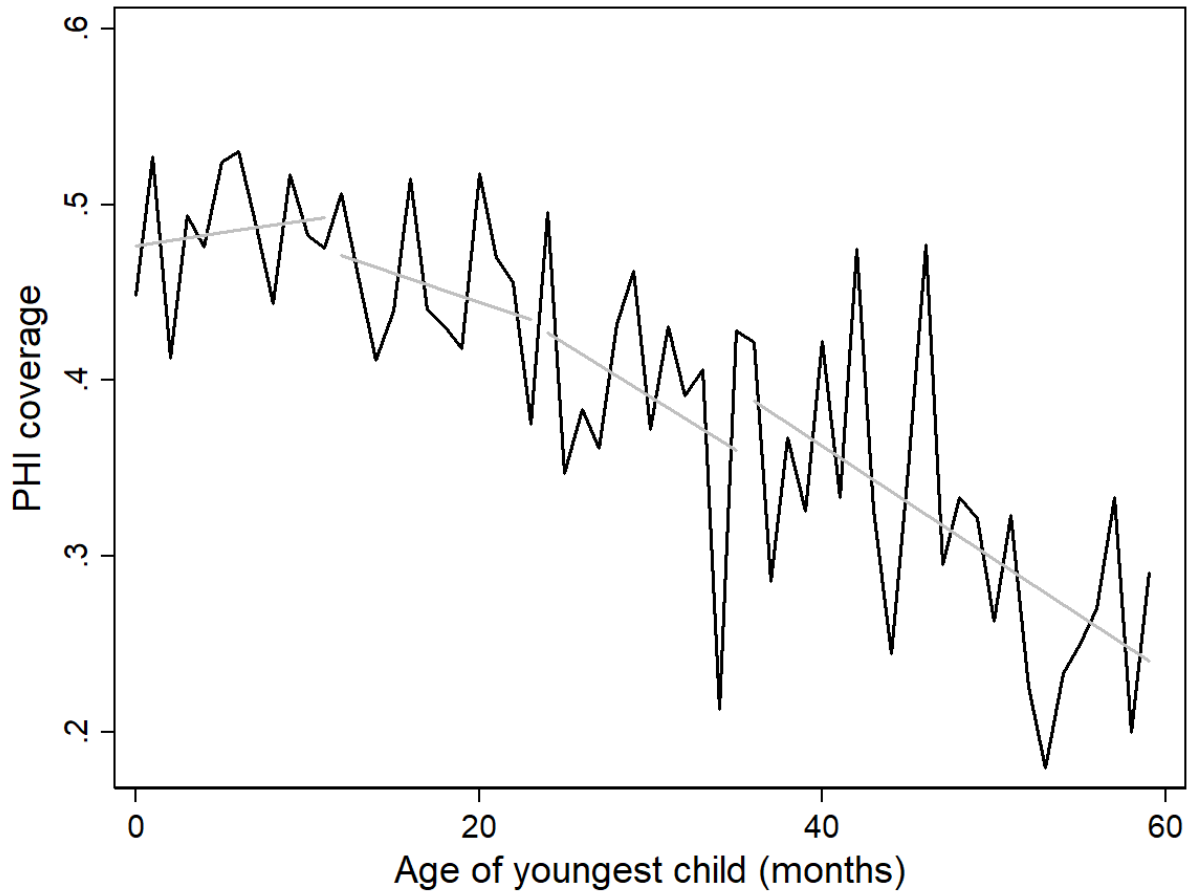
Note: Sample size for each wave is 6624 women.

Figure 1: PHI coverage by age, desired and actual children



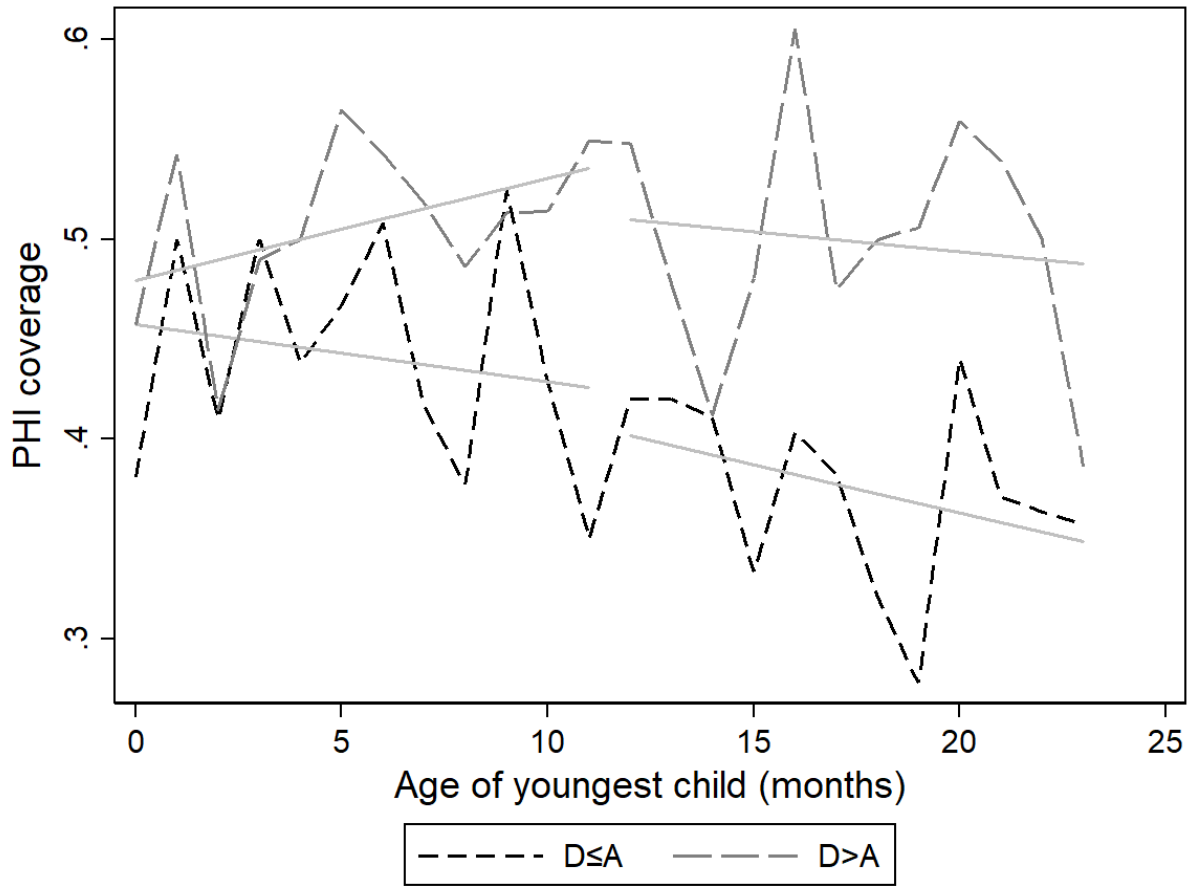
Note: Includes all respondents in wave 2-4 aged 22-32 years (N=19739). Age has been rounded down to the last birthday.

Figure 2: PHI coverage by age of youngest child



Note: Includes all respondents in wave 2-4 who do not have a youngest child 60 months or older (N=18965). Age of youngest child has been rounded down to nearest month.

Figure 3: PHI coverage by age of youngest child and desire for more children



Note: Includes all respondents in wave 2-4 who do not have a youngest child 24 months or older (N=16997). Age of youngest child has been rounded down to nearest month.

Table 2: Regression results: main variables

	Non-linear models						Linear models	
	Wooldridge dyn. prob.		Pooled prob.		IPW pooled prob.		Fixed Effects	Arellano-Bond
Variable	Coeff.	APE	Coeff.	APE	Coeff.	APE	Coeff.	Coeff.
D≥A	0.126**	0.026**	0.127***	0.031***	0.127***	0.031***	0.023**	0.013
	(0.057)	(0.012)	(0.033)	(0.008)	(0.034)	(0.008)	(0.010)	(0.010)
Has child[a]	−0.074	0.006	−0.174**	−0.023	−0.172**	−0.023	0.017	−0.003
	(0.102)	(0.020)	(0.071)	(0.016)	(0.074)	(0.017)	(0.017)	(0.018)
No. children	−0.267***	−0.054***	−0.095***	−0.023***	−0.105***	−0.026***	−0.053***	−0.073***
	(0.052)	(0.011)	(0.028)	(0.007)	(0.029)	(0.007)	(0.009)	(0.010)
Youngest <1yr	0.187**	0.038**	0.165***	0.041***	0.158***	0.039***	0.012	0.020
	(0.075)	(0.017)	(0.053)	(0.013)	(0.055)	(0.014)	(0.012)	(0.013)
Age youngest[b]	−0.023***	−0.046***	−0.117***	−0.029***	−0.118***	−0.029***	−0.058***	−0.048***
	(0.006)	(0.013)	(0.030)	(0.007)	(0.031)	(0.007)	(0.008)	(0.009)
Hinc*has child[c]	0.198***	0.026	0.155***	−0.005	0.153***	−0.005	0.063***	0.042***
	(0.075)	(0.022)	(0.050)	(0.018)	(0.051)	(0.018)	(0.014)	(0.015)
$PHI_{i,t-1}$	1.080***	0.254***	1.543***	0.491***	1.560***	0.496***		0.228***
	(0.048)	(0.013)	(0.028)	(0.009)	(0.029)	(0.009)		(0.017)
PHI_0	0.714***	0.159***	0.156***	0.040***	0.158***	0.040***		
	(0.063)	(0.009)	(0.029)	(0.007)	(0.030)	(0.007)		
N	19872		19872		19872		19872	13248

Note: Average partial effects (APEs) are the shift in predicted probabilities averaged across the sample when changing status from 0 to 1 for discrete variables, and the average of the vector of marginal effects for the sample for continuous variables. Standard errors in parenthesis. For coefficients, asymptotic standard errors are reported, clustered across individuals. Standard errors for APEs are calculated using non-parametric bootstrap with 200 replications for the pooled probit specifications and 100 replications for the Wooldridge dynamic probit. All models include controls in Table A2. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

[a] In the non-linear models, the APE takes into account the interaction between Has child and Hinc and therefore reflects the overall effect of presence of children.

[b] Row values have been multiplied by 1000.

[c] In the non-linear models, the APE is measuring the shift in predicted probability from having a child conditional on being rich both before and after having the child.

Appendix

Policy incentives

In 1996, PHI coverage was around 34% in Australia and on a downward trajectory. In response, three policy initiatives were introduced in the late 1990s to increase take-up. These were a tax penalty imposed on uninsured higher income earners (the Medicare Levy Surcharge), a 30% rebate for the cost of insurance and a mandate requiring insurers to increase premiums by 2% for every year uninsured after turning 31 (Lifetime Health Cover (LHC) loading). The impact of these reforms has been studied in Palangkaraya and Yong (2005) and Ellis and Savage (2008). Following these reforms, PHI coverage increased markedly from a historical low of 30% to 46% by September 2000. Coverage then decreased slightly to 43.6% in December 2006 but has since increased to around 45%.

We consider how to account for the policy incentives mentioned above in our specification of PHI demand. The rebate was not means tested during our study period and therefore does not require modelling adjustment. Studies suggest that LHC loading did have a significant effect on insurance participation in 2000 when it was introduced (Palangkaraya & Yong, 2005; Ellis & Savage, 2008). Our sample is generally too young for this policy to affect their incentives for insurance and only a subset of women reach the age of 30 during the analysis period. Nevertheless, we examined our coverage data for those women around the age of 31 and could see no clear discontinuity in the pattern of insurance at this age. Also, the inclusion of a year dummy will capture residual advertising effects. Controlling for the MLS would require more detailed income records than are available to us. We do not consider this as a serious limitation since recent research indicates that the surcharge has not significantly affected overall demand for PHI (Stavrunova & Yerokhin, 2014). Overall, we do not expect these policy incentives to bias our results.

Attrition

In this section, we provide additional details on sensitivity analysis to deal with potential non-random attrition. To address this concern, we use inverse probability weighting (IPW) as detailed in Wooldridge (2002).

We denote s_i as an indicator for whether a respondent who is present in the first wave ultimately becomes part of our analysis sample. We assume that, conditional on a set of observables in the first time period, $S_{i,1}$, the probability of participation in the sample is independent of other future observables such that

$$P(s_i = 1|X_{it}, PHI_{it}, PHI_{i,t-1}, PHI_{i,1}, S_{i,1}) = P(s_i = 1|S_{i,1}) \quad t = 2, 3, 4 \tag{A.1}$$

where the vector $S_{i,1}$ includes all variables measured in the initial period that explain selectivity into the analysis sample. This assumption effectively states that selection is only on observables. Importantly, $S_{i,1}$ can contain all variables in X as well as additional variables not used in the main equation.

IPW is implemented in two steps. First, a probit model is estimated for equation (A.1) using data from the first wave of the survey. Second, our data is weighted by the inverse predicted probability so that a greater weighting is placed on respondents with a higher probabil-

ity of attrition. Wooldridge (2002) shows that IPW produces consistent, $\sqrt{N}$-asymptotically normal estimators in models where the likelihood can be written as a sum of contributions across all observations. This is not the case for the random effects probit model and so we use a pooled probit specification. A pooled probit regression without weighting is presented for comparison. One shortcoming of our IPW procedure is that some variables in X_{it} are not observed in wave 1 and therefore cannot be included in $S_{i,1}$. We are however able to include additional wave 1 information that does not enter the main equation. Our complete set of controls used to create our probability weights include all the controls in Table A2 except for the high income indicator, health care card indicator, hospital access indicators and variables on the desire for and age of children. Additional controls include life satisfaction, a continuous measure for mental stress; variables created from the SF-36 survey questions including number of activities has limited ability to do, in past four weeks whether physical/mental health interfered with work, how much body pain was experienced, whether pain interfered with work, health perception relative to others and expected health; how well gets along with others; whether friends understand you; educational aspirations and whether wants children at some point.

References

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.

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.

Stavrunova, O., & Yerokhin, O. (2014). Tax incentives and the demand for private health insurance. *Journal of Health Economics*, *34*, 121–130.

Wooldridge, J. M. (2002). Inverse probability weighted M-estimators for sample selection, attrition, and stratification. *Portuguese Economic Journal*, *1*, 117–139.

Table A1: Transitions in desire for additional children ($1 = D > A$)

Group	Wave 2	Wave 3	Wave 4	Proportion (%)
1	1	1	1	59.59
2	1	1	0	15.41
3	1	0	1	2.91
4	1	0	0	8.97
5	0	0	0	7.40
6	0	1	1	2.68
7	0	0	1	1.39
8	0	1	0	1.48

Note: Sample size for each wave is 6624 women.

Table A2: Variable definitions

Variable	Definition
Family formation	
$D \geq A$	Desire more children than have currently
Has child	Has at least one child
Number child	Total number of children
Time child - 1yr	Had a child less than one year ago
Time child	Days since youngest child turned 1 year
Hinc*has child	High income and has child
$D \geq A$ miss	$D \geq A$ missing
Financial resources	
Hinc	Average gross household income > $1000 per week
Inc alone	Household income is also personal income
Inc miss	Income measure missing
M_stress extreme	Over last 12 months, extremely stressed about money
M_stress very	Over last 12 months, very stressed about money
M_stress mod	Over last 12 months, moderately stressed about money
M_stress some	Over last 12 months, somewhat stressed about money
M_stress none*	Over last 12 months, not stressed about money
M_stress miss	Money stress missing
IER	Index of Economic Resources score for area
IER miss	IER missing
Employment status	
Study	Student not in work
Work	Worker not in study
Work study	Works and studies
No work study*	Neither works nor studies
Work miss	Work/study status missing
Qualifications	
Tertiary	Highest qualification is degree or higher
Diploma	Highest qualification is diploma or certificate
Trade	Highest qualification is a trade or apprenticeship
Only school*	Highest qualification HSC, school certificate or none
Qual miss	Highest qualification missing
Relationship status	
Married	Married
Defacto	In defacto relationship
Other*	Single, divorced, widowed or separated
Rel miss	Relationship status missing
Health and risk	
SF36 general	SF36 general health score
Health miss	SF36 general health score missing
Smoke	Smokes daily

Smoke miss	Smoker status missing
Alc risk	Risky or high risk drinker according to NHMRC guidelines
Alc miss	Alcohol intake missing
BMI	Body mass index (height/weight^2)
BMI miss	BMI missing
Drug use	Ever used a prohibited drug other than marijuana (Wave 4)
Drug miss	Drug use missing
Region	
Urban*	Urban centre population $\geq$ 100000
Rural	Urban centre population between 10000-99999
Remote	Urban centre population $<$ 10000
Area miss	Remoteness classification missing
State	Full set of dummies for Australia's 8 states and territories
Access	
Access 1	Perceived hospital access is excellent or very good
Access 2	Perceived hospital access is good
Access 3*	Perceived hospital access is fair or poor
Access miss	Perceived hospital access is missing
Other	
Age	Age in years
Health card	Holds a government Health Care Card
Card miss	Health Care Card status missing
Aus	Country of birth is Australia
COB miss	Country of birth is missing

Note: * indicates a control group.

Table A3: Variable means by wave

Variable	Wave 1	Wave 2	Wave 3	Wave 4
Family formation				
$D \geq A$		0.87	0.79	0.67
Has child	0.06	0.17	0.31	0.49
Number child	0.08	0.25	0.51	0.90
Time child - 1yr	0.03	0.06	0.10	0.13
Time child	28.44	115.98	244.69	474.25
Hinc*has child		0.03	0.11	0.29
$D \geq A$ miss		0.01	0.01	0.02
Financial resources				
Hinc		0.39	0.48	0.58
Inc alone		0.05	0.06	0.08
Inc miss		0.22	0.19	0.15
M_stress extreme	0.03	0.08	0.08	0.08
M_stress very	0.13	0.15	0.14	0.13
M_stress mod	0.31	0.23	0.24	0.22
M_stress some	0.38	0.36	0.38	0.39
M_stress miss	<0.01	0.01	<0.01	<0.01
IER	1004.17	999.04	995.70	983.81
IER miss	<0.01	<0.01	<0.01	0.03
Employment status				
Study	0.34	0.04	0.04	0.03
Work	0.40	0.61	0.56	0.61
Work study	0.14	0.25	0.23	0.16
Work miss	0.02	<0.01	0.02	0.03
Qualifications				
Tertiary	0.14	0.42	0.46	0.48
Diploma	0.15	0.21	0.22	0.24
Trade	0.02	0.03	0.03	0.03
Qual miss	<0.01	0.03	0.02	<0.01
Relationship status				
Married	0.09	0.26	0.43	0.55
Defacto	0.12	0.21	0.20	0.18
Rel miss	<0.01	<0.01	<0.01	<0.01
Health and risk				
SF36 general	69.79	70.71	72.49	73.70
Health miss	<0.01	<0.01	<0.01	<0.01
Smoke	0.22	0.17	0.15	0.13
Smoke miss	<0.01	<0.01	<0.01	<0.01
Alc risk	0.05	0.03	0.03	0.03
Alc miss	0.01	<0.01	<0.01	<0.01
BMI	20.49	21.92	21.95	24.54

BMI miss	0.10	0.08	0.11	0.03
Drug use	0.26	0.26	0.26	0.26
Drug miss	0.01	0.01	0.01	0.01
Region				
Rural	0.41	0.42	0.38	0.36
Remote	0.04	0.04	0.04	0.04
Area miss	<0.01	<0.01	<0.01	0.01
Access				
Access 1		0.51	0.55	0.61
Access 2		0.29	0.28	0.24
Access miss		0.08	0.06	0.05
Other				
Age	20.83	24.62	27.59	30.62
Health card		0.19	0.15	0.14
Card miss		0.03	<0.01	<0.01
Aus	0.93	0.93	0.93	0.93
COB miss	<0.01	<0.01	<0.01	<0.01

Note: Sample size for each wave is 6624 women.

Table A4: Coefficient estimates: controls

Variable	Wooldridge	Pooled Probit	Fixed-effects	Arellano-Bond
Financial resources				
Hinc	0.124***	0.257***	−0.003	0.001
Inc alone	−0.045	0.145***	−0.023*	−0.028*
M_stress extreme	−0.240***	−0.259***	−0.037**	−0.049***
M_stress very	−0.162**	−0.197***	−0.032***	−0.026*
M_stress mod	−0.090	−0.125***	−0.023**	−0.014
M_stress some	−0.023	−0.090***	−0.007	−0.003
IER	0.000	0.001***	0.000	0.000
Employment status				
Work	0.011	0.082**	−0.001	0.001
Work study	0.085	0.133***	0.010	0.018
Qualifications				
Tertiary	−0.329***	0.088***	−0.064***	−0.062***
Diploma	−0.129*	−0.021	−0.032**	−0.032**
Relationship status				
Married	0.369***	0.349***	0.076***	0.088***
Health and risk				
Smokes	0.028	−0.170***	−0.004	0.005
BMI	−0.005	−0.006***	−0.001	0.000
Drug use	−0.206***	−0.157***		
Region				
Rural	−0.285***	−0.229***	−0.049***	−0.059***
Access				
Access 1	0.370***	0.356***	0.068***	0.073***
Access 2	0.164***	0.136***	0.027**	0.027**
Other				
Age	0.129*	0.079***	0.014	0.033**
Health card	0.118***	−0.094***	0.027***	0.028**
Year 2000	−0.345	−0.151***	−0.160**	−0.038
Year 2003	0.014	0.136***	−0.066	0.012
N	19872	19872	19872	13248

Note: Only variables that are statistically significant in at least one regression are shown (state and territory and controls for missing values exclusive). Statistical significance for coefficient estimates is based on asymptotic standard errors. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A5: Regression results no first wave: main variables

	Non-linear models						Linear models	
	Wooldridge dyn. prob.[a]		Pooled prob.		IPW pooled prob.		Fixed Effects	Arellano-Bond
Variable	Coeff.	APE	Coeff.	APE	Coeff.	APE	Coeff.	Coeff.
D≥A	0.094	0.019	0.162***	0.036***	0.152***	0.035***	0.020*	0.013
	(0.078)	(0.014)	(0.039)	(0.009)	(0.040)	(0.009)	(0.012)	(0.010)
Has child[b]	−0.124	0.008	−0.237***	−0.041**	−0.210**	−0.036**	0.013	−0.003
	(0.139)	(0.023)	(0.085)	(0.017)	(0.089)	(0.018)	(0.021)	(0.018)
No. children	−0.447***	−0.088***	−0.071**	−0.016**	−0.086***	−0.019**	−0.075***	−0.073***
	(0.081)	(0.016)	(0.031)	(0.007)	(0.033)	(0.007)	(0.012)	(0.010)
Youngest <1yr[c]	0.144	0.029	0.212***	0.048***	0.203***	0.047***	0.011	0.020
	(0.094)	(0.018)	(0.061)	(0.014)	(0.063)	(0.014)	(0.013)	(0.013)
Age youngest	−0.228***	−0.045***	−0.079**	−0.018**	−0.087***	−0.020***	−0.056***	−0.048***
	(0.085)	(0.015)	(0.032)	(0.007)	(0.033)	(0.007)	(0.011)	(0.009)
Hinc*has child[d]	0.303***	0.035	0.080	−0.035*	0.082	−0.029	0.050***	0.042***
	(0.110)	(0.022)	(0.062)	(0.019)	(0.063)	(0.020)	(0.018)	(0.015)
$PHI_{i,t-1}$	1.500***	0.380***	1.790***	0.544***	1.804***	0.545***		0.228***
	(0.092)	(0.017)	(0.033)	(0.009)	(0.034)	(0.009)		(0.017)
PHI_0	0.704***	0.159***	0.361***	0.087***	0.362***	0.086***		
	(0.156)	(0.029)	(0.049)	(0.008)	(0.033)	(0.008)		
N	13248		13248		13248		13248	13248

Note: Average partial effects (APEs) are the shift in predicted probabilities averaged across the sample when changing status from 0 to 1 for discrete variables, and the average of the vector of marginal effects for the sample for continuous variables. Standard errors in parenthesis. For coefficients, asymptotic standard errors are reported, clustered across individuals. Standard errors for APEs are calculated using non-parametric bootstrap with 200 replications for the pooled probit specifications and 100 replications for the Wooldridge dynamic probit. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

[a] Due to convergence issues, some variables with frequencies less than 100 observations were omitted from the estimation. These were indicators for missing region, marital status, smoking status, alcohol use, health card, country of birth, money stress and SF36 health score. For the health score missing observations were given the sample average and for all other variables were allocated to the reference group.

[b] In the non-linear models, the APE takes into account the interaction between Has child and Hinc and therefore reflects the overall effect of presence of children.

[c] Row values have been multiplied by 1000.

[d] In the non-linear models, the APE is measuring the shift in predicted probability from having a child conditional on being rich both before and after having the child.