Lund University
Department of Economics
NEKP01
Master Thesis 2

Can the Hilda survey offer additional insight on the impact of the Australian lifetime health cover policy?

A regression discontinuity approach using panel data

Author: Karl-Olof Welin
Supervisor: Ulf Gerdtham
Seminar: 5 September 2013

Abstract

The focus of this paper is on the partial replication and reevaluation of the work of Palangkaraya and Yong (2007) regarding the impact of the lifetime health cover policy in Australia. The analysis is performed using data on single person households obtained from the Hilda survey. The method used is a regression discontinuity design applied to a panel data material with age as the forcing variable. The analysis consist of two parts, the first is a strict replication of the linear probability estimate of Palangkaraya and Yong (2007). In the second part a regression discontinuity approach more suitable for the data material is utilized. The general result is inconclusive, the strict replication of the linear probability model indicate an insignificant treatment effect. The regression discontinuity analysis show varying levels of significance depending on the bandwidth used. As indicated by the contradictory estimates, it cannot be ruled out that the regression discontinuity design is invalid and therefore offer no additional insight on the effect of the lifetime health cover policy.

Keywords: Regression discontinuity design, lifetime health cover, panel data, Hilda survey

A note on the Hilda survey

This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the author and should not be attributed to either FaHCSIA or the Melbourne Institute.

Table of content

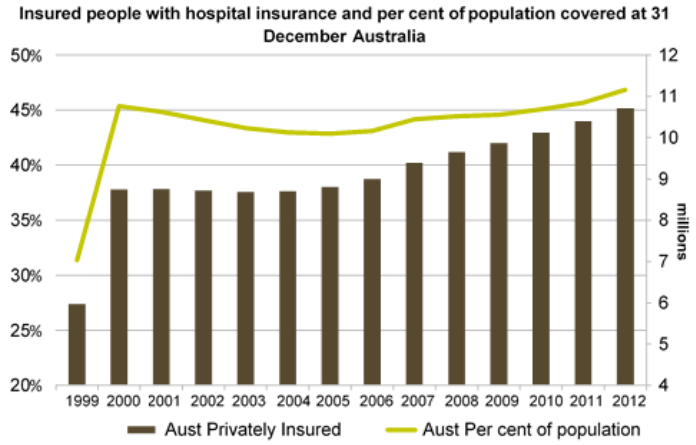
1. Introduction

1.1 Background: the Australian Health Care System, Medicare and Policy Changes

The Australian health care system offers "universal health care coverage for citizens and permanent residents" (Thomson, Osborn, Squires & Jun, 2012, p.11). This is done through the public health insurance system called Medicare. According to Thomson et al (2012) this can be generalized as most medical services being either free of charge or publicly subsidized. Note that general in-hospital care is always provided free of charge but private health insurance (PHI) can be utilized to obtain more options in terms of where, when and how care is provided. PHI is not a requirement for care but is encouraged by the government and public health care sector.

During the period 1997 to 2000 three major incentives for individuals to obtain PHI were introduced by the Australian government. The three are described by Palangkaraya and Yong (2007) as an additional fee for high income earners without insurance, a thirty percent rebate on the premium of PHI and finally the lifetime health cover policy. The addition fee for high income earners is one percent on taxable income for individuals and families with incomes greater than AUD$80´000 and AUD$160´000[1] respectively if not covered by PHI (Thomson et al, 2012, p.12).

Figure 1: PHI coverage at age thirty-one



Obtained from the Private Health Insurance Administration Council (2013).

The lifetime health cover (LHC) policy increases the price of PHI by 2% for every year an individual postpones obtaining PHI, starting the 1st of July after turning thirty-one. The price increase accumulates up to 70%, if PHI is obtained that price is maintained for ten years of continuous insurance, after which it is removed as long as the individual retains the insurance (Private Health Insurance Ombudsman).

[1] These are the values for 2012, the values have been revised since implementation.

The proportion with PHI at age thirty-one from 1999 to 2012 is displayed in figure 1. According to Thomson et al (2012) 54.4% of the Australian population had general coverage while 46.8% had the hospital cover that is required by the LHC.

Following the implementation several attempts have been made to estimate the actual effect of the LHC on the acquisition of PHI[2]. The perhaps most notable contribution is that of Palangkaraya and Yong (2007), where a regression discontinuity design is utilized. The main argument and focus is that the effect of the three policies ought to be separated when analysis is performed, in order to establish which policy causes the respective change in PHI. The work of Palangkaraya and Yong (2007) will be extended upon in section 2.1.

[2] See Palangkaraya and Yong (2007) for some examples.

1.2 Description of the Problem

There are several issues in trying to determine the effect of the LHC, as stated by Palangkaraya and Yong (2007) the implementation of policies with similar goals make it hard to separate the effect of the LHC from the other policies. The method used by Palangkaraya and Yong (2007) is a prime example of making the most of the available information and combining several techniques in order to make a convincing argument. There is no doubt regarding the value of the work but as far as the regression discontinuity (RD) design and data goes there is room for improvement. The broad age groups and using groups as far apart from the threshold renders a fairly weak RD design due to the possibly large difference in various characteristics that influence the insurance decision.

In light of these issues the Hilda survey offers the option to attempt a similar analysis using a different set of data. The new data has the potential to improve on the weakness of the design of Palangkaraya and Yong (2007) in terms of much more elaborate data on income and age of individuals while offering a variety of additional covariates. The work here will consist of trying to replicate the general result of Palangkaraya and Yong (2007) and to reevaluate the result using a design more appropriate for the current data-set.

1.3 Purpose

The purpose of this study is primarily to replicate and reevaluate the result of Palangkaraya and Yong (2007). This study can contribute in two ways; by testing the previous result by performing a similar analysis using a different data set and by further investigating whether the effect of the LHC can be estimated using a specification appropriate for the Hilda dataset.

The contributions are important in terms of further evaluation of the LHC policy and in working towards further development of RD design by applying recent suggestions to real world data.

2. Theoretical Framework and Previous Research

2.1 Previous Research

In response to the PHI incentives introduced by the Australian government Palangkaraya and Yong (2007) performs an analysis using a RD design. The main goal being to determine the effect of the lifetime health cover (LHC) and to separate it from the other PHI inducing policies. Palangkaraya and Yong (2007) utilizes the National Health Survey (NHS) data from 2001 and 1995. This allows for analysis both around the threshold age of turning thirty-one after the LHC is implemented and a sort of difference-in-difference estimate comparing the PHI data in the 2001 NHS sample to that in the 1995 sample.

There are a few restrictions and assumptions in effect for the analysis performed by Palangkaraya and Yong (2007), some of these are important to highlight when trying to replicate part of the analysis. The first restriction can be regarded as fairly severe in terms of a pure RD approach, namely that the NHS data is divided into age groups with five year intervals. This "coarse classification of age groups means that a significant amount of heterogeneity is buried within each age group" (Palangkaraya & Yong, 2007, p.1365) meaning there likely exist meaningful differences between the individuals in the same age group. It is unfortunate due to the threshold being at age 31, which renders the 30-34 group invalid since it contains both treated and untreated individuals. The actual discontinuity is formulated in a rather rough manner with the 25-29 group being compared to the one with 35-39 years of age. There is a significant risk that individuals in the two groups are quite different. Due to restrictions in the dataset regarding number of individuals in each household and the distribution of adults and children, Palangkaraya and Yong (2007) have opted to only use single individuals, this is further motivated by "insurance decisions of families are also potentially more complex" (Palangkaraya & Yong, 2007, p.1364).

A direct connection to the coarse age groups can be found in the fact that income generally tend to increase with age, which possibly has a significant effect on the PHI decision. As a response to this Palangkaraya and Yong (2007) has divided their sample into a high and low income group and the analysis is performed for each group. In addition to the general analysis, Palangkaraya and Yong (2007) argue that individuals with greater risk to their health react differently to the LHC. The analysis is thus also performed on groups with worse than average health, defined as having three or more chronic conditions.

The findings of Palangkaraya and Yong (2007) can be summarized as the LHC having a positive effect on the proportion of individuals who acquire PHI. Income is show to have a significant positive effect on the acquisition of PHI and having three or more chronic conditions increases the probability of possessing PHI. It is also shown that the effect of LHC appear to mainly affect individuals with a higher income.

2.2 Theoretical Framework

There are a number of recent publications with focus on RD design, a distinct group of these stand out for their suggestions and advances in the field. Among these a number with heavy focus on empirical research have emerged which act as guidelines for applied work. The most extensive work in this category is that of Lee and Lemieux (2010), which here will be supplemented by the work of Van der Klaauw (2008) as well as Imbens and Lemieux (2008). Lee and Card (2008) is utilized for the suggestions focused on RD designs with specification error.

The general methodology dealt with in this analysis is the regression discontinuity design. A RD design can be described as a quasi-experimental design where a specific effect is applied conditional on meeting a certain criteria. The criteria commonly follows from policy or administrative decisions guided by a clearly formulated assignment rule (Imbens and Lemieux, 2008, p.616). What this translates to in practice is individuals who reach a certain level in some variable, commonly called the assignment or forcing variable, are eligible for a certain program or rule. Being exposed to said program or rule is known as being treated, which is why any effect of the change is known as the treatment effect. The treatment status d_i can formally be described as:

$$d_i = \begin{cases} 1 & if\ a_i \geq \bar{a} \\ 0 & otherwise \end{cases} \quad (1)$$

Where $\bar{a}$ is the threshold value of the assignment variable which enables treatment. Let $Y_i(d_i)$ be the outcome for individual i, the treatment effect can thus be written as:

$$\tau_i = Y_i(1) - Y_i(0) \quad (2)$$

Which is the treatment effect for individual i, calculated as the difference in outcome when treated and when untreated. The issue in using this methodology according to Van der Klaauw (2008) is the impossibility of observing an individual as both treated and untreated due to the two states being mutually exclusive, an individual can only exist in one at the time. In order to estimate the treatment effect Van der Klaauw (2008) explains that local averages are used based on observations in close proximity on both sides of the threshold. This emulates randomization of the treatment variable as ”if it is reasonable to assume that persons close to the threshold with very similar x values are comparable, then we may view the design as almost experimental near $\bar{x}$“ (Van der Klaauw, 2008, p.224). Similar arguments can be found in Imbens and Lemieux (2008) and Lee and Lemieux (2010). The generalized average treatment effect using a sample within ϵ on either side of the threshold can be expressed as:

$$\tau = E[Y|\bar{a} + \epsilon] - E[Y|\bar{a} - \epsilon] \quad (3)$$

Equation (3) is based on the work of Imbens and Lemieux (2008) but modified to fit this context.

The analysis, however, is more complicated than the model above would suggest. Restricting ϵ to a very small value provides the best model in terms of how similar individuals are, and the virtual randomization is therefore strong. As pointed out by Imbens and Lemieux (2008) however, a very large sample is required for this to fully function. This poses a practical threat to the RD method as even when utilizing very large sets of data the number of observations in the area around the threshold is often low. In practice there are thus two different versions of RD design. The first is the non-parametric version which is described above, due to the strong randomization in the area around the threshold additional covariates can basically be ignored and the causal effect estimated by local linear regression or a similar method at the boundary (Van der Klaauw, 2008). The second method is called a parametric approach and is based on a broader range of ϵ, which implies observations further from the threshold are used in order to utilize additional information. According to Van der Klaauw (2008) including more distant observations mean a higher precision is attained at the cost of possible misspecification and bias due to individuals being increasingly different as the bandwidth increases. A fundamental difference is that for a parametric approach functional form has to be considered in order to account for the observations further away from the threshold. For parametric approaches additional covariates are included to control for differences between individuals. Defining an appropriate functional form is thus crucial to correctly specifying the design and obtaining unbiased estimates.

An additional distinction has to be made regarding whether a sharp or fuzzy design is appropriate. The difference between the two is that in a sharp design crossing the threshold means an individual is treated, treatment is a deterministic based on assignment. In the fuzzy case passing the threshold only increases the probability of receiving treatment, treatment is not certain just because of being assigned treatment (Van der Klaauw, 2008)

The work of Lee and Lemieux (2010) provide additional insight on three specific areas relevant for this analysis. The first section is concerned with discrete assignment variables. Lee and Lemieux (2010) point out that the theoretical appeal of the RD design is based on a continuous assignment variable while many practical applications utilize a discrete assignment variable. A discrete variable would violate the idea of individuals infinitely close to the discontinuity. Following this violation “one must use regressions to estimate the conditional expectation of the outcome variable at the cutoff point by extrapolation. /…/ however, in practice we always extrapolate to some extent, even in the case of a

continuous assignment variable. (Lee and Lemieux, 2010, p.336). In addition to the extrapolation being performed for continuous variables, it is also clear that "the fact we must do so in the case of a discrete assignment variable does not introduce particular complications from an econometric point of view, provided the discrete variable is not too coarsely distributed" (Lee and Lemieux, 2010, p.336)

The second section of importance is regarding RD design and analyzing panel data. According to Lee and Lemieux (2010) it is tempting to estimate a model with elements from fixed effects models, "note, however, that including fixed effects is unnecessary for identification in an RD design" (Lee & Lemieux, 2010, p.337). The authors even go as far as arguing that "imposing a specific dynamic structure introduces more restrictions without any gain in identification" (Lee and Lemieux, 2010, p.337). Instead it is suggested to simply pool the panel material and treat it as a cross-sectional dataset while "taking care to account for the within-individual correlation of the errors over time using clustered standard errors" (Lee and Lemieux, 2010, p.337). The method utilizing the clustered standard errors was developed by Lee and Card (2008), who state that using a discrete variable tends to overestimate the treatment effect due to the impossibility of minimizing the bandwidth on either side of the threshold. It is noted that "`clustered´ standard errors will generally lead to wider confidence intervals that reflect the imperfect fit of the parametric function away from the discontinuity" (Lee & Card, 2008, p.656). It appear that clustered standard errors allow for a more realistic estimate, which for the current context is in line with Lee and Lemieux (2010). Note that the method proposed by Lee and Card (2008) concerning the use of clustered standard errors has a similar reasoning to the method proposed by Lee and Lemieux (2010) but is in fact different. The method by Lee and Lemieux (2010) is more suitable for this context due to the panel nature of the data.

The third and final section of relevance is regarding specific issues in using age as the assignment variable. Lee and Lemieux (2010) argue that utilizing age as the forcing variable can be an issue due to the inevitability of treatment. First of all, it is argued that effects of age-based policies might not take effect immediately and will thus not prompt a discontinuity. In addition, there is a second effect which potentially is more relevant for this context, due to the nature of age treatment is perfectly predictable and individuals might adapt prior to actually being treated, thus making the effect unobservable around the discontinuity (Lee & Lemieux, 2010). No general recommendations are offered other than that the context has to dictate how difficulties like this should be regarded and that analysis has to be performed with some caution.

2.3 Model

According to the propositions of Lee and Lemieux (2010) the following model will be used. First of all, due to treatment being certain after age thirty-one this is a sharp design where the treatment status is determined by:

$$d_i = \begin{cases} 1 & if\ a_i \geq 31 \\ 0 & otherwise \end{cases} \quad (4)$$

d_i indicates the treatment status and a_i is the age of individual *i* which act as the assignment variable. Due to the discrete nature of the forcing variable a parametric RD design will be used. The general model in turn is defined by function (5) below.

$$E[y_i] = d_i\tau + \beta X\ + \epsilon_i \quad (5)$$

Where y_i is the variable indicating if individual *i* has private health insurance while X is a vector of covariates, τ is the treatment effect and ϵ_i is the residual. The standard errors are clustered around the variable indicating unique individuals in order to allow for within individual correlation over time due to the repeat sampling of individuals in the panel data, in accordance with the suggestions by Lee and Lemieux (2010).

When it comes to estimation of the model two different methods will be employed, a linear probability model similar to that of Palangkaraya and Yong (2007) and a probit regression model to verify the results and to more efficiently utilize the information in the binary dependent variable.

3. Data

3.1 The Hilda Survey

Data is obtained from the Household, Income and Labour Dynamics in Australia Survey (Hilda). The Hilda material is based on a longitudinal survey with a new wave coming out every year, as of 2013 Hilda include eleven waves (Summerfield, Freidin, Hahn, Ittak, Li, Macalalad, Watson, Wilkins & Wooden, 2012). The initial wave covered 19 914 individuals across 7 682 households and while the initial sample is maintained it is allowed to evolve by deaths, births, migration and household connections (Melbourne Institute). Another 5 477 individuals spread over 2 153 households were added for wave eleven. This results in a rather extensive data-set with over 200 000 individual observations over eleven years. The data set has a rather complex setup with multiple individuals in each household and individual relationships coded into the material. Each wave consist of a few thousand variables, it should be noted that a substantial part of these are technical in nature in order to allow for processing of data and simply contain structural and technical descriptions rather than direct information on each individual. For this analysis waves four through eleven are utilized due to limited information on PHI in waves one through three.

3.2 Sample

The sample will largely be constructed in a similar way to that of Palangkaraya and Yong (2007) due to trying to replicate their OLS result, but also due to the research they offer in terms of which variables that affect PHI decisions. The initial idea was to include a broader sample in terms of not restricting the analysis to single person households. However, the comments and work of Palangkaraya and Yong (2007) proved true, the insurance decision of multi-person households, be it couples or families, is much more complicated than that of singles. The number of unobservable characteristics increase and issues such as the partner of an individual crossing the threshold quickly become a significant limitation to the RD design. In addition to the problem of modeling more complicated insurance decisions there is a more practical limitation in terms of data. The Hilda material contains data on PHI expenditure on the household level rather than the individual level. The PHI expenditure data is used as a proxy for having PHI, for individuals in multi-person households it is impossible to determine whether the individual in question is covered by the insurance or not. A scenario where a household has positive expenditure but a specific individual being without cover is thus possible. The second restriction on the sample is concerning which age to include in the analysis. I have decided to limit the age to the span of twenty to sixty-five. This is motivated by it being the general working age in Australia, and also the age when a significant proportion of younger individuals have started living on their own and thus are not covered to the same extent of any PHI of their parents. The Hilda material only contains data

on PHI expenditure for waves four through eleven, the first three waves are thus excluded from the analysis.

The sample is thus defined as individuals in single person households between twenty and sixty-five years of age in waves four through eleven.

3.3 Variables of Interest

Age is defined in two different ways in the Hilda survey. The version that will be used here is defined as "the age at last birthday as of 30 June immediately preceding the fieldwork for that wave" (Summerfield et al, 2012, p.31). Note that using the alternative definition of age potentially affects the result for this kind of study since age acts as the forcing variable for treatment. The reason for using the above definition can be found in how the LHC is implemented. The effect of the LHC starts on 1 July after an individual turns thirty-one (Private Health Insurance Ombudsman), it is therefore safe to assume that the age of an individual on 30 June is the same as on 1 July. This way any potential loss of information regarding date of birth in relation to response date and whether the individual in fact is treated is avoided, all individuals with an age of thirty-one or more can be considered treated.

When considering income there are several different options that can be used, each with its own set of advantages and disadvantages. The version that is utilized for this analysis is disposable income, defined as "total income after receipt of government benefits and deduction of income tax" (Summerfield et al, 2012, p.56). The reason for using disposable income is the belief that of the income variables available disposable income ought to be the one which best corresponds to the income variable used by Palangkaraya and Yong (2007).

To determine whether an individual is covered by private health care insurance the variable regarding expenditure on PHI is used. This variable is treated as a proxy for having PHI and is transformed into a binary variable which indicates an individual as having PHI if the expenditure on PHI is more than zero. The fourth wave of Hilda is focused on health insurance and therefore has a specific variable for PHI which is utilized for that year.

Two variables regarding the individual's health condition are included, one which indicates if the individual suffers from some chronic condition and the other one is a measure of self-assessed health. The variable regarding self-assessed health is defined as a reply to "As healthy as anybody I know" and is answered on a scale of "definitely true, mostly true, don't know, mostly false and definitely false".

This is transformed into a binary variable defined as the individual regarding their health as bad if "mostly false" or "definitely false" is provided as answer. In addition to the above a couple of additional covariates are included; gender and whether an individual smokes.

3.4 Variables for Replication

A replication of the OLS results of Palangkaraya and Yong (2007) will be attempted in order to determine if the current data set indicate similar results when their basic linear probability procedure is followed as closely as possible. While an identical replication would be ideal it is unfortunately not quite possible due to certain variables being defined differently in the two sets of data. The variables that differ greatly are considered in this section and their respective definitions are compared.

Income is defined as taxable income and is adjusted for the common deductible of AUD$6000, additionally income is reported in intervals of AUD$5000 rather than specific amounts (Palangkaraya & Yong, 2007, p.1371). The income variable opted for the replication is disposable income, meaning it is adjusted for deductibles and benefits. Gross income is also available but due to Palangkaraya and Yong (2007) having adjusted the gross income for the common deductible, disposable income is used here. Also note that exact income figures are available in the Hilda survey as opposed to the NHS used by Palangkaraya and Yong (2007).

The variable indicating chronic conditions differ slightly between the two studies, Palangkaraya and Yong (2007) uses the number of chronic conditions from zero to five while the current analysis uses a binary variable indicating if an individual suffers from any chronic condition. The difference is due to limitations in the data, the Hilda survey simply does not contain information on the specific number of chronic conditions.

Regarding the dummy variable indicating if self-assessed health is worse than average, it is unclear from the description exactly how it is defined other than it is indeed self-assessed health. The specification utilized here is done in accordance with the description in section 3.3. Sex, smoking and the dummy variable indicating the two age groups, i.e. the treatment effect, are similar enough to not warrant any further attention.

4. Analysis

The analysis is divided into three parts, in the first part the validity of the RD design for the current data is investigated using various graphical tests. The second part is an attempted replication of the OLS result of Palangkaraya and Yong (2007) to determine if the effect of LHC can be shown using their methodology on the current data. The final part consists of performing the RD analysis using the current material.

4.1 Graphical Validation

The first and most basic test of validity is to plot the forcing variable against the outcome in order to visually illustrate the discontinuity as suggested by Van der Klaauw (2008) as well as Imbens and Lemieux (2008) amongst others. This also provides a general overview of the correlation between the outcome and forcing variable and indicate fluctuations and volatility (Lee & Lemieux, 2010).

Figure 2: Proportion with PHI by age

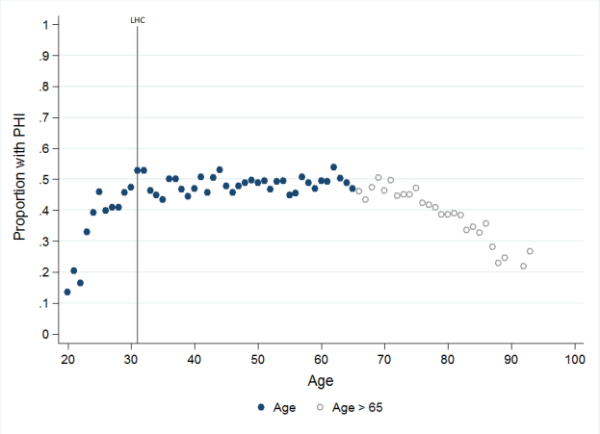
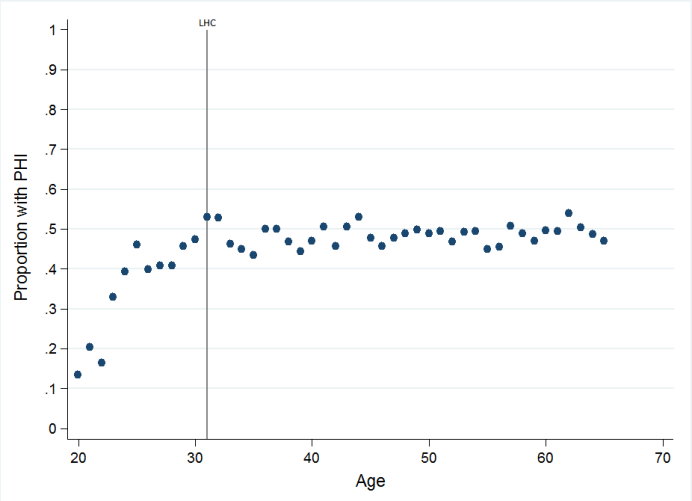


Figure 2 provides an overview of the proportion of individuals at each age with PHI. It is difficult to draw any conclusions based on the figure 2, a few things worth mentioning is that the proportion with PHI appear to be a fairly consistent over the various ages, and a continuous relationship between age and PHI is no impossibility, further note the decline in proportion with PHI after the legal retirement age of 65.

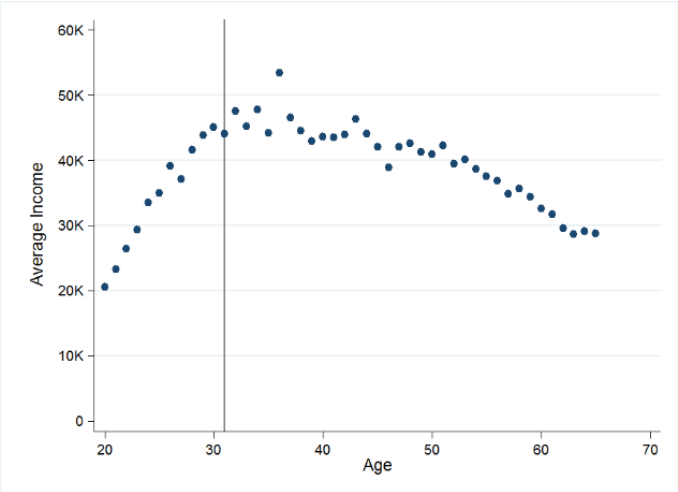
Figure 3 describes the same relationship as figure 2, the only difference being that figure 3 is limited to the sample age of twenty to sixty-five, and therefore the sample that is actually utilized in the analysis. The most important thing to note is the ambiguity regarding a possible discontinuity at the threshold age of thirty-one, indicated by the line LHC in the figure.

Figure 3: PHI by forcing variable



Evaluating covariates by forcing variable is done in order to investigate whether any discontinuities around the threshold exist in the covariates. If this is the case the design might prove invalid since the treatment effect could be caused by the discontinuity in the covariate rather than the actual treatment Lee and Lemieux (2010). Income is used as the covariate here as it is the only continuous variable apart from the assignment variable.

Figure 4: Income by forcing variable

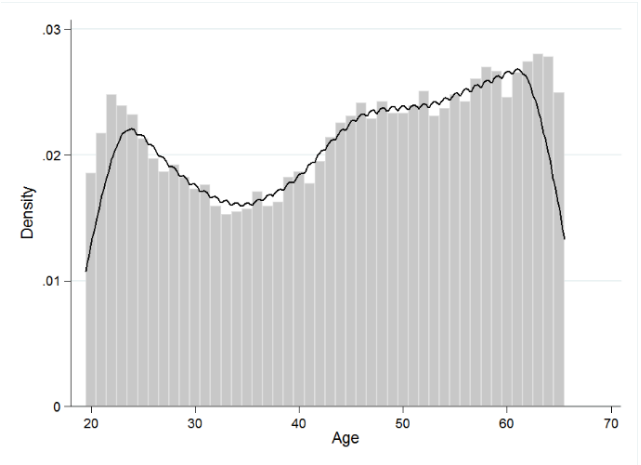


As can be seen in figure 4, no apparent discontinuity can be seen in average disposable income, which supports the validity of the design. The relationship also vaguely resembles the one describing PHI, suggesting a potential relationship between the two.

The density of the forcing variable is examined as it indicates if assignment has been manipulated. Manipulation is indicated by the density close to the threshold being significantly different on either side of the cutoff. While the work of Lee and Lemieux (2010) implies this is unlikely to be a problem

with age as the forcing variable it is included for completeness. Looking at figure 5, there is no discontinuity or jump in the density of age on either side of turning thirty-one. The age variable does not seem to have been manipulated, this scenario was expected due to the inevitability of age, but reassuring for the validity of the design nonetheless.

Figure 5: Density of Age



4.2 Results

The first result presented is the replication of the linear probability model by Palangkaraya and Yong (2007). The left part of table 1 is the estimate of Palangkaraya and Yong (2007) while the right side contains the estimate using the Hilda sample.

Table 1: Replication of the linear probability model of Palangkaraya and Yong 2007

	2001 NHS			**Hilda-survey**		
Dependent Variable: PHI	Estimate	SE	t-value	Estimate	SE	t-value
Treatment: 1 = LHC	0.152	0.0254	6.01	0.0269	0.03549	0.76
Sex: 1 = Female	0.031	0.0254	1.22	0.1422 *	0.03634	3.91
Income	0.060	0.0061	9.68	5.43e-06 *	1.01e-06	5.39
Chronic condition	0.017	0.0080	2.17	-0.0112	0.04304	-0.26
Self-assessed health, 1 = bad	0.137	0.0375	3.65	-0.0006	0.04597	-0.01
Smoker: 1 = Yes	0.134	0.0258	5.20	-0.2066 *	0.03294	-6.27
Constant	0.077	0.0358	2.16	0.2232 *	0.05547	4.02
Observations	1150			1584 (830 clusters)		

Note that no significance levels are provided for the result of Palangkaraya and Yong (2007)
Indications of significance: * 1%, ** 5%, *** 10%.

As can be seen in table 1 the results are rather different, the only similarity is found in the effect of income where both magnitude and sign are close while remaining significant. The most important thing to note is the treatment effect not being significant when using the Hilda material.

In order to further investigate the effect of the LHC policy using a specification more appropriate than the replication above, lead to the following results. Four estimates are provided here, using four different bandwidths around the threshold. The first estimate is presented in table 2 and includes the entire sample.

Table 2: The entire sample

	OLS			**Probit**		
Dependent Variable: PHI	Estimate	SE	t-value	Estimate	SE	z-value
Treatment: 1 = LHC	0.1179 *	0.1902	6.20	0.3277 *	0.0570	5.75
Sex: 1 = Female	0.1362 *	0.0187	7.28	0.4223 *	0.0572	7.38
Income	6.46e-06 *	5.34e-07	12.10	2.49e-05 *	1.45e-06	17.22
Chronic condition	-0.0613 *	0.0190	-3.23	-0.1272 **	0.0568	-2.24
Self-assessed health, 1 = bad	-0.0101	0.0192	-0.53	-0.0172	0.0605	-0.28
Smoker: 1 = Yes	-0.1791 *	0.0185	-9.67	-0.5244 *	0.0581	-9.03
Constant	0.1261 *	0.0277	4.56	-1.3085 *	0.0797	-16.41
Observations	8774 (2869 clusters)			8774 (2869 clusters)		

Indications of significance: * 1%, ** 5%, *** 10%.

As can be seen in table 2 the estimate for the treatment effect is significant for both the linear probability model and the probit model. As for the remaining covariates most are significant, with the exception of self-assessed health. When including only the observations on either side of the threshold the results of table 3 are obtained.

Table 3: Sample age thirty to thirty-one

	OLS			**Probit**		
Dependent Variable: PHI	Estimate	SE	t-value	Estimate	SE	z-value
Treatment: 1 = LHC	0.0691 ***	0.0396	1.75	0.2253 ***	0.1288	1.75
Sex: 1 = Female	0.1500 **	0.0596	2.52	0.4967 **	0.1936	2.57
Income	8.78e-06 *	1.33e-06	6.61	3.35e-05 *	5.57e-06	6.02
Chronic condition	0.0676	0.0766	0.88	0.3109	0.2458	1.26
Self-assessed health, 1 = bad	-0.0813	0.0788	-1.03	-0.2896	0.2756	-1.05
Smoker: 1 = Yes	-0.2006 *	0.0630	-3.18	-0.6438 *	0.2018	-3.19
Constant	0.0750	0.0857	0.87	-1.5840 *	0.3048	-5.20
Observations	307 (229 clusters)			307 (229 clusters)		

Indications of significance: * 1%, ** 5%, *** 10%.

The result for the minimal bandwidth indicates a significant treatment effect, but only at the 10% level. In addition to treatment only sex, income and smoking are significant. In order to further investigate this result the bandwidth is widened to three years on each side of the threshold leading us to the following output.

Table 4: Sample age twenty-eight to thirty-three

	OLS			**Probit**		
Dependent Variable: PHI	Estimate	SE	t-value	Estimate	SE	z-value
Treatment: 1 = LHC	0.0450	0.0334	1.35	0.1484	0.1026	1.45
Sex: 1 = Female	0.1333 *	0.0438	3.05	0.4195 *	0.1334	3.15
Income	8.04e-06 *	1.19e-06	6.76	2.54e-05 *	5.81e-06	4.38
Chronic condition	0.0551	0.0527	1.05	0.1663	0.1613	1.03
Self-assessed health, 1 = bad	0.0307	0.0565	0.54	0.1265	0.1727	0.73
Smoker: 1 = Yes	-0.1930 *	0.0428	-4.51	-0.5832 *	0.1327	-4.39
Constant	0.0934	0.0676	1.38	-1.2796 *	0.2887	-4.43
Observations	910 (458 clusters)			910 (458 clusters)		

Indications of significance: * 1%, ** 5%, *** 10%.

Table 4 contains the estimate for age twenty-eight to thirty-three. Compared to table 3 above the treatment effect is actually not significant while the estimate of the covariates is consistent. The final specification include a bandwidth of ten years on each side of the threshold and is presented in table 5.

Table 5: Sample age twenty to forty-one

	OLS			Probit		
Dependent Variable: PHI	Estimate	SE	t-value	Estimate	SE	z-value
Treatment: 1 = LHC	0.0825 *	0.0255	3.24	0.2034 *	0.0757	2.69
Sex: 1 = Female	0.1197 *	0.0254	4.71	0.3730 *	0.0783	4.76
Income	6.75e-06 *	7.81e-07	8.65	2.52e-05 *	2.53e-06	9.95
Chronic condition	-0.0136	0.0297	-0.46	-0.0136	0.0943	-0.14
Self-assessed health, 1 = bad	-0.0064	0.0291	-0.22	0.0127	0.0903	0.14
Smoker: 1 = Yes	-0.1698 *	0.0233	-7.29	-0.5055 *	0.0748	-6.76
Constant	0.1142 *	0.0367	3.11	-1.3178 *	0.1191	-11.06
Observations	3607 (1517 clusters)			3607 (1517 clusters)		

Indications of significance: * 1%, ** 5%, *** 10%.

Here the estimates are comparable to those of the entire sample except for the dummy variable indicating a chronic condition. When considering the various specifications above the significance of the treatment effect appear somewhat unstable. To further investigate this every possible bandwidth in the span twenty to forty-one is investigated. The significance of the treatment effect changes quite a bit depending on the bandwidth. Most bandwidths close to the threshold show insignificant values for the treatment effect even at the 10% significance level, the notable exception being when only including age thirty and thirty-one. To complicate things further the significance of the treatment effect increases when including observations further away from the threshold in terms of age. When reaching the bandwidth twenty to forty-one years of age the treatment effect is highly significant, which is then consistent all the way up to when the entire sample is included as can be seen in table 2. The details can be found in table A1 in appendix A. The complete regression outputs are included in appendix A.

4.3 Discussion

The first part of the analysis is the validation of the RD design. There are no direct indications that contradict the validity of the design in terms of the density of the forcing variable or the continuity of the covariate investigated. The main issue when it comes to validation is illustrated in figure 3, where it is difficult to determine whether there is an actual discontinuity in the proportion of individuals with PHI. The troublesome part being the sharp rise in PHI up to age thirty-one and then sudden change where the proportion remains more or less constant up to the retirement age. While there is some doubt concerning the validity of the design, it is impossible to disregard how improbable it is for PHI to peak at the exact same age as the proposed discontinuity. The validity of the design is therefore questionable and caution is advised when interpreting the results.

The sample slightly overstates the PHI proportion due to the data in the Hilda survey not making it possible to differentiate between hospital and general cover. According to Thomson et al (2012) 54.4% of the Australian population have general coverage but only 46.8% have hospital cover as required by the LHC. The Hilda sample hovers around 50% with PHI after age thirty-one, the proportion is therefore slightly inflated, but only by a few percentage points which likely does not affect the outcome.

When it comes to the result of Palangkaraya and Yong (2007) the attempted replication using data from the Hilda survey failed. The treatment effect is not significant and most covariates lack significance or contradict the result of Palangkaraya and Yong (2007). There is a multitude of possible reasons why this might be the case. First of all the bandwidth is the best available for the NHS 2001 data, for the Hilda survey however, it is arbitrary and the estimate suffer from loss of valid observations in the interval thirty to thirty-four. Additionally, the difference in the number of chronic conditions can possibly explain the lack of significance in this variable, one chronic condition as in the Hilda survey might simply not be a great enough factor to affect the insurance decision. The three chronic conditions of Palangkaraya and Yong (2007) on the other hand could well provide a much more clear indication of individuals with increased health risk and therefore affecting the insurance decision. The effect of disposable income on PHI can be found in both estimates, with similar magnitude, confirming the expected effect of income on PHI. The attempted use of the Hilda material does not support the results of the linear probability model of Palangkaraya and Yong (2007), on the other hand the current result is not strong enough to actually question the previous findings. In evaluating the findings of the work of Palangkaraya and Yong (2007) the result is inconclusive at best.

In trying to analyze the more general RD estimate first note that the probit model and the linear probability model show perfectly consistent estimates for all specifications, indicating that any potential problems are not caused by estimation technique but rather by specification or method. Regarding the estimates of various specifications, the significance of the treatment effect is ambiguous and varies depending on the bandwidth used. The reason for this is difficult to discern, the sample in the region just around the threshold might be too small and thus not contain enough information to allow for reliable estimation or the RD design itself might not be a valid method. Do note that no specific parametric form was assumed and that the analysis is carried out using only the probit and linear probability model. The reason for not using a specific functional form in the analysis is the ambiguity due to lack of a distinct discontinuity at the threshold. Depending on which functional form was attempted very different results were obtained, some proposing a positive discontinuity and some indicating a negative jump. With no real support being available for any of these functional forms the analysis is considered too volatile and the result too dependent on assumptions of the specification and was therefore left out due to being inconclusive and potentially invalid. Examples when using different functional forms can be found in appendix B.

The lack of conclusive evidence in both the attempted parts of the analysis force us to return to the work of Lee and Lemieux (2010). As stated by Lee and Lemieux (2010), using age as the forcing variable causes a problem due to its inevitability and predictability. The argument is based on individuals being able to act preemptively and adapt prior to reaching the threshold. Note that this does not mean there is no effect, it simply translates into the effect not being observable or possible to estimate using a RD design. While this may prove to be the reason for the inconclusive RD estimates it offers some light on the very sharp increase in the proportion of the population with PHI leading up to the threshold and the stagnation after that point. It is possible that the LHC might in fact force individuals to decide prior to turning thirty-one if they are interested in PHI or not, and those that are adapt and attain PHI earlier than just prior to the cutoff, rendering the effect unobservable. A possible explanation for the stable proportion with PHI after age thirty-one can be found in the positive effect of income on PHI being dampened, or indeed cancelled, by the yearly 2% increase in price of PHI after turning thirty-one.

While we know that the LHC exists it is not possible to estimate the effect using the Hilda survey and the RD design. This forces the question if the RD design estimate of Palangkaraya and Yong (2007) is indeed valid. There is a possibility that the effect captured by their linear probability model is a combination of heterogeneity between treated and untreated individuals and quite a large difference in age over a period where characteristics likely differ substantially. Note that the overall result of Palangkaraya and Yong (2007) is not being disputed, only the application of RD design in the context.

5. Conclusion

5.1 Conclusion

The attempted replication of the OLS result of Palangkaraya and Yong (2007) using the Hilda sample offers no conclusive evidence in either support or in contradiction to their RD design. The treatment effect is insignificant and the covariates are generally inconclusive or very different in terms of magnitude or sign. The notable exception being income which is similar in both significance and magnitude. When it comes to the RD specification appropriate for the Hilda sample the significance of the treatment effect largely depends on the bandwidth which the analysis is confined to. The sensitivity of the estimated treatment effect when it comes to bandwidth and generally inconclusive estimates indicate that the design is either invalid or inappropriate for the context. While the design did not appear invalid when performing the graphical validation there is a significant possibility that the RD method is inappropriate due to the inevitability and predictability of treatment. The general conclusion is that the design offers inconclusive evidence and potentially suffers from an unobservable treatment effect caused by individuals adapting ahead of time in response to predictable treatment. The Hilda survey offers no conclusive evidence or additional insight on the effect of the LHC using the current method.

5.2 Further research

The issues encountered using the RD methodology for this analysis illustrate the need of additional research on how to best model specifications where RD design suffer due to inevitability and predictability of treatment. Possible solutions could be based on investigating estimates of an effect for a section just prior to the forcing threshold. The question regarding the persistence of the effect of the LHC remain unanswered and therefore also require additional analysis.

References

Imbens, G. W. & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142, p. 615-635.

Lee, D. S. & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142, p. 655-674.

Lee, D. S. & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48, p. 281-355.

Melbourne Institute of Applied Economic and Social Research, University of Melbourne. Available at: <www.melbourneinstitute.com/hilda> [Accessed 22 August 2013].

Palangkaraya, A. & Yong, J. (2007). How effective is "lifetime health cover" in raising private health insurance coverage in Australia? An assessment using regression discontinuity. *Applied Economics*, 39, p. 1361-1374.

Private Health Insurance Administration Council. (2012). *Privately Insured People with Hospital Treatment Cover - Annual Analysis. Sex, Age and State*. Kingston.

Private Health Insurance Ombudsman, Australian Government. Available at: <http://www.privatehealth.gov.au/healthinsurance/incentivessurcharges/lifetimehealthcover.htm> [Accessed 14 August 2013].

Summerfield, M., Freidin, S., Hahn, M., Ittak, P., Li, N., Macalalad, N., Watson, N., Wilkins, R. & Wooden, M. (2012). *HILDA User Manual – Release 11*. Melbourne Institute of Applied Economic and Social Research, University of Melbourne.

Thomson, S., Osborn, R., Squires, D. & Jun, M. (2012). *International Profiles of Health Care Systems, 2012.* The Commonwealth Fund, New York.

Van Der Klaauw, W. (2008), Regression–Discontinuity Analysis: A Survey of Recent Developments in Economics. *Labour*, 22, p. 219–245.

Appendix A – Regression output

Replication of the linear probability model by Palangkaraya and Yong (2007)

Dependent Variable: PHI	Coef.	Robust SE	t	P>|t|	95% Conf. Interval	
Treatment: 1 = LHC	.0269281	.0354879	0.76	0.448	-.0427287	.0965849
Sex: 1 = Female	.142226	.0363436	3.91	0.000	.0708898	.2135622
Income	5.43e-06	1.01e-06	5.39	0.000	3.45e-06	7.41e-06
Chronic condition	-.0111954	.0430406	-0.26	0.795	-.0956769	.073286
Self-assessed health, 1 = bad	-.0005562	.0459658	-0.01	0.990	-.0907793	.0896669
Smoker: 1 = Yes	-.2065689	.0329369	-6.27	0.000	-.2712183	-.1419194
Constant	.2231504	.0554656	4.02	0.000	.1142808	.3320199
Observations	1584 (830 clusters)					
F(6, 829)	20.96					
Prob > F	0.0000					
R-squared	0.1596					
Root MSE	.45674					

Linear probability model - entire sample

Dependent Variable: PHI	Coef.	Robust SE	t	P>|t|	95% Conf. Interval	
Treatment: 1 = LHC	.117943	.0190181	6.20	0.000	.0806524	.1552336
Sex: 1 = Female	.1361979	.018711	7.28	0.000	.0995094	.1728864
Income	6.46e-06	5.34e-07	12.10	0.000	5.41e-06	7.51e-06
Chronic condition	-.0613091	.0190079	-3.23	0.001	-.0985796	-.0240386
Self-assessed health, 1 = bad	-.0101219	.0192139	-0.53	0.598	-.0477964	.0275526
Smoker: 1 = Yes	-.1791365	.0185171	-9.67	0.000	-.2154448	-.1428283
Constant	.1261063	.027677	4.56	0.000	.0718374	.1803752
Observations	8774 (2869 clusters)					
F(6, 2868)	83.55					
Prob > F	0.0000					
R-squared	0.2018					
Root MSE	0.44495					

Probit model – entire sample

Dependent Variable: PHI	Coef.	Robust SE	z	P>|z|	95% Conf. Interval	
Treatment: 1 = LHC	.3277222	.056971	5.75	0.000	.216061	.4393834
Sex: 1 = Female	.4222804	.057184	7.38	0.000	.3102018	.5343591
Income	.0000249	1.45e-06	17.22	0.000	.0000221	.0000277
Chronic condition	-.1272315	.0567831	-2.24	0.025	-.2385244	-.0159387
Self-assessed health, 1 = bad	-.0171579	.0604965	-0.28	0.777	-.135729	.1014131
Smoker: 1 = Yes	-.5244182	.0581035	-9.03	0.000	-.6382989	-.4105374
Constant	-1.308472	.0797209	-16.41	0.000	-1.464722	-1.152222
Observations	8774 (2869 clusters)					
Wald chi2(6)	496.10					
Prob > chi2	0.0000					
Pseudo R2	0.1798					
Log pseudolikelihood	-4956.7218					
Iteration 0	-6043.6736					
Iteration 1	-4976.2739					
Iteration 2	-4956.7314					
Iteration 3	-4956.7218					
Iteration 4	-4956.7218					

Linear probability model – age 30 – 31

Dependent Variable: PHI	Coef.	Robust SE	t	P>|t|	95% Conf.	Interval
Treatment: 1 = LHC	.0691146	.0395647	1.75	0.082	-.0088446	.1470737
Sex: 1 = Female	.1500426	.0596024	2.52	0.013	.0326007	.2674846
Income	8.78e-06	1.33e-06	6.61	0.000	6.17e-06	.0000114
Chronic condition	.0675608	.076561	0.88	0.378	-.0832968	.2184183
Self-assessed health, 1 = bad	-.0813396	.0788076	-1.03	0.303	-.2366238	.0739447
Smoker: 1 = Yes	-.2006268	.0630244	-3.18	0.002	-.3248116	.0764421
Constant	0.749773	.0857093	0.87	0.383	-.0939063	.2438609
Observations	307 (229 clusters)					
F(6, 228)	15.97					
Prob > F	0.0000					
R-squared	0.2669					
Root MSE	.43307					

Probit model – age 30 – 31

Dependent Variable: PHI	Coef.	Robust SE	z	P>|z|	95% Conf.	Interval
Treatment: 1 = LHC	.2252814	.1288138	1.75	0.080	-.027189	.4777518
Sex: 1 = Female	.4967282	.1935759	2.57	0.010	.1173263	.87613
Income	.0000335	5.57e-06	6.02	0.000	.0000226	.0000444
Chronic condition	.310922	.2458156	1.26	0.296	-.1708677	.7927117
Self-assessed health, 1 = bad	-.2896398	.2756481	-1.05	0.587	-.8299002	.2506206
Smoker: 1 = Yes	-.6437536	.2017523	-3.19	0.000	-1.039181	-.2483263
Constant	-1.583977	.3048359	-5.20	0.168	-2.181445	-.9865097
Observations	307 (229 clusters)					
Wald chi2(6)	56.37					
Prob > chi2	0.0000					
Pseudo R2	0.2391					
Log pseudolikelihood	-161.90997					
Iteration 0	-212.79456					
Iteration 1	-162.33148					
Iteration 2	-161.91028					
Iteration 3	-161.90997					
Iteration 4	-161.90997					

Linear probability model – age 28 – 33

Dependent Variable: PHI	Coef.	Robust SE	t	P>|t|	95% Conf.	Interval
Treatment: 1 = LHC	.0450036	.0334219	1.35	0.179	-.0206761	.1106833
Sex: 1 = Female	.133325	.0437645	3.05	0.002	.0473203	.2193297
Income	8.04e-06	1.19e-06	6.76	0.000	5.70e-06	.0000104
Chronic condition	.0550638	.0526811	1.05	0.296	-.0484635	.158591
Self-assessed health, 1 = bad	.0306554	.0564534	0.54	0.587	-.0802851	.1415959
Smoker: 1 = Yes	-.1929996	.0427755	-4.51	0.000	-.2770607	-.1089384
Constant	.0933808	.0676352	1.38	0.168	-.0395337	.2262953
Observations	910 (458 clusters)					
F(6, 457)	21.66					
Prob > F	0.0000					
R-squared	0.2273					
Root MSE	.44066					

Probit model – age 28 – 33

Dependent Variable: PHI	Coef.	Robust SE	z	P>|z|	95% Conf.	Interval
Treatment: 1 = LHC	.1483655	.1026255	1.45	0.148	-.0527769	.3495078
Sex: 1 = Female	.4195478	.1333,785	3.15	0.002	.1581308	.6809648
Income	.0000254	5.81e-06	4.38	0.000	.0000141	.0000368
Chronic condition	.1663171	.1612799	1.03	0.302	-.1497856	.4824198
Self-assessed health, 1 = bad	.1264757	.1727021	0.73	0.464	-.2120142	.4649656
Smoker: 1 = Yes	-.5831534	.1326952	-4.39	0.000	-.8432312	-.3230755
Constant	-1.279641	.288737	-4.43	0.000	-1.845555	-.7137264
Observations	910 (458 clusters)					
Wald chi2(6)	64.32					
Prob > chi2	0.0000					
Pseudo R2	0.1868					
Log pseudolikelihood	-512.01074					
Iteration 0	-629.6008					
Iteration 1	-513.08128					
Iteration 2	-512.01161					
Iteration 3	-512.01074					
Iteration 4	-512.01074					

Linear probability model – age 20 – 41

Dependent Variable: PHI	Coef.	Robust SE	t	P>|t|	95% Conf.	Interval
Treatment: 1 = LHC	.0824881	.025496	3.24	0.001	.032477	.1324992
Sex: 1 = Female	.1196768	.0254353	4.71	0.000	.0697847	.1695688
Income	6.75e-06	7.81e-07	8.65	0.000	5.22e-06	8.29e-06
Chronic condition	-.0136172	.0296873	-0.46	0.647	.0446154	.0446154
Self-assessed health, 1 = bad	-.0064109	.0290767	-0.22	0.826	.0506239	.0506239
Smoker: 1 = Yes	-.169781	.0232838	-7.29	0.000	.1241091	-.1241091
Constant	.1142018	.036701	3.11	0.002	.186192	.186192
Observations	3607 (1517 clusters)					
F(6, 1516)	39.26					
Prob > F	0.0000					
R-squared	0.1862					
Root MSE	.44322					

Probit model – age 20 – 41

Dependent Variable: PHI	Coef.	Robust SE	z	P>|z|	95% Conf.	Interval
Treatment: 1 = LHC	.203393	.0756976	2.69	0.007	.0550284	.3517575
Sex: 1 = Female	.3730399	.0783089	4.76	0.000	.2195572	.5265226
Income	.0000252	2.53e-06	9.95	0.000	.0000203	.0000302
Chronic condition	-.0136154	.0942908	-0.14	0.885	-.198422	.1711912
Self-assessed health, 1 = bad	.0127253	.0902621	0.14	0.888	-.1641853	.1896358
Smoker: 1 = Yes	-.5054645	.0747863	-6.76	0.000	-.652043	-.358886
Constant	-1.317775	.1191341	-11.06	0.000	-1.551294	-1.084297
Observations	3607 (1517 clusters)					
Wald chi2(6)	193.78					
Prob > chi2	0.000					
Pseudo R2	0.1656					
Log pseudolikelihood	-2031.1408					
Iteration 0	-2434.3564					
Iteration 1	-2037.0772					
Iteration 2	-2031.1424					
Iteration 3	-2031.1408					
Iteration 4	-2031.1408					

Table A1:

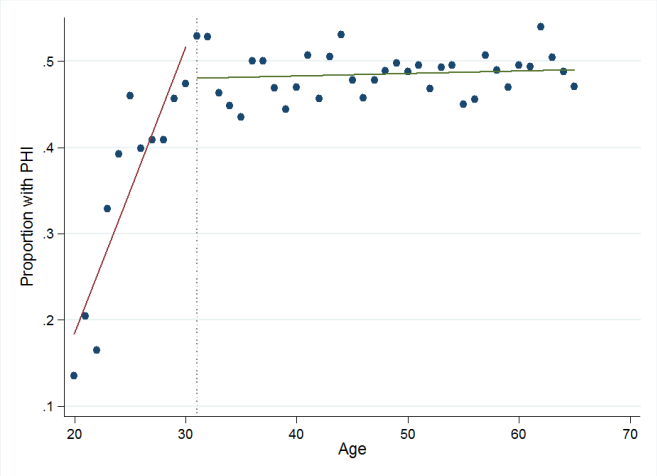
Significance of the treatment effect depending on bandwidth

Dependent Variable: PHI			**OLS**		**Probit**	
Bandwidth	No. Obs.	Clusters	Estimate	Significance	Estimate	Significance
30 - 31	307	229	0.0691	0.082	0.2253	0.080
29 - 32	607	336	0.0691	0.206	0.1400	0.201
28 - 33	910	458	0.0450	0.179	0.1484	0.002
27 - 34	1210	578	0.0306	0.334	0.1051	0.286
26 - 35	1521	702	0.0305	0.319	0.1006	0.289
25 - 36	1858	841	0.0246	0.411	0.0590	0.516
24 - 37	2202	977	0.0295	0.307	0.0677	0.432
23 - 38	2555	1113	0.0343	0.224	0.0751	0.368
22 - 39	2933	1265	0.0555	0.039	0.1302	0.103
21 - 40	3288	1404	0.0684	0.009	0.1659	0.032
20 - 41	3607	1517	0.0825	0.001	0.2034	0.007
Observations						

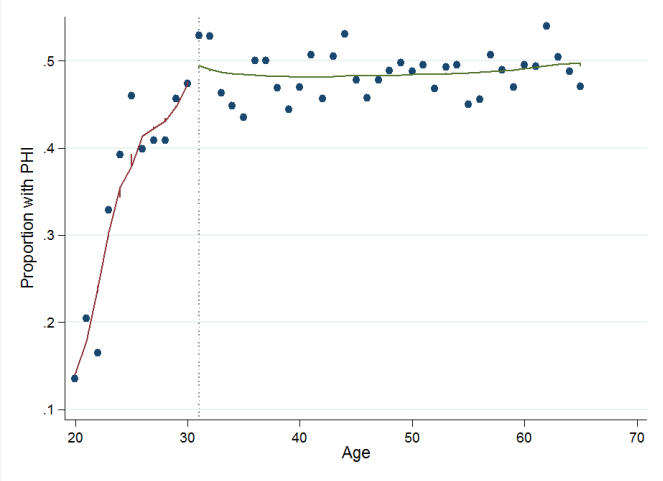
Indications of significance: * 1%, ** 5%, *** 10%.

Appendix B

Linear fit



Local linear regression (using the lowess command in Stata)



Quadratic fit on the entire material (not limited to the sample)