Expanding Health Insurance for the Elderly of the Philippines *

Michael R.M. Abrigo† Timothy J. Halliday‡ Teresa Molina§

July 16, 2021

Abstract

This paper evaluates a Filipino policy that expanded health insurance coverage of its senior citizens, aged 60 and older, in 2014. We employ an instrumental variables estimator in which the first stage is a difference-in-differences specification that exploits the age discontinuity at age 60, along with data from before and after the policy. First stage results show the expansion increased insurance coverage by approximately 16 percentage points. The compliers, those induced by the policy to obtain insurance, were disproportionately female and largely from the middle of the socioeconomic distribution. Second stage regressions indicate that out-of-pocket medical expenditures more than doubled among the compliers. We argue that this is most likely driven by an outward shift in the medical demand curve.

Key Words: Health Insurance, Compliers, Philippines, Medical Demand

JEL Codes: I13, I15, I10

*We thank Matt Harris, Amanda Kowalski, Owen O'Donnell, Jonathan Skinner, Raf van Gestel, and seminar participants at University of the Philippines, Essen Health Conference, Vanderbilt, AHEW, Kobe-Hawaii Applied Econometrics Conference, Barcelona GSE Summer Forum, Erasmus University Rotterdam, UH Manoa, and Keio University for useful comments. All errors are our own. Declarations of interest: none.

†Philippine Institute for Development Studies

‡University of Hawai'i at Mānoa and IZA

§University of Hawai'i at Mānoa and IZA. Corresponding Author. Address: 2424 Maile Way; 515A Saunders Hall; Honolulu, HI 96822. e-mail: tmolina@hawaii.edu

1 Introduction

There is a large, well-established literature that investigates how health insurance affects health-related decisions and outcomes in developed countries (Card et al., 2008; Finkelstein et al., 2012; Levy and Meltzer, 2008; Manning et al., 1987; Shigeoka, 2014). As low- and middle-income countries begin to establish and expand their various nationally-sponsored health insurance programs, we have seen the rapid emergence of a similar literature that focuses on the developing world.[1] The vast majority of programs in developing countries, however, have targeted the poor, which means that little is known about how health insurance affects people from higher up in the socioeconomic distribution of lower income countries.[2]

Notably, it is also the case that very few programs in low-income countries make special provisions for the elderly, a vulnerable group who face frequent and often serious health shocks. One exception is the 2014 amendment to the Expanded Senior Citizens Act (ESCA) in the Philippines, which granted free health insurance to all individuals aged 60 and older. This policy provides us with an opportunity to study the effects of heath insurance for the elderly in a lower income country. Unlike their counterparts in rich nations, many of the individuals who became eligible due to this expansion had been uninsured, and therefore without regular access to medical care or advice, for large portions of their lives. Expansion of insurance to a group of people that has had relatively little access to the healthcare system might have very different effects compared to similar policies in the developed world, where people have more consistent access to medical care throughout their lives.

This paper investigates how the ESCA amendment affected insurance coverage, who was most affected, and how their expenditures and utilization changed as a result of the new insurance coverage. To estimate effects on coverage, we exploit the age eligibility cutoff and use data from before and after the policy. This allows us to use a difference-in-differences strategy, comparing individuals just above and below the age cutoff of 60, before and after the policy was implemented.

[1] For excellent reviews, see Giedion et al. (2013) and Acharya et al. (2012).

[2] The "30 Baht" program in Thailand did expand insurance coverage to the middle class. However, Gruber et al. (2014) discuss how this program provided supply-side incentives for care for the poor and show how this aspect of the policy may have mattered the most. In keeping with this, they show that a major impact of the policy was to reduce infant mortality for the poor.

Using data from the Annual Poverty Indicators Survey (APIS) and the Demographic and Health Survey (DHS), we find that the policy increased insurance coverage by 16 percentage points.

Next, we ask who was induced by the policy to take up insurance. We find that the compliers in this natural experiment were disproportionately female and largely from the middle of the income and education distributions. This is consistent with several features of the Filipino insurance system, in which male senior citizens were more likely to be insured in the absence of the policy and individuals at both the bottom and top of the socioeconomic distribution had other channels through which they could have been insured before the policy. We also find some evidence of adverse selection. That is, the compliers in our context had higher inpatient utilization (in the absence of the policy) than the never-takers, those who did not take up insurance despite being eligible for free insurance under the ESCA amendment.

Identifying the characteristics of compliers, which is not frequently done in policy evaluations of health insurance programs, is a valuable exercise that helps place our study in the context of the broader literature.[3] In particular, our finding that the ESCA amendment primarily affected those in the middle of the socioeconomic distribution highlights the unique nature of this policy compared with other programs implemented in similar countries, which largely target the poor. In this sense, our study could be useful for policymakers considering expansions of their national insurance programs to broader (higher-income) populations.

In addition, we estimate the impact of insurance on expenditures and utilization, thereby providing important insights into how the impact of insurance might change as incomes increase in the developing world. We use an instrumental variables strategy, where the first stage is the regression used to estimate effect of the ESCA amendment on insurance. We find that insurance actually *increased* household per capita out-of-pocket (OOP) expenditures on health. This is in contrast with the majority of evidence in both the developed and developing world (Acharya et al., 2012; Finkelstein et al., 2012; Giedion et al., 2013; Shigeoka, 2014), with a few important exceptions from China, Indonesia, and Peru (Bernal et al., 2017; Sparrow et al., 2013;

[3]Other studies that analyze the characteristics of compliers who are induced to take up health insurance (see, for example, Kowalski (2018), Kowalski (2016), and Baillon et al. (2019)) work with a true experiment rather than a natural experiment.

Wagstaff and Lindelow, 2008). Interestingly, the increase we document is driven by expenditures on outpatient services and medicines, which are not typically included in the primarily inpatient benefits of the insurance coverage in this context.

We argue that the increase in expenditures is driven by an outward shift in the demand for medical care, rather than just a movement along the curve. As Bernal et al. (2017) show was the case in Peru, this outward shift could be due to insured individuals' increased contact with the healthcare system and complementarities between inpatient services and non-covered outpatient services or drugs. This is consistent with our finding that insurance increased the likelihood of being diagnosed with a chronic condition, specifically, hypertension. In a setting where these chronic diseases are vastly underdiagnosed, an increase in diagnoses is likely indicative of an increase in testing. This could explain why insured individuals end up spending more on outpatient services (like laboratory tests) and drugs (for treatment). Importantly, this could also mean that the shift in the demand curve is beneficial for health, although we do not examine health outcomes in this study.

Consistent with this explanation, we find evidence suggesting that insurance increased *intensive* margin utilization though we do not find significant effects on *extensive* margin utilization. Specifically, we do not detect any significant effects of insurance on the probability of visiting a health facility in the last month or having a hospital stay in the past year (though we note that these coefficient estimates are positive though imprecisely estimated and that the 95% confidence intervals include large positive effects on utilization).

Our IV estimates provide us with a local average treatment effect (the effect of insurance on the compliers), so it is important to ask whether we would see similar effects as more of the eligible elderly population begins to receive coverage. We explore treatment effect heterogeneity using marginal treatment effect methods proposed by Brinch et al. (2017). The magnitudes of our estimates suggest that never takers have larger treatment effects than compliers, though these differences are not statistically significant.

2 Health Insurance in the Philippines

The Philippine Health Insurance Corporation (PHIC) was created in 1995 with the aim of achieving more comprehensive health insurance coverage throughout the country. During our study period, there were three types of membership to PHIC: paying, lifetime, and sponsored.[4] Paying members pay their own premiums either in their entirety or shared with their employers, who are required by law to contribute. These premiums range between four and 17 USD per month depending on the employee's salary and are typically split by the employee and employer. In addition, own-account workers are encouraged to become members of the PHIC by contributing between four and seven USD per month depending on their income. Next, lifetime members are those who have paid at least 120 monthly premiums and have reached the official minimum age of retirement of 60. Finally, sponsored members have their premiums paid by a third party such as a local or the national government.

The Philippines expanded health insurance coverage in 2013 with the passage of the National Health Insurance Act (NHIA) (see Pantig (2013) for additional details). Under the 2013 NHIA, the national government was tasked to fully subsidize the premium contributions of households identified as poor by a proxy means test (many of whom were already eligible for PHIC benefits since 2010 through a conditional cash transfer program). In addition, a corollary law from 2014 that applied to the elderly population, the Republic Act 10645, which amended an earlier law, the Expanded Senior Citizens Act (ESCA), was instituted. Because of this amendment to ESCA, all individuals aged 60 and above were automatically eligible for the PHIC's sponsored program. Previously, senior citizens were only guaranteed coverage if they were lifetime members of the PHIC or if they were indigent.[5]

Although all senior citizens are automatically eligible, they are not automatically enrolled.

[4] With the adoption of the Philippine Universal Health Care Act in 2019, the NHI primary membership types were reduced to two: direct contributors, whose insurance premiums are paid directly by members with their employers, if applicable, and indirect contributors, whose premiums are paid by the national government through general taxes and other government incomes.

[5] In addition, parents of PHIC members who were over 60 years old and below a non-specified income threshold were technically considered to be covered dependents, prior to 2014. However, the data seems to suggest that knowledge of this law was limited, as we see see no jump in coverage rates at age 60 prior to 2014, which is what we would expect if senior citizen parents were enrolling as dependents.

To enroll, individuals are required to fill out a two-page form and provide proof of identification. This can be done at local PHIC offices, the Office of Senior Citizen Affairs, and hospitals. There is no waiting period. Anecdotally, the vast majority of seniors who sign up for PHIC do so at the Office of Senior Citizen Affairs. PHIC's point-of-service registration program, which started in 2013, is a relatively small program with only 146,210 enrollees in 2018.[6] In the extreme but very unlikely case that these are all senior citizens, this represents about 1.6 percent of the 8.8 million enrolled members aged 60 and above in the same year. We therefore estimate the true share of point-of-service enrollees among seniors to be under 1%.

For seniors who gained coverage as a result of the ESCA amendment, the main benefit is the coverage of basic inpatient services. Drugs are generally not covered by the PHIC, unless they are included as part of an inpatient stay. Sponsored senior members can also obtain free primary care benefits, though these are also fairly accessible and inexpensive for those who do not have insurance.

3 Data

We employ data from two different sources: the Annual Poverty Indicators Survey (APIS), which is a nationally representative consumption survey, and the Demographic and Health Survey (DHS), which collects information on demographic and health indicators. The APIS has good quality expenditures data, while the DHS has basic information on individual utilization which is not available in the APIS.[7]

3.1 APIS

The APIS contains information on household expenditures, including OOP medical expenditures, income, demographics, education, and access to government programs. Given that one of the main goals of health insurance is to reduce OOP spending, our main outcome of interest is OOP

[6]Data is not available for earlier years.

[7]We rely on the APIS for expenditures because features of the expenditures questions in the DHS make it difficult to separately identify total and OOP expenditures.

medical expenditures, which we divide by household size to obtain expenditures per capita. This includes spending on inpatient care, outpatient care, medicines, therapeutic gadgets, and equipment.

Importantly, the APIS has information on PHIC membership at the individual level.[8] We employ the APIS from survey years 2014 and 2016. Because the APIS is conducted in July, and because mandatory coverage for senior citizens was only officially implemented in November of 2014, the 2014 round of the APIS serves as pre-policy data.

As we report in Appendix Table A1, 34% of respondents aged 50-59 and only 26% of those aged 60-69 were enrolled in the PHIC in 2014. In 2016, however, we see a substantial jump in PHIC membership at age 60: from 36% to 46%. In both years, households of individuals younger than 60 had lower per capita OOP health expenditures than households with individuals aged 60 and older. The health expenditure share of total income was also higher for households with senior citizens (4%, compared to 3% for non-senior households) in both years.

In addition to total health expenditures, the APIS also records inpatient, outpatient, and medical product expenditures separately, although there are data quality issues related to these variables in the 2014 survey.[9] Average spending for inpatient expenditures and medical products was much higher than for outpatient expenditures. For all categories, spending was higher for households with senior citizens.

We report in Appendix Table A4 the average shares of health expenditures going to each of the three categories at the household-level. In 2016, households spent more than 80% of their health expenditures on medical products on average. We also complement this information with data from another source, the 2015 Family Income and Expenditures Survey (FIES), which

[8] As we discuss later, individuals are only recorded as members if they are the primary member of PHIC, not if they are qualified dependents.

[9] For 30% of individuals in the 2014 survey, there are discrepancies between total health expenditures and the sum of the three expenditures categories. This is due to respondents reporting a value for total health expenditures but being uncertain about expenditures on one or more of the categories. Instead of being coded as missing, these values are coded as zeros, resulting in a total health expenditure value that exceeds the sum of the expenditure categories. This problem does not exist in the 2016 survey. To deal with this issue, in all expenditure-related regressions, we control for an indicator equal to one for households with total health expenditures that exceed the sum of their inpatient, outpatient, and medical product expenditures. We also run specifications that exclude these households from the analysis, which drops about 2,000 individuals from the age 50-69 sample.

provides an even more detailed breakdown of the health expenditure categories.[10] In the second column of Table A4, we show that the vast majority of spending on medical products was on pharmaceuticals; very little was spent on other products (like bandages or knee braces). In addition, we show that while over half of (the small share of) outpatient spending was on "general" services like consultations, a non-trivial portion went to specialized services like X-rays or electrocardiograms.

In addition to these expenditure variables, we also employ data on education, income, and gender from the APIS. We construct separate dummies for three educational categories: incomplete primary, complete primary, and complete secondary. We also construct dummies for three income categories: the 3^{rd} decile and below, the 4^{th} to the 7^{th} deciles, and the 8^{th} decile and higher. Descriptive statistics for these categorical variables and gender are also reported in Appendix Table A1. Finally, in Appendix Figure A1, we display a histogram of the log of per capita medical expenditures (adding one before taking logs). The figure reveals that approximately 9% of our sample had zero health expenditures. For households with positive spending, the distribution is slightly right skewed.

3.2 DHS

The DHS contains information on medical utilization, PHIC membership, and other demographics. We use the 2013 and 2017 rounds. Unlike in the APIS, where PHIC enrollment refers only to primary members, the DHS also includes individuals who are covered under the PHIC as dependents of primary members.[11] This accounts for the higher percentages of people covered by the PHIC in the DHS than the APIS. In 2013, 63% percent of people aged between 50 and 59 were covered by the PHIC, while 61% of those aged between 60 and 69 were. In 2017, however, we see a substantial jump in PHIC coverage among those aged 60 to 69 to 83%, while coverage in the younger group increased only slightly to 67%. In both years, individuals aged 60 to 69 were more

[10] Unfortunately, the FIES does not have information on the insurance status or ages of individual household members and therefore cannot be used in the main analysis.

[11] Non-working spouses and underage children count as qualified dependents and can also avail of benefits. However, they are not automatically enrolled and have more limited benefits than the primary member.

likely to have visited a health facility or had a hospital stay. This may be a consequence of the higher prevalence of diagnosed chronic and acute conditions among the elderly, also documented in this table.

As with the APIS, we employ information on socioeconomic status and gender from the DHS. We generate the same three categorical education dummies as in the APIS. However, because the DHS does not have comparable income information to the APIS, we employ a wealth index that is constructed from a Principal Components Analysis of a battery of questions on asset ownership. We construct three wealth categorical variables corresponding to the 1st and 2nd quintiles, the 3rd and 4th quintiles, and the highest quintile. Descriptive statistics for these covariates from the DHS are also reported in Table A2.

4 Research Design

4.1 Conceptual Framework

We consider the impact of health insurance coverage on medical utilization and other outcomes within the framework first discussed by Heckman and Vytlacil (1999) and Vytlacil (2002) and subsequently applied to the context of health insurance by Kowalski (2016).[12] We let Y_1 and Y_0 denote potential outcomes both with and without the treatment (insurance coverage in our case). Treatment status is denoted by $D \in \{0, 1\}$ where unity represents treatment. As is standard, the econometrician observes

$$Y = DY_1 + (1 - D)Y_0.$$

Consistent with the previous literature, treatment status is determined by a latent variable of the form

$$I = p_Z - U. \tag{1}$$

[12]Specifically, Vytlacil (2002) shows how the local average treatment effect framework developed by Imbens and Angrist (1994) is equivalent to selection model.

Here, p_Z (which enters the expression positively) can be interpreted as the benefits of treatment and U (which enters negatively) as the costs of treatment, so that individuals with lower values of U will take up treatment prior to those with higher values. We normalize U to be a uniform random variable on the unit interval without loss of generality. In our context, p_Z depends on a binary instrumental variable (IV) denoted by $Z \in \{0, 1\}$. Accordingly, p_Z can take on one of two values: p_1 (the probability of treatment for those with $Z = 1$) and p_0 (the probability of treatment for those with $Z = 0$). We will assume that the vector (Y_1, Y_0, U) and Z are distributed independently. Without loss of generality, we assume that $p_1 \geq p_0$. Participation is then determined by the relation $D = 1(I \geq 0)$.

We are interested in $\Delta = Y_1 - Y_0$ which is the individual-specific treatment effect. The IV research design discussed in the next sub-section will allow us to identify the local average treatment effect (LATE) which is defined as

$$LATE = E\left[\Delta | p_0 \leq U \leq p_1\right].$$

This corresponds to the average treatment effect for the subpopulation whose participation is affected by the IV. These individuals will be treated when $Z = 1$, but untreated when $Z = 0$. This group of people is typically called the compliers and constitutes $100 \times (p_1 - p_0)\%$ of the population.

4.2 Instrumental Variables Estimator

To estimate the effect of health insurance coverage on utilization, we employ a two stage least squares (2SLS) estimator which is equivalent to the IV estimator in our case. In the first stage, we estimate $p_1 - p_0$, which corresponds to the effects of the IV on participation, as discussed above. To do this, we use pre and post-policy data and exploit the discontinuity generated by the fact that all people 60 and older were able to enroll in the PHIC after November of 2014. In the second stage, we compute the 2SLS or IV estimator which, as discussed by Imbens and Angrist (1994) and Lee and Lemieux (2010), identifies the LATE.

First Stage

In the first stage, we estimate the effects of the ESCA amendment on PHIC membership. Building on earlier notation, we let D_{iat} denote whether individual i of age a (which is reported in the survey in discrete years) at time t is insured by the PHIC. Next, we define an indicator for being age 60 or older (which we denote using $SENIOR_a$), and a dummy variable $POST_t$ that is equal to one in survey years after the ESCA was amended (2016 for the APIS and 2017 for the DHS). Restricting the estimation to individuals within some bandwidth b of age 60, we then estimate the following regression:

$$\begin{aligned} D_{iat} =& \alpha_0 + \alpha_1 SENIOR_a \times POST_t + \alpha_2 SENIOR_a + \alpha_3 POST_t + \\ & g_1(a-60) + SENIOR_a \times g_2(a-60) + \nu_{iat}, \end{aligned} \tag{2}$$

where $g_1(a-60)$ and $g_2(a-60)$ are either linear or quadratic functions of $a-60$. By allowing for the age polynomials to vary above and below the cutoff, this specification takes the local linear (or quadratic) approach recommended by Gelman and Imbens (2019). The variable $SENIOR_a \times POST_t$ is equal to one for individuals who are 60 or older in the post period. This variable, which we will refer to as Z_{at} in subsequent discussions, is our IV. The parameter α_1 governs the strength of our IV and determines the magnitude of $p_1 - p_0$.

It is important to note that insurance enrollment (D_{iat}) is distinct from insurance eligibility ($SENIOR_a \times POST_t$) in this context. While seniors (post-2014) are automatically eligible for insurance once they turn 60, and while they can technically enroll at point-of-service (though we estimate less than 1% of seniors do this), there are two main reasons why seniors may not behave as though they are automatically covered. First, many do not know they are eligible. In the Philippines, de jure coverage rates calculated from national administrative data are much higher than those calculated from household surveys, which suggests that "many of those who are listed in the PhilHealth database as being entitled to free health insurance do not know of this entitlement" (Bredenkamp et al., 2017, p.12). This is likely to be especially true for senior citizens in the first few years after the ESCA amendment, when the policy was still new. Second, even those who know they are eligible may not initially have the proper identification to officially enroll. The World Bank Identification for Development Global Database estimate that about 15% of the Filipino population lacked valid identification in 2018 (World Bank, 2018).

Second Stage

In the second stage, we estimate the effects of PHIC membership on medical expenditures, hospital stays, and outpatient visits. Building on earlier notation, we let Y_{iat} denote the outcome of interest for individual i of age a at time t. Our second stage regression can then be written as

$$\begin{aligned} Y_{iat} = \theta_0 + \theta_1 \hat{D}_{iat} + \theta_2 SENIOR_a + \theta_3 POST_t + \\ f_1(a-60) + SENIOR_a \times f_2(a-60) + \mu_{iat}, \end{aligned} \tag{3}$$

where $\hat{D}_{iat}$ is predicted coverage obtained from our first stage in equation (2) and $f_1(a-60)$ and $f_2(a-60)$ are either linear or quadratic functions of $a-60$. The parameter θ_1 identifies the LATE provided that Z_{at} affects PHIC membership monotonically and satisfies the usual exclusion restriction.[13] This estimator identifies the effect of health insurance coverage on the compliers: the subset of the population that acquired insurance coverage as a consequence of the ESCA amendment.[14]

The exclusion restriction is satisfied if Z_{at} only affects outcomes (Y_{iat}) through its effect on insurance (D_{iat}), conditional on the control variables. In this context, specifically, our identifying assumption is that the difference in outcomes between those above and below the age 60 cutoff (conditional on flexible age controls) would have been the same in the post period as in the pre period, if the policy had not been implemented.

In Appendix Table A3, we provide evidence supporting the validity of our instrument. We estimate a variant of equation (2) in which we replace PHIC membership with different covariates

[13] In our setting, where the instrument increases the probability of PHIC membership (treatment) on average, monotonicity means that being over 60 after the policy would not discourage treatment for any individual. This is implied by the selection equation (1).

[14] Our set-up differs from that of the otherwise very related Bernal et al. (2017) study, where researchers do not have measures of insurance coverage for the individuals in their data and instead run reduced form regressions of their outcomes on insurance eligibility. In their context, eligibility may be more tightly linked to actual coverage than in our setting, where (as discussed above) lack of awareness about eligibility is a concern. Their estimates are intent-to-treat estimates, while our specification estimates the LATE.

as our dependent variables: indicator variables for being male and in the lowest and middle education and SES categories. We estimate the system as a Seemingly Unrelated Regression model and compute an F-test of the null that all of the interaction parameters (α_1 in equation (2)) are zero. None of the five estimates is individually significant at conventional levels, and the F-test fails to reject the null that all five estimates are zero. In sum, none of these covariates exhibits a different age discontinuity before and after the policy, which would have been indicative of a potential violation of the identification assumption.

In addition, we test for manipulation of the running variable using McCrary test methods adapted for discrete running variables (Frandsen, 2017). We find no evidence of manipulation in the APIS (from which most of our main results are drawn), but do find some evidence of manipulation in the DHS.[15] However, the "manipulation" in the DHS is most likely the result of age-heaping – the tendency of respondents to report round numbers – and it is clear from the age distribution in Appendix Figure A3 that this is coming primarily from 61-year-olds reporting their age as 60. This should not generate any measurement error in our $SENIOR$ indicator of interest.

Importantly, the use of an instrumental variable helps alleviate typical endogeneity concerns that would arise from an OLS regression of utilization or expenditures on insurance coverage. Adverse selection could lead to a positive correlation between the error term and the insurance indicator (unhealthy people choose to get insurance and also have higher utilization), while advantageous selection would result in the opposite (people who care about their health choose to get insurance and have lower utilization because they are healthier in general). Another potential source of endogeneity is reverse causality; individuals may not know they are eligible for free insurance until they go to a health facility. Using an instrument helps alleviate these concerns. We argue that the instrument, the interaction between $SENIOR_a$ and $POST_t$, is uncorrelated with the error term after we control for the main effects of $SENIOR_a$, $POST_t$, and a flexible function of age that varies above and below age 60. Senior citizens might have systematically different utilization from non-seniors, individuals the post period might have systematically dif-

[15]For the DHS, we reject the null of no manipulation for values of k, a parameter chosen by the researcher, from 0 to 0.08 at the 95% level. In the APIS, however, we cannot reject the null at any value of $k < 0.1$.

ferent utilization from those in the pre period, and utilization might vary (flexibly) with age – all for reasons unrelated to the ESCA. However, we argue that any *differential* difference in outcomes across years, right at the cutoff between seniors and non-seniors, can only be due to the policy's expansion of insurance to the elderly. Under this assumption, θ_1 provides us with a consistent estimate of the LATE.

In all expenditure-related regressions, to address the expenditure data quality issues discussed above, we control for an indicator equal to one for households with total health expenditures exceeding the sum of their inpatient, outpatient, and medical product expenditures.[16] Aside from this data quality indicator, we do not control for any other covariates because these should be exogenous to our instrument, an assumption that is supported by the results in Table A3. Our results, however, are robust to the inclusion of controls for gender, education, and socioeconomic status (not reported but available upon request).

We cluster standard errors at the household level. In addition, we cluster by age and report p-values obtained from wild cluster bootstrap methods to account for the small number of clusters.[17]

4.3 Complier Characteristics

As in any study of health insurance, it is important to ask who selects into insurance. For example, are the poor selecting in before the rich? Are those with the highest potential utilization or the lowest potential utilization selecting in first? In this section, we discuss methods from the literature that provide answers to these questions.

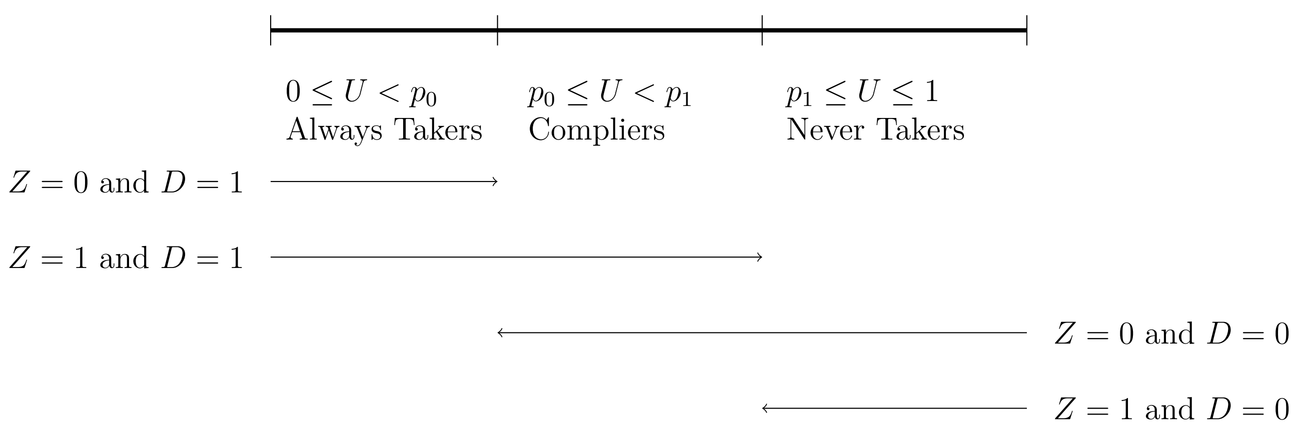
We use Figure 1 to discuss how the take-up of treatment, insurance in our case, varies across the population in our research design. This is very similar to the discussion in Kowalski (2016). First, individuals with $0 \leq U < p_0$ will always obtain treatment. These are the *always takers*, and in our context, they are individuals who take up insurance even if they are not affected by the policy change. In the data, if $Z = 0$ and $D = 1$ then the individual is an always taker (note that this condition is sufficient but not necessary to be an always taker). Second, the *compliers*

[16] We also estimate regressions that exclude these households from the analysis (Appendix Table A5).

[17] We impose the null hypothesis of $\beta = 0$ in the resampling, use 999 replications, and use Rademacher weights except in regressions with a bandwidth a less than 5 (for which we use Webb weights).

have $p_0 \leq U < p_1$. For these individuals, we will have that $Z = D$. These individuals obtain insurance if they are eligible under the the ESCA amendment ($Z = 1$) but do not if they are not eligible ($Z = 0$). Unlike the always takers, the compliers cannot be separately identified in the data. Individuals with $Z = 1$ and $D = 1$ are comprised of always takers and treated compliers. Similarly, the set of individuals for whom $Z = 0$ and $D = 0$ (who are not insured while ineligible) contains both the untreated compliers and *never takers*. Never takers have $p_1 \leq U \leq 1$ and will never seek treatment in this experimental design. Individuals who do not have insurance ($D = 0$) despite being eligible due to the ESCA ($Z = 1$) must be never takers.

Figure 1: Identifying the Compliers



We calculate the characteristics of the compliers using methods similar to those used in the existing literature (Abadie, 2002, 2003; Imbens and Rubin, 1997; Katz et al., 2001; Kowalski, 2016). Identification entails computing $E[X|D = d, p_0 \leq U < p_1]$ for $d \in \{0, 1\}$ and then taking a weighted average. In the Appendix, we show that the average characteristics of the untreated compliers will be given by the following formula (which also appears in Kowalski (2016)):

$$\begin{aligned} &\mu_X(0) \equiv E[X|D = 0, p_0 \leq U < p_1] = \\ &\frac{1}{p_1 - p_0}\left[(1 - p_0)E[X|D = 0, Z = 0] - (1 - p_1)E[X|D = 0, Z = 1]\right]. \end{aligned} \tag{4}$$

Similarly, we also show that the average of the characteristics for the treated compliers is given by

$$\begin{aligned}\mu_X(1) &\equiv E\left[X|D=1, p_0 \le U < p_1\right] = \\ &\frac{1}{p_1 - p_0}\left[p_1 E[X|D=1, Z=1] - p_0 E[X|D=1, Z=0]\right].\end{aligned} \tag{5}$$

Next, we can take a weighted sum of the averages of the untreated and treated compliers to recover the overall averages of the compliers' observables.[18]

Estimates of the conditional expectations in the above equations can be calculated directly from the data. If we let X_{iat} denote some individual characteristic, then a simple way is to estimate the following regression:

$$X_{iat} = \lambda_{NT} + \lambda_{AT} 1(AT_{iat}) + \lambda_{AT+TC} 1(ATTC_{iat}) + \lambda_{NT+UC} 1(NTUC_{iat}) + u_{iat}, \tag{6}$$

where NT identifies the never takers, AT identifies the always takers, $ATTC$ identifies the composite of the always takers and treated compliers, and $NTUC$ identifies the composite of the never takers and the untreated compliers.The different subgroups can be identified in the data as discussed in Figure 1. Each coefficient provides us with one of the conditional expectations needed for the calculations in equations (4) and (5).[19]

Estimation of equation (6) using outcomes such as medical expenditures or utilization as the dependent variables also sheds light on the presence of adverse selection. With utilization outcomes, the parameter λ_{NT+UC} identifies selection into the treatment because a non-zero value for this parameter indicates a difference in average outcomes across compliers and never takers. Because both groups are untreated, differences between the two cannot be due to heterogeneous treatment effects (which could be the case in any comparisons involving λ_{AT} or λ_{AT+TC}). If $\lambda_{NT+UC} = 0$ then there are no differences between the two groups and therefore there is no selection into treatment. If, on the other hand, $\lambda_{NT+UC} > 0$, this indicates that never takers have lower utilization than the compliers, which indicates adverse selection. Finally, $\lambda_{NT+UC} < 0$ indicates lower utilization among the compliers which indicates advantageous selection.

[18]We chose the weight optimally to minimize the variance of the estimates.

[19]Estimation of the means of the complier characteristics combined estimation of equation (6) with estimates of the propensity scores. To conduct inference, we employed the bootstrap with 1000 replications.

5 The Effect of ESCA on Insurance Coverage

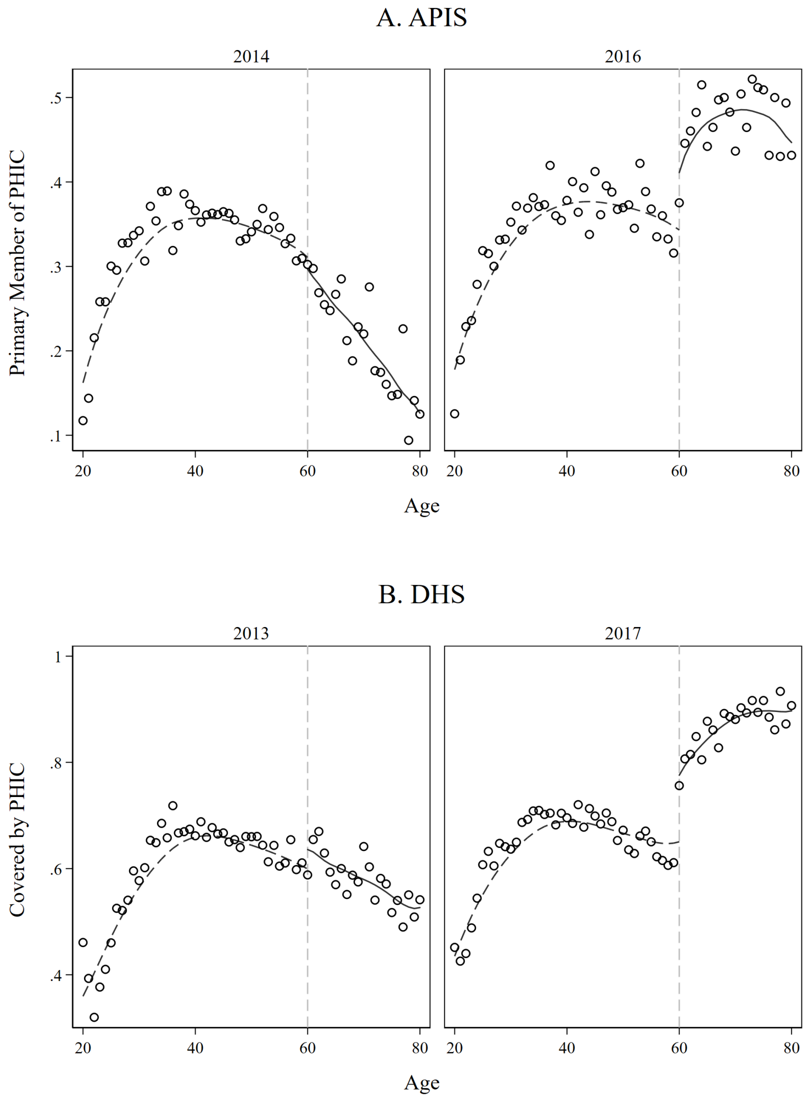
We begin with a graphical illustration of the effects of the ESCA amendment on PHIC membership. In Figure 2, we plot the share of individuals enrolled in the PHIC, by age, and graph the lowess-smoothed relationship separately for 2014 and 2016. We use the APIS in Panel A and the DHS in Panel B. Although the two surveys capture slightly different measures of PHIC coverage (described in section 3), both figures depict similar patterns. Before the policy was implemented, the relationship between coverage and age appears to be fairly smooth through age 60.[20] In contrast, in the post period for both the APIS and the DHS, there is a large discontinuous increase in insurance coverage at age 60. This comparison of the age patterns before and after the policy offers evidence that the ESCA amendment was effective at increasing coverage rates for those over 60 years old. However, it is worth noting that, in principle, the policy could have increased coverage rates of elderly Filipinos to 100% but did not. Limited awareness, sign-up time costs, or low perceived benefits of insurance may have served as barriers to enrollment, which reiterates the need to consider insurance eligibility and enrollment as two distinct variables in this setting.[21]

In Table 1, we report estimates of the corresponding regressions presented in equation (2), which serves as the first stage equation for the subsequent IV analysis. Every cell in this table reports the coefficient on the age 60 discontinuity interacted with the post dummy, denoted by $Z_{at} = POST \times SENIOR$. This is the effect of the ESCA amendment on PHIC membership. This variable is also our instrument for insurance coverage. Each column uses a different bandwidth of either 2, 5 or 10 years around the cutoff. Each row uses a different order for the polynomial $g(a-60)$. We also report the optimal order for each bandwidth. Consistent with the graphical evidence above, these results demonstrate that the policy had a sizable effect on insurance coverage rates, with estimates ranging from nine to 19 percentage points depending on the bandwidth. Within each bandwidth, our estimates are consistent across different polynomial

[20] It is worth noting that both the 2013 DHS and 2014 APIS represent two different pre-policy years, and that we note a discontinuity of essentially zero in both years (i.e., no evidence for significant pre-trends in the discontinuity), providing support for our identifying assumption. Although there is a small upward discontinuity in the 2013 DHS, this jump is small in magnitude and not statistically significant.

[21] Capuno et al. (2016) find that subsidies, information, and application assistance all increase PHIC enrollment rates, which highlights the importance of limited awareness and sign-up costs.

Figure 2: Insurance Coverage by Age and Year



Notes: Dots represent age-specific means, and lines represent the lowess-smoothed age-coverage relationship, above and below age 60. In the APIS, only primary members of PHIC are recorded as enrolled in PHIC. In the DHS, both primary members and dependents are identified as enrolled in PHIC.

Table 1: First Stage Estimates: Effect of Policy on Insurance Coverage

	APIS			DHS		
Bandwidth	2	5	10	2	5	10
Polynomial Order						
Zero	0.088	0.16	0.18	0.12	0.15	0.19
	(0.036)**	(0.024)***	(0.018)***	(0.027)***	(0.018)***	(0.014)***
	[0.05]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
One	0.091	0.16	0.18	0.12	0.15	0.19
	(0.036)**	(0.024)***	(0.018)***	(0.027)***	(0.018)***	(0.014)***
	[0.05]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Two		0.16	0.18		0.15	0.19
		(0.024)***	(0.018)***		(0.018)***	(0.014)***
		[0.00]	[0.00]		[0.00]	[0.00]
Optimal Order	0	1	2	1	2	2
N	2811	6707	13116	5435	13650	27356

Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. Each cell represents a different regression, defined by the specified bandwidth and polynomial order. The dependent variable in all regressions is an indicator for PHIC membership. We only report the coefficient (and standard error) for the $POST \times SENIOR$ interaction, but all regressions control for the main effects of $POST$, $SENIOR$, and a flexible polynomial for age that varies above and below the cutoff.

orders. Across bandwidths, estimates do vary, but they are fairly consistent across the 5 and 10-year bandwidths. Importantly, despite the fact that estimates of the levels of coverage differ across the DHS and APIS, the estimates of the impacts of the ESCA amendment are almost identical across datasets.

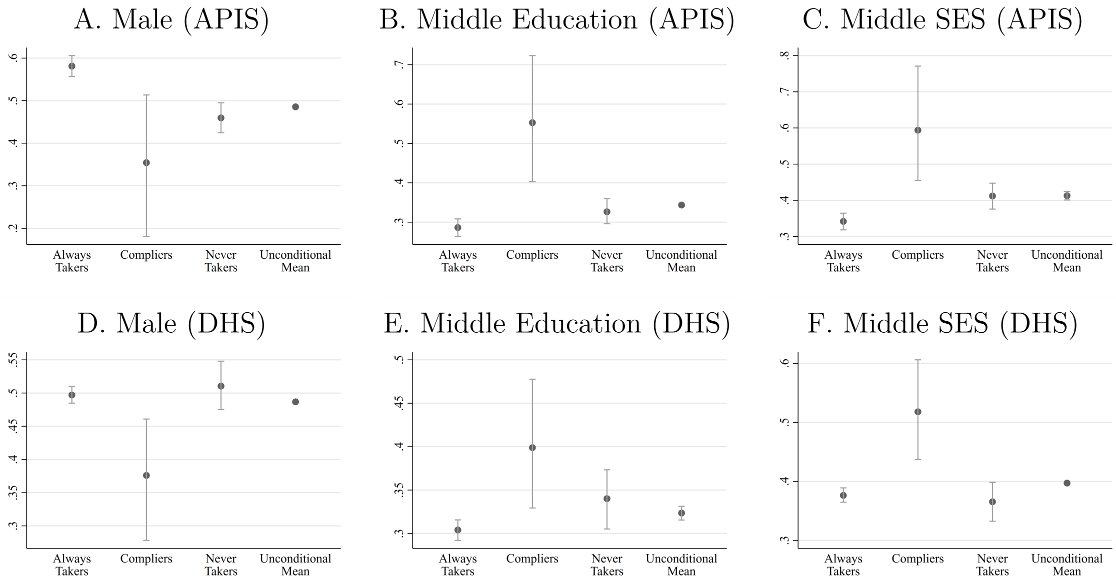
6 Describing the Compliers

6.1 Complier Characteristics

Before moving on to estimate the effect of increased insurance coverage on healthcare utilization, we compute the average characteristics of the compliers and compare them with those of the never takers and always takers using the regression parameters from equation (6). The covariates we examine are indicators for male, low/middle/high education, and low/middle/high SES. We employ a 5-year bandwidth throughout this analysis.

Figure 3 summarizes the main takeaways of this analysis, reporting the results for the male, middle education, and middle SES indicators (remaining results are reported in Appendix Figure

Figure 3: Observable Characteristics for Always Takers, Compliers, and Never Takers



Notes: Means and 95% confidence intervals are computed from 1000 bootstrapped re-samples using a 5-year bandwidth around age 60. Middle education corresponds to complete primary (but incomplete secondary). For the APIS, middle SES corresponds to the 4th to 7th deciles of the national per capita income distribution. For the DHS, middle SES corresponds to the 3rd to 4th quintiles of the wealth index distribution.

A2). Each panel corresponds to a different covariate and dataset and contains the means and 95% confidence intervals for the always takers, compliers, never takers, and full sample, computed from 1000 bootstrapped re-samples. Because the means of the compliers involve the computation of many auxiliary parameters and because compliers make up a smaller share of the sample, these estimates are noisier than those for the always and never takers.

In both datasets, the compliers were disproportionately female (panels A and D) and by-and-large from the middle of the education distribution (panels B and E) and the middle of the SES distribution (panels C and F). The complier means for the male indicator are lower than the means of the always takers and never takers. The complier means for the middle education and SES indicators are higher than the means of the always takers and never takers. In addition, as we show in Appendix Figure A2, the compliers were less likely to be in the low education and SES categories than the never takers (and in some cases, the always takers). They were also less likely to be in the high SES and education groups than the always takers.

Why was the middle class most impacted by this policy? Prior to the ESCA amendment, the

Philippines already provided health insurance coverage to the poor, though the low insurance enrollment rates among the lowest income group suggest that many were not aware of this. In addition, health insurance in the Philippines is strongly tied to employment, which means that individuals in the highest socioeconomic categories already had high coverage rates – substantially higher than the rest of the population – prior to the policy.[22] In short, the rich already had insurance, while the poor either already had insurance or were unaware of their eligibility for it. This helps explain why the compliers were less likely to be drawn from the lowest socioeconomic groups than the never takers, and less likely to be from the highest socioeconomic groups compared to the always takers.

The link between employment and insurance also explains why the compliers were more likely to be women, who had lower labor force participation than men and were therefore less likely to be insured, either as paying or lifetime members once they turned 60 (in the absence of the policy). Although married women can be covered as dependents if their spouses are PHIC members, this is not automatic and benefits are slightly less generous if there are additional dependents involved.

6.2 Testing for Selection

We now test whether there is any selection into health insurance as a consequence of the policy change. That is, were those with the highest potential utilization the ones that chose to obtain insurance first? In Table 2, we regress expenditure and utilization measures on the independent variables in equation (6): indicators for always takers, always takers with treated compliers, and never takers with untreated compliers, which leaves never takers only as the omitted category. We estimate this regression using a 5-year bandwidth and the optimal polynomial orders identified in Table 1.

As previously discussed, selection into the treatment is identified by the coefficient on the never taker/untreated complier (NT-UC) indicator. For medical expenditures and outpatient

[22] As described above, individuals who have paid at least 120 monthly premiums and are at least 60 years old become lifetime members of PHIC. Those in the highest socioeconomic groups are likely to fall in this category simply from being consistently employed in jobs where employers helped subsidize their premiums.

Table 2: Selection Tests

	Log Health Expenditures (APIS)	Hospital Stay Last Year (DHS)	Health Visit Last Month (DHS)
Always-Takers	0.28	0.064	0.038
	(0.15)*	(0.0092)***	(0.014)***
	[0.30]	[0.00]	[0.00]
Always-Takers & Treated Compliers	0.52	0.065	0.044
	(0.14)***	(0.0065)***	(0.011)***
	[0.03]	[0.04]	[0.04]
Never-Takers & Untreated Compliers	-0.11	0.026	0.0074
	(0.14)	(0.0092)***	(0.015)
	[0.56]	[0.06]	[0.45]
Polynomial Order	1	2	2
Never-Taker Dep. Var Mean	4.901	0.014	0.075
N	6707	13650	13650

Notes: Standard errors reported in parentheses ($^{***}p < 0.01,^{**}p < 0.05,^{*}p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. All regressions utilize a 5-year bandwidth. The omitted category is the Never-Taker only group. All regressions include the main effects of $POST$, $SENIOR$, and a flexible polynomial for age that varies above and below the cutoff.

health visits, we see no significant difference between the never-takers and compliers, suggesting no selection into insurance based on these two outcomes.[23]

For hospital stays, however, we see a significant difference between the never takers and the NT-UC group. The NT-UC were twice as likely to have had a hospital stay last year as the group of only never-takers (a difference of three percentage points). This indicates that, compared to the never takers, compliers were significantly more likely to have had a hospital stay in the past year. Given that never-takers make up the majority of this sample's NT-UC group, the difference in hospital stays between never-takers and compliers is likely substantially larger than three percentage points. In short, those who were more likely to select into insurance (the compliers), had substantially higher hospital utilization prior to being eligible for PHIC, which provides some evidence of adverse selection.

The groups with insurance (always takers and always takers with compliers, or AT-TC) had significantly higher expenditures and utilization than those without insurance (never takers

[23]Regressions on health expenditure shares as well as the various components of total health expenditures reveal the same results.

and NT-UC), which could be due to the treatment effect of insurance as well as selection into insurance. Across all three columns, there are no significant differences between the always-takers and the AT-TC. Because always-takers select into treatment before compliers, and because both groups are treated (i.e. insured), the difference between the two groups also combines a selection effect with treatment effect heterogeneity. For example, the absence of significant differences between the always-takers and AT-TC could be a result of no selection and no treatment effect heterogeneity, or a combination of adverse selection (with always-takers having higher baseline utilization and expenditures) and treatment effects that are increasing in the unobserved costs of insurance (i.e. larger treatment effects for the compliers).

7 The Effect of Insurance Coverage on Utilization

7.1 Expenditures and Utilization Results

In Table 3, we report second-stage IV estimates of the effect of insurance on various expenditure variables. We study five outcomes of interest, and we report estimates from a 5-year and 10-year bandwidth for each outcome, using the optimal polynomial orders identified in Table 1. We also report test statistics of weak and under identification based on Kleibergen and Paap (2006) below each of the point estimates. Appendix Table A5 shows that results are very similar when we exclude individuals with data quality issues (that is, with total health expenditures that exceed the sum of the three expenditure categories). We also show that results are not sensitive to bandwidth choice (Figure A4).

The first outcome of interest is household per capita medical expenditures from the APIS. This outcome includes OOP spending on inpatient and outpatient services, as well as medical products (which, as shown in Table A4, are primarily drugs). We find that total OOP expenditures increased. The estimate from the 10-year bandwidth is 1.230, which indicates that insurance coverage more than doubled medical expenditures. For regressions on health expenditure shares, there are positive coefficients that are large relative to the mean in both specifications, and a statistically significant coefficient for the 10-year bandwidth, indicating that health insurance led

Table 3: IV Estimates: Effect of Health Insurance on Expenditures (APIS)

	Log Health Expenditures		Health Expenditure Share		Log Inpatient Expenditures		Log Outpatient Expenditures		Log Medical Product Exp.	
Bandwidth	5	10	5	10	5	10	5	10	5	10
Enrolled in PHIC	1.42	1.23	0.027	0.046	-0.52	0.11	1.75	1.54	1.53	1.82
	(0.77)*	(0.52)**	(0.025)	(0.019)**	(0.68)	(0.47)	(0.72)**	(0.49)***	(0.71)**	(0.49)***
	[0.16]	[0.04]	[0.30]	[0.02]	[0.47]	[0.83]	[0.01]	[0.00]	[0.11]	[0.01]
Weak identification F	43.3	100.4	43.3	100.4	43.3	100.4	43.3	100.4	43.3	100.4
Underidentification F	42.9	99.0	42.9	99.0	42.9	99.0	42.9	99.0	42.9	99.0
Polynomial Order	1	2	1	2	1	2	1	2	1	2
Mean of Dep. Var.	4.99	4.94	0.036	0.036	0.62	0.62	1.23	1.25	3.77	3.69
N	6707	13116	6707	13116	6707	13116	6707	13116	6707	13116

Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. Each cell represents a different regression, defined by the specified dependent variable, bandwidth, and polynomial order. We report the results of instrumental variables regressions, with 1(Enrolled in PHIC) as the endogenous variable of interest, and $POST \times SENIOR$ as our instrument. All regressions control for the main effects of $POST$, $SENIOR$, a flexible polynomial for age that varies above and below the cutoff, and a data quality indicator equal to one for individuals from households whose total health expenditures exceed the sum of their inpatient, outpatient, and medical product expenditures.

to a 4.6 percentage point increase in the health share of expenditures.

It is notable that newly enrolled individuals in the PHIC spent more on medical expenditures despite gaining insurance coverage. If the only effect of the policy was to move consumers down their demand curves, then this estimate implies an elasticity of demand for medical care that is greater than unity, much larger than existing estimates in the literature (e.g., -0.2 from the RAND Health Insurance Experiment (Keeler and Rolph, 1988)). Consequently, we suspect that our large coefficient estimates reflect an outward shift of the demand curve and therefore caution against using these estimates to compute elasticities.

If the rise in expenditures were only reflecting movement along the demand curve, then we should see the largest increases in expenditures for inpatient services, as the main benefit of PHIC membership is coverage of inpatient services. On the contrary, the remaining columns of Table 3 show large and statistically significant increases in expenditures for outpatient services and drugs, and not for inpatient expenditures.[24] This provides further evidence that insurance in this context did not simply move individuals along the demand curve but rather, shifted the

[24]Notably, Bernal et al. (2017) also find that access to health insurance in Peru increases OOP expenditures on medications.

Table 4: IV Estimates: Effect of Health Insurance on Utilization (DHS)

	Hospital Stay Last Year		Health Visit Last Month		Hospital Stay Last Month		
Bandwidth	5	10	5	10	5	10	15
Enrolled in PHIC	0.0056	0.0043	0.051	0.016	0.24	0.19	0.24
	(0.060)	(0.034)	(0.079)	(0.046)	(0.35)	(0.17)	(0.13)*
	[0.91]	[0.88]	[0.38]	[0.74]	[0.12]	[0.23]	[0.05]
Weak identification F	66.6	190.0	66.6	190.0	4.57	20.6	33.1
Underidentification F	65.9	186.3	65.9	186.3	4.57	20.4	32.7
Polynomial Order	2	2	2	2	2	2	2
Mean of Dep. Var.	0.060	0.056	0.11	0.10	0.13	0.14	0.14
N	13650	27356	13650	27356	1469	2827	4081

Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01, ^{**} p < 0.05, ^{*} p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. Each cell represents a different regression, defined by the specified dependent variable, bandwidth, and polynomial order. We report the results of instrumental variables regressions, with 1(Enrolled in PHIC) as the endogenous variable of interest, and $POST \times SENIOR$ as our instrument. All regressions control for the main effects of $POST$, $SENIOR$, and a flexible polynomial for age that varies above and below the cutoff. The Hospital Stay Last Month variable is only available for those with Health Visit Last Month equal to one.

demand curve out.

What could have caused this shift in the demand curve for primarily non-covered healthcare? One explanation is that insured individuals spent more time with doctors and the healthcare system in general, which resulted in better knowledge about their health status and/or higher demand for healthcare (Bernal et al., 2017). This could have happened if increased interactions with doctors revealed new information about insured patients' health, or if insured patients used more inpatient services that have complementarities with outpatient services and drugs. To examine the plausibility of this hypothesis, we investigate utilization measures from the DHS. In Table 4, we investigate whether insurance increased the likelihood of an individual using inpatient services in the past year or visiting any health facility in the past month. We find no significant effects of insurance on either of these variables, but note that the coefficients are positive but imprecisely estimated (95% confidence intervals cannot rule out large positive effects relative to the means).

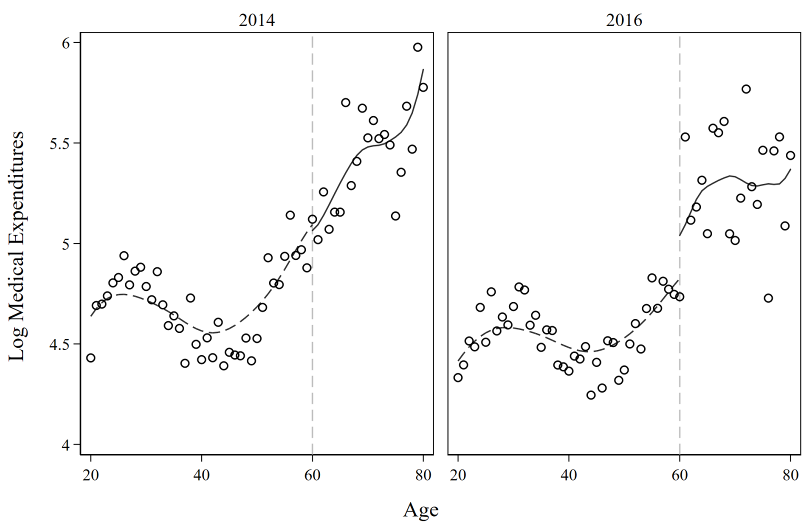
We do, however, find some evidence of increased intensity of utilization. In the last three columns of Table 4, the outcome variable is an indicator for individuals who went to a hospital after visiting a health facility. Note that this outcome is conditional on having visited a health facility in the past month. Because of the substantially smaller sample size available for this variable (which results in lower first-stage F-statistics), we also report the results from a 15-year bandwidth. Across all three specifications, estimated coefficients are around 0.2, which represents a large magnitude relative to the dependent variable mean of 0.14. Although these estimates are not significant in either the 5 or 10-year bandwidth specifications, the estimate in the 15-year bandwidth specification is significant at the 10% level. Insurance does appear to be have increased the intensity of utilization, which could have driven the outward shift in the demand curve through the mechanisms described above.

In Appendix section A.3, we also discuss results from conditional quantile IV regressions, which provide estimates of the effect of insurance on total health expenditures at different deciles of the expenditure distribution. Table A6 and Figure A5 show that the effect of insurance is increasing across deciles. In fact, the effects are only significant (at the 10% level) for the top three deciles of the expenditures distribution, implying more movement at the upper end of the distribution. This is exactly what we would see if insurance changed the intensive but not the extensive margin of utilization (though it could also be indicative of individuals going from very low to very high spending as a result of insurance).

These results are also consistent with our earlier characterization of the compliers as from the middle of the socioeconomic distribution. Low-income individuals would be unlikely to have the means to pay for follow-up treatments or care, while higher-income individuals would be more likely to have been receiving this higher-intensity care prior to the policy. Finally, previous results from Table 2 indicate that the untreated compliers were more likely to have been hospitalized relative to the never takers, once again indicating that the compliers were likely to have had some medical utilization even in the absence of PHIC membership.

We have interpreted our results thus far as evidence that insurance causally increased health expenditures for individuals who gained insurance coverage as a result of the policy. However, the positive coefficients that we estimate in Table 3 could also be the result of individuals just below the age cutoff *reducing* expenditures in anticipation of gaining free coverage at age 60; in fact, summary statistics in Appendix Table A1 show that individuals under age 60 had lower health expenditures in 2016 than 2014. To investigate this possibility, we plot average log health expenditures by age and graph the lowess-smoothed relationship separately for 2014 and 2016. We first note that, consistent with our estimates in Table 3, Figure 4 shows a positive discontinuity at age 60 that is larger in 2016 than in 2014.[25] We also note that the patterns in this figure are inconsistent with the idea that individuals close to the cutoff might be delaying their utilization until turning 60. First, total health expenditures were lower in 2016 than 2014 for all age groups – not only for those slightly younger than age 60.[26] Second, while those just under 60 years old are the ones who should be most likely to delay expenditures, it appears to be those in their early fifties with the lowest expenditures in 2016.

Figure 4: Health Expenditures by Age and Year



Notes: Dots represent age-specific means of log per capita health expenditures, and lines represent the lowess-smoothed age-expenditure relationship, above and below age 60.

In Appendix section A.4, we explore the plausibility of another alternative interpretation – increases in utilization leading to point-of-service enrollment (as opposed to insurance leading to more utilization) – and argue it is unlikely. We note, however, that reduced form regressions of

[25]This figure also shows that age gradients just below the cutoff appear similar in 2014 and 2016: indeed, in regressions that allow for year-specific polynomials (available upon request), we find no significant differences in the pre-period age gradients across years. This helps validate our use of a single polynomial across years in the regressions above.

[26]This is likely due to sampling variation, as other sources do not show large decreases in OOP health expenditures over this time period.

outcomes on the instrument still provide the causal effect of the policy change. These regressions are reported in Tables A7 and A8. Finally, Appendix section A.4 also shows that IV results are similar when we use age fixed effects instead of age polynomials, which results in a more standard generalized difference-in-differences specification (Tables A9 and A10).

7.2 Chronic Conditions Results

Table 5: Effect of Health Insurance on Chronic and Acute Conditions (DHS)

	Chronic Condition		Acute Condition		Hypertension		Diabetes		Cancer	
Year	2013	2017	2013	2017	2013	2017	2013	2017	2013	2017
Age 60 and Older	-0.015	0.015	-0.000	0.012	-0.012	0.013	-0.003	0.000	0.000	0.001
	(0.012)	(0.009)*	(0.014)	(0.010)	(0.010)	(0.007)**	(0.006)	(0.004)	(0.002)	(0.001)
	[0.47]	[0.05]	[1.00]	[0.36]	[0.45]	[0.03]	[0.60]	[0.99]	[0.95]	[0.36]
Mean of Dep. Var.	0.074	0.067	0.13	0.10	0.054	0.050	0.019	0.015	0.0017	0.0017
N	9410	17946	9410	17946	9410	17946	9410	17946	9410	17946

Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. Each cell reports the coefficient on $SENIOR$ from a different regression, defined by the specified year and outcome variable. All regressions use a 10-year bandwidth and control for a first-order polynomial for age that varies above and below the cutoff.

In Table 5, we explore one specific reason why insurance might have led to higher spending, by investigating how insurance affected chronic condition diagnoses using the DHS. We first note that the share of individuals with a chronic disease diagnosis in our sample is very low (7% overall, 5% for hypertension, 2% for diabetes, and less than 0.2% for cancer), not because actual prevalence is this low but because of severe underdiagnosis.[27] True prevalence rates for hypertension, for example, are estimated to be 35-40% for this age group (FNRI-DOST, 2013). In this setting of drastic underdiagnosis, changes in the share of individuals with a chronic disease diagnosis should be interpreted as a change in the share that have been tested (rather than a change in actual prevalence).

Because chronic condition diagnoses are so rare, we use a more parsimonious, reduced-form approach and estimate two separate regression discontinuity specifications for 2013 and 2017.

[27]Measurement error could also contribute to the low rates in this context, as the survey respondent may not have full information about all household members.

The results in Table 5 show that after the implementation of the policy, individuals above the age of 60 were significantly more likely to be diagnosed with a chronic condition. This relationship does not exist prior to the policy, and does not exist for acute conditions in either year. When we disaggregate the different types of chronic conditions, the effect appears to be driven by increases in hypertension diagnoses (the most common of the three). Hypertension, like the other two chronic conditions, are often treated with medication which is, generally, not covered by the PHIC. This could be an important reason for the increase in drug spending that we document. In addition, the finding that individuals were more likely to be tested for various conditions could also explain the increase in outpatient spending, which includes various types of medical tests.

8 External Validity

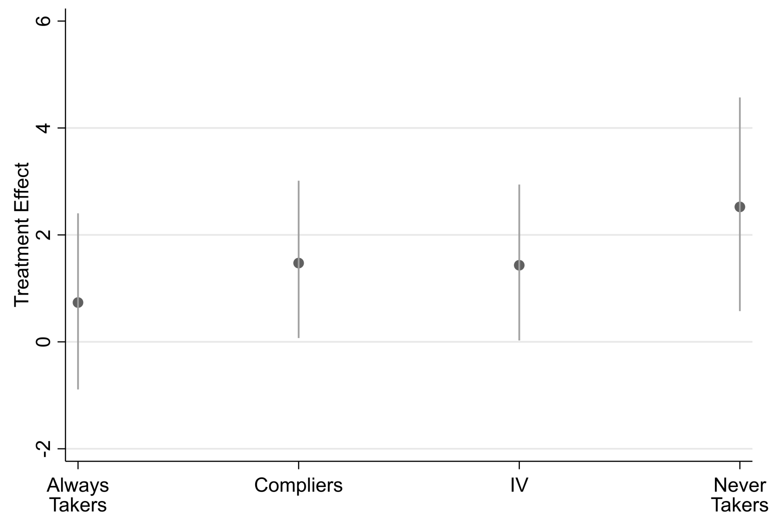
An important question is whether or not our estimate of the LATE is externally valid (i.e., can be applied to the never takers or the always takers). To answer this, we employ methods from Brinch et al. (2017) who build on Heckman and Vytlacil (2007) and approach the question of external validity by using the marginal treatment effect (MTE):

$$MTE(p) = E\left[\Delta | U = p\right] = E\left[Y_1 - Y_0 | U = p\right].$$

The MTE, first employed by Björklund and Moffitt (1987), can be used to estimate treatment effects for the never and always takers in addition to the compliers. We also formally test for heterogeneity across these groups. We report the main results here and reserve the technical details for Appendix section A.5.

In Figure 5, we report estimates of the treatment effects for the always takers, compliers, and never takers, using total health expenditures as the outcome variable. We also report the IV estimate of the LATE (which is the same as the treatment effect for the compliers). For inference, we report means and 95% confidence intervals using 1000 bootstrapped samples. The p-value associated with a test of no heterogeneity across groups is 0.132 which means we cannot reject

Figure 5: Treatment Effects for Always Takers, Compliers, and Never Takers



Notes: Means and 95% intervals are computed from 1000 bootstrapped re-samples using a 5-year bandwidth around age 60. The plots for Always Takers, Compliers, and Never Takers correspond to estimated treatment effects (of insurance on log health expenditures) using the methods discussed in Section 8. IV is the 2SLS estimate.

the null of no treatment effect heterogeneity. However, the estimates in Figure 5 are suggestive of larger treatment effects for never takers and smaller treatment effects for always takers.[28]

9 Discussion

This paper evaluates the effects of a health insurance expansion in the Philippines, which provided free health insurance to all individuals ages 60 and older starting in 2014. We find that the policy increased insurance rates by roughly 16 percentage points and that the compliers were largely drawn from the middle of the income distribution. Our results highlight important policy considerations for low- and middle-income countries considering providing free health insurance for the elderly, or – given the characteristics of the compliers – expanding national insurance programs to higher-income groups.

[28]These results should be interpreted with caution. It is possible that we are underpowered to detect treatment effect heterogeneity. In addition, these tests make the assumption that outcomes are linear in the unobserved cost of treatment. Nevertheless, this evidence does suggest that we would expect to see similar (or even larger) increases in health expenditures as more of the eligible elderly population (the never takers) receive coverage.

Interestingly, IV estimates reveal that this increase in insurance coverage led to an increase in OOP expenditures. This increase in spending was driven by outpatient and drug expenditures, which are not typically covered by this insurance. We argue that these large expenditure increases were due to an outward shift of the demand curve, rather than movements down the curve in response to reductions in price.

This outward shift could have been due to insurance increasing the amount of contact that individuals have with the healthcare system.[29] As Bernal et al. (2017) argue, increased interaction with doctors might have provided insured patients with more information about their own health or increased their valuation of health. Consistent with this, we find the insurance expansion increased diagnoses of hypertension, a condition that is often treated with medication that would contribute to higher expenditures. Another related explanation is the possibility that insured patients used more covered (inpatient) services that have complementarities with non-covered services. Consistent with this, we find evidence of increased utilization of inpatient services on the intensive margin and higher spending for outpatient care and drugs. Regardless of which of the above explanations is the dominant one, our results show that the newly insured elderly in our setting spent more on healthcare, including categories that were not covered by insurance – which means they were willing to pay for it themselves.

Given the now mounting evidence that health insurance can increase OOP spending for households in developing countries (Bernal et al., 2017; Sparrow et al., 2013; Wagstaff and Lindelow, 2008), it is important to ensure that the increased expenditures reflect use of necessary care and that physicians are not charging higher prices to insured patients. However, we find that the higher medical expenditures we document are at least partially driven by the diagnosis and treatment of otherwise undetected chronic conditions. Given that this is likely to be welfare-enhancing, these findings suggest that government efforts to encourage more contact between the elderly and healthcare providers would be beneficial.

[29] There are two other related explanations for our results. First, it could be the case that doctors treat insured patients differently, recommending more aggressive treatments or medication. Another possibility is that insured patients make different health spending decisions due to the knowledge that potentially large future inpatient costs will be covered. Some combination of all of these explanations may be in play.

References

Abadie, A. (2002). Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American statistical Association*, 97(457):284–292.

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of econometrics*, 113(2):231–263.

Acharya, A., Vellakkal, S., Taylor, F., Masset, E., Satija, A., Burke, M., and Ebrahim, S. (2012). The impact of health insurance schemes for the informal sector in low-and middle-income countries: a systematic review. *The World Bank Research Observer*, 28(2):236–266.

Baillon, A., Capuno, J. J., O'Donnell, O., Tan, J. C. R., and van Wilgenburg, K. (2019). Persistent effects of temporary incentives: Evidence from a nationwide health insurance experiment.

Bernal, N., Carpio, M. A., and Klein, T. J. (2017). The effects of access to health insurance: evidence from a regression discontinuity design in peru. *Journal of Public Economics*, 154:122–136.

Björklund, A. and Moffitt, R. (1987). The estimation of wage gains and welfare gains in self-selection models. *The Review of Economics and Statistics*, pages 42–49.

Bredenkamp, C., Gomez, M. V. A., and Bales, S. (2017). Pooling health risk to protect people: an assessment of health insurance coverage in philippines. Technical report, The World Bank.

Brinch, C. N., Mogstad, M., and Wiswall, M. (2017). Beyond late with a discrete instrument. *Journal of Political Economy*, 125(4):985–1039.

Capuno, J. J., Kraft, A. D., Quimbo, S., Tan Jr, C. R., and Wagstaff, A. (2016). Effects of price, information, and transactions cost interventions to raise voluntary enrollment in a social health insurance scheme: A randomized experiment in the philippines. *Health economics*, 25(6):650–662.

Card, D., Dobkin, C., and Maestas, N. (2008). The impact of nearly universal insurance coverage on health care utilization: evidence from medicare. *American Economic Review*, 98(5):2242–58.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., and Group, O. H. S. (2012). The oregon health insurance experiment: evidence from the first year. *The Quarterly journal of economics*, 127(3):1057–1106.

FNRI-DOST (2013). *Philippine Nutrition Facts and Figures 2013: Clinical and Health Survey.* Food and Nutrition Research Institute-Department of Science and Technology (FNRI-DOST), FNRI Bldg., DOST Compound, Bicutan, Taguig City, Metro Manila, Philippines.

Frandsen, B. R. (2017). Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete. In *Regression discontinuity designs*. Emerald Publishing Limited.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Giedion, U., Andrés Alfonso, E., and Díaz, Y. (2013). The impact of universal coverage schemes in the developing world: a review of the existing evidence. World Bank Report.

Gruber, J., Hendren, N., and Townsend, R. M. (2014). The great equalizer: Health care access and infant mortality in thailand. *American Economic Journal: Applied Economics*, 6(1):91–107.

Heckman, J. J. and Vytlacil, E. J. (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the national Academy of Sciences*, 96(8):4730–4734.

Heckman, J. J. and Vytlacil, E. J. (2007). Econometric evaluation of social programs, part ii: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. *Handbook of econometrics*, 6:4875–5143.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Imbens, G. W. and Rubin, D. B. (1997). Estimating outcome distributions for compliers in instrumental variables models. *The Review of Economic Studies*, 64(4):555–574.

Katz, L. F., Kling, J. R., and Liebman, J. B. (2001). Moving to opportunity in boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics*, 116(2):607–654.

Keeler, E. B. and Rolph, J. E. (1988). The demand for episodes of treatment in the health insurance experiment. *Journal of health economics*, 7(4):337–367.

Kleibergen, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of econometrics*, 133(1):97–126.

Kowalski, A. E. (2016). Doing more when you're running late: Applying marginal treatment effect methods to examine treatment effect heterogeneity in experiments. Technical report, National Bureau of Economic Research.

Kowalski, A. E. (2018). Extrapolation using selection and moral hazard heterogeneity from within the oregon health insurance experiment. Technical report, National Bureau of Economic Research.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48:281–355.

Levy, H. and Meltzer, D. (2008). The impact of health insurance on health. *Annu. Rev. Public Health*, 29:399–409.

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., and Leibowitz, A. (1987). Health insurance and the demand for medical care: evidence from a randomized experiment. *The American economic review*, pages 251–277.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Pantig, I. M. T. (2013). Sustainability of the national government premium subsidy for indigents. *Philippine Journal of Development*, 40(1/2):35.

Shigeoka, H. (2014). The effect of patient cost sharing on utilization, health, and risk protection. *American Economic Review*, 104(7):2152–84.

Sparrow, R., Suryahadi, A., and Widyanti, W. (2013). Social health insurance for the poor: Targeting and impact of indonesia's askeskin programme. *Social science & medicine*, 96:264–271.

Vytlacil, E. (2002). Independence, monotonicity, and latent index models: An equivalence result. *Econometrica*, 70(1):331–341.

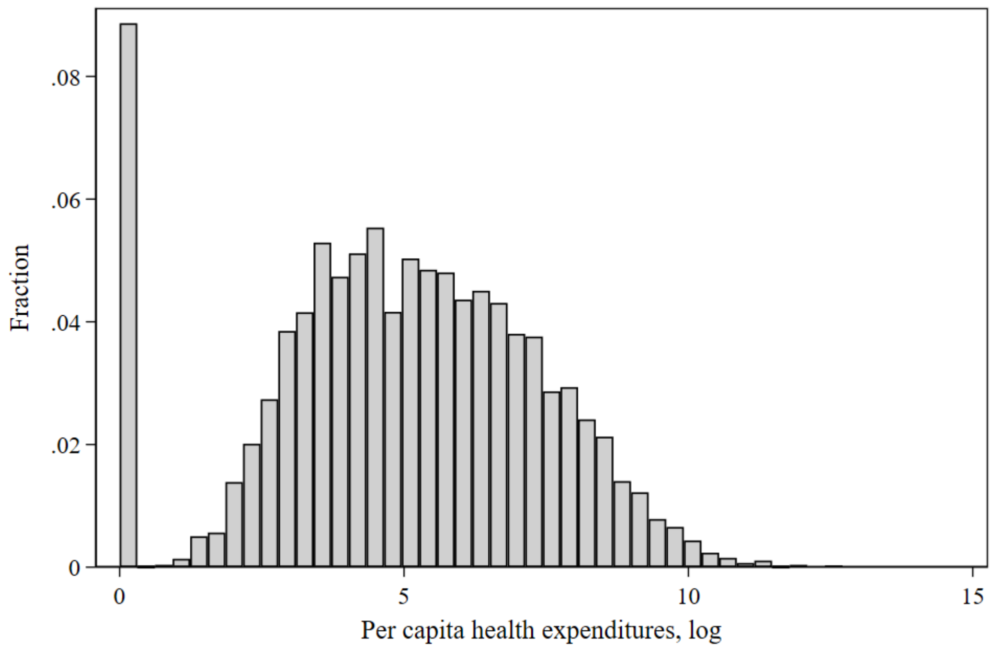
Wagstaff, A. and Lindelow, M. (2008). Can insurance increase financial risk?: The curious case of health insurance in china. *Journal of health economics*, 27(4):990–1005.

World Bank (2018). Global id4d dataset 2018.

A Online Appendix

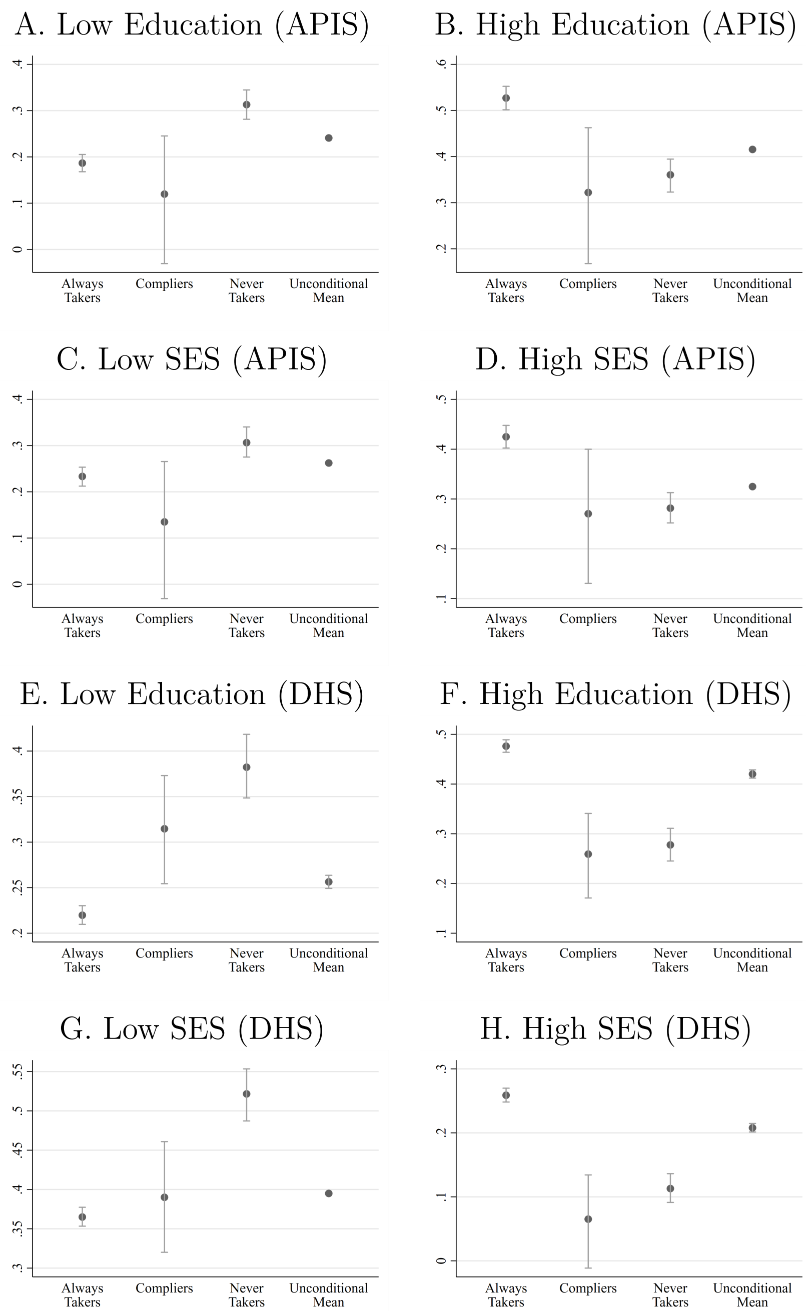
A.1 Appendix Figures and Tables

Figure A1: Distribution of PC Medical Expenditures (log)



Notes: Sample includes individuals aged 50 to 69 in the 2014 and 2016 waves of the APIS.

Figure A2: Additional Observable Characteristics for Always Takers, Compliers, and Never Takers



Notes: Means and 95% confidence intervals are computed from 1000 bootstrapped re-samples using a 5-year bandwidth around age 60. Low and high education correspond to incomplete primary and complete secondary education, respectively. For the APIS, low and high SES correspond to the 1st to 3rd and 8th to 10th deciles of the national per capita income distribution, respectively. For the DHS, low and high SES correspond to the 1st-2nd and 5th quintiles of the wealth index distribution, respectively.

Table A1: Summary Statistics, APIS

	2014		2016	
	Age 50-59	Age 60-69	Age 50-59	Age 60-69
Primary member of PHIC (=1)	0.34	0.26	0.36	0.46
	(0.47)	(0.44)	(0.48)	(0.50)
Health expenditure per capita (pesos)	1378.99	1874.15	1075.33	1931.24
	(7369.51)	(9026.70)	(6088.83)	(6745.78)
Health expenditure share	0.03	0.04	0.03	0.04
	(0.07)	(0.08)	(0.07)	(0.09)
Inpatient exp. per capita (pesos)	519.77	592.38	424.17	482.58
	(6337.35)	(5071.41)	(4665.57)	(3806.97)
Outpatient exp. per capita (pesos)	140.21	169.57	100.65	284.58
	(1413.12)	(932.85)	(754.49)	(2895.45)
Medical products exp. per capita (pesos)	266.04	369.58	550.50	1164.08
	(1150.56)	(1424.41)	(1954.64)	(3646.46)
Male (=1)	0.49	0.47	0.49	0.49
	(0.50)	(0.50)	(0.50)	(0.50)
Education: Incomplete Primary (=1)	0.19	0.29	0.21	0.29
	(0.39)	(0.45)	(0.41)	(0.45)
Education: Complete Primary (=1)	0.35	0.38	0.31	0.32
	(0.48)	(0.48)	(0.46)	(0.47)
Education: Complete Secondary (=1)	0.46	0.34	0.48	0.39
	(0.50)	(0.47)	(0.50)	(0.49)
Low-income household (=1)	0.28	0.27	0.28	0.26
	(0.45)	(0.45)	(0.45)	(0.44)
Middle-income household (=1)	0.41	0.41	0.41	0.40
	(0.49)	(0.49)	(0.49)	(0.49)
High-income household (=1)	0.31	0.31	0.31	0.34
	(0.46)	(0.46)	(0.46)	(0.47)
N	4247	2756	3718	2395

Notes: Low, middle, and high income correspond to the 1st to 3rd, 4th to 7th, and 8th to 10th deciles of the national per capita income distribution, respectively.

Table A2: Summary Statistics, DHS

	2013		2017	
	Age 50-59	Age 60-69	Age 50-59	Age 60-69
Covered by PHIC (=1)	0.63	0.61	0.67	0.83
	(0.48)	(0.49)	(0.47)	(0.37)
Hospital stay last year (=1)	0.05	0.07	0.05	0.07
	(0.21)	(0.26)	(0.21)	(0.26)
Health visit last month (=1)	0.10	0.13	0.09	0.12
	(0.30)	(0.34)	(0.28)	(0.33)
Illness: Chronic condition (=1)	0.06	0.11	0.05	0.10
	(0.24)	(0.34)	(0.23)	(0.33)
Illness: Acute condition (=1)	0.12	0.14	0.09	0.11
	(0.33)	(0.35)	(0.29)	(0.32)
Male (=1)	0.49	0.46	0.49	0.47
	(0.50)	(0.50)	(0.50)	(0.50)
Education: Incomplete Primary (=1)	0.22	0.28	0.22	0.29
	(0.41)	(0.45)	(0.42)	(0.46)
Education: Complete Primary (=1)	0.31	0.35	0.29	0.33
	(0.46)	(0.48)	(0.46)	(0.47)
Education: Complete Secondary (=1)	0.47	0.37	0.48	0.38
	(0.50)	(0.48)	(0.50)	(0.48)
Low-wealth household (=1)	0.37	0.35	0.43	0.41
	(0.48)	(0.48)	(0.49)	(0.49)
Middle-wealth household (=1)	0.40	0.41	0.38	0.39
	(0.49)	(0.49)	(0.49)	(0.49)
High-wealth household (=1)	0.23	0.24	0.19	0.20
	(0.42)	(0.43)	(0.39)	(0.40)
N	6045	3365	11145	6801

Notes: Low, middle, and high wealth correspond to the 1st and 2nd, 3rd to 4th, and 5th quintiles of the wealth index distribution, respectively.

Figure A3: Age Distributions

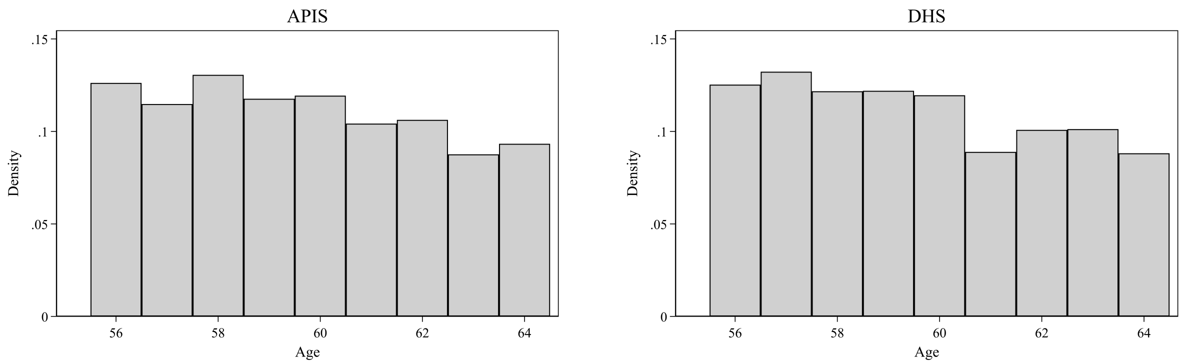
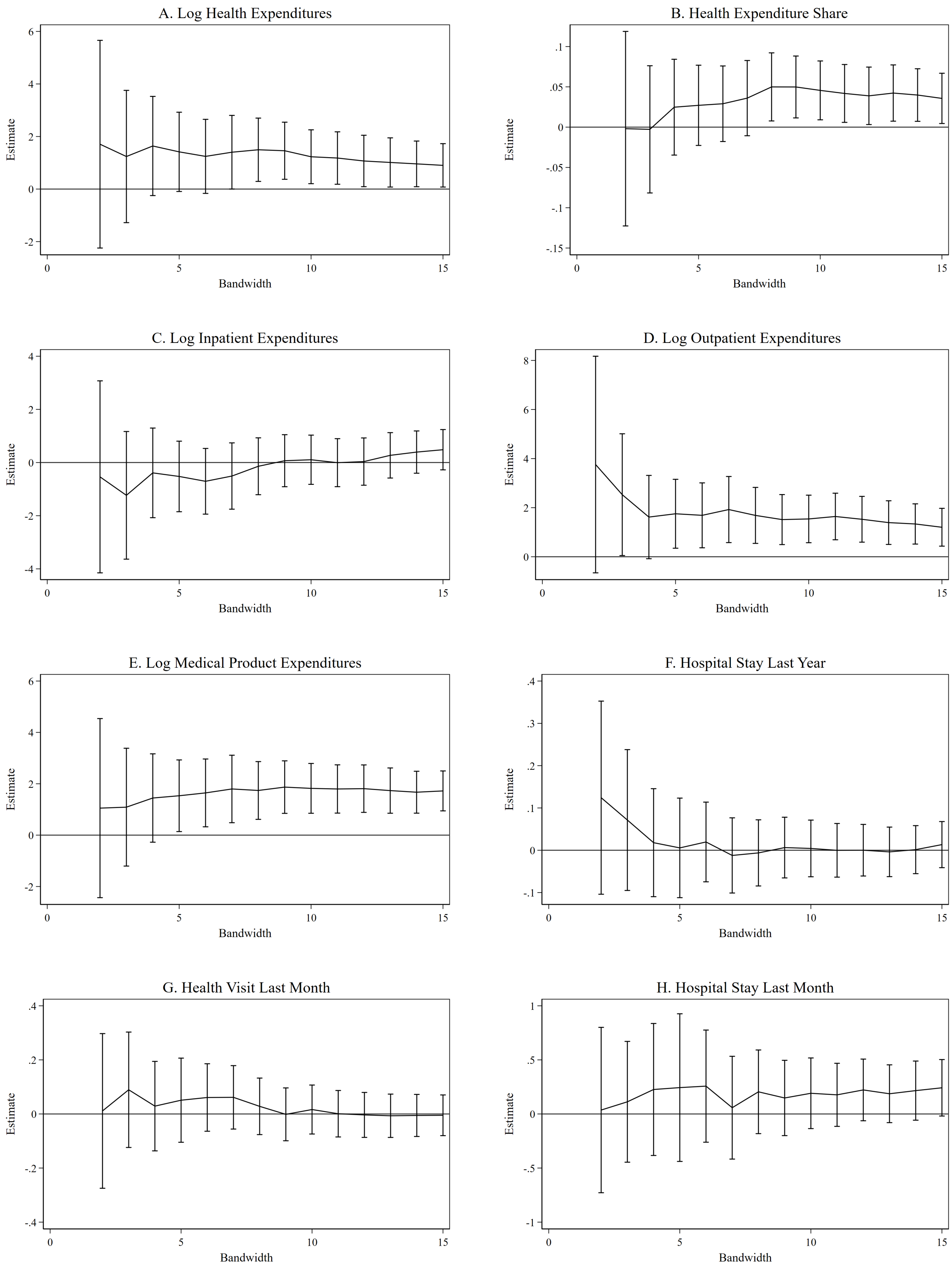


Table A3: Instrument Validity Tests

	APIS			DHS		
Bandwidth	2	5	10	2	5	10
Dependent Variable:						
Male	0.016	0.030	0.0068	-0.020	-0.021	0.013
	(0.038)	(0.025)	(0.018)	(0.028)	(0.018)	(0.013)
Low SES	0.022	-0.011	-0.011	-0.0032	0.00015	0.0034
	(0.033)	(0.022)	(0.016)	(0.028)	(0.018)	(0.013)
Middle SES	-0.035	-0.029	-0.019	0.0039	0.00031	0.0038
	(0.037)	(0.024)	(0.018)	(0.028)	(0.018)	(0.013)
Low Education	-0.026	-0.023	-0.024	-0.038	-0.00092	0.0010
	(0.033)	(0.021)	(0.015)	(0.025)	(0.016)	(0.011)
Middle Education	-0.031	-0.011	-0.0094	0.022	0.022	0.0042
	(0.036)	(0.023)	(0.017)	(0.027)	(0.017)	(0.012)
Polynomial Order	0	1	2	1	2	3
F-statistic	3.93	5.43	4.59	3.22	3.35	1.61
p-value	0.56	0.37	0.47	0.67	0.65	0.90
N	2811	6707	13116	5435	13650	27356

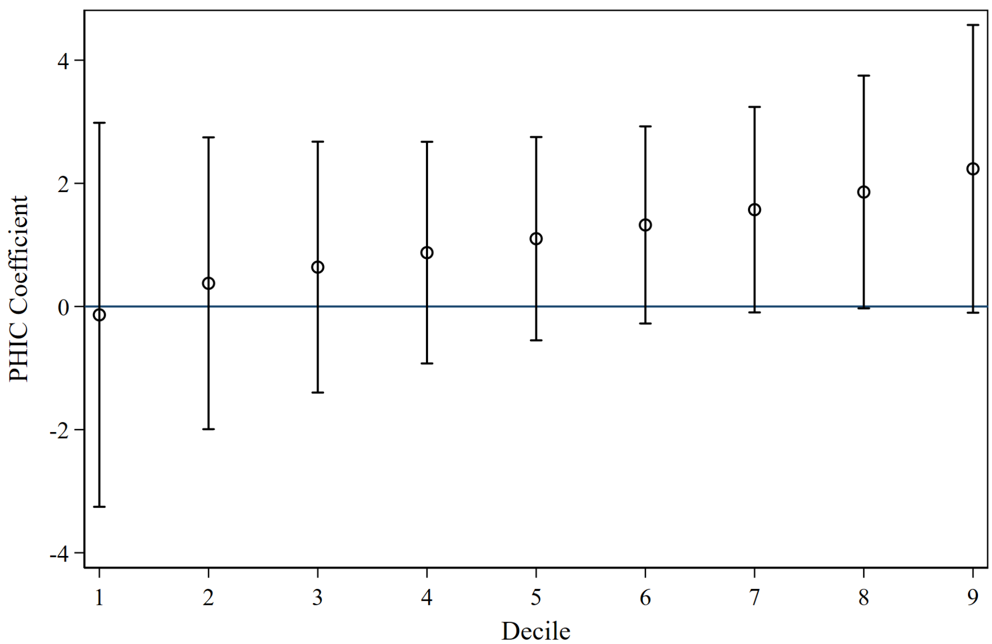
Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01$,$^{**}p < 0.05$,$^{*}p < 0.1$). Each cell represents a different regression, defined by the specified dependent variable, bandwidth, and polynomial order. We only report the coefficient (and standard error) for the $POST \times SENIOR$ interaction, but all regressions control for the main effects of $POST$, $SENIOR$, and a flexible polynomial for age that varies above and below the cutoff. In both datasets, low and middle education correspond to incomplete primary and complete primary, respectively. In the APIS, low and middle SES correspond to the 1st to 3rd and 4th to 7th deciles of the national per capita income distribution, respectively. In the DHS, low and middle SES correspond to the 1st to 2nd and 3rd to 4th quintiles of the wealth index distribution, respectively.

Figure A4: IV Estimates, Different Bandwidths



Notes: Each panel reports the coefficient on PHIC membership (and 95% confidence interval) from the IV specification in (3), for different bandwidths. Standard errors clustered at the household level. Panels A to E use the APIS, while panels F to H use the DHS.

Figure A5: Quantile Regression Coefficients



Notes: Each point represents the coefficient estimate (and 95% confidence interval) of the effect of PHIC membership on per capita medical expenditures from an IV quantile instrumental regression for the specified decile. All regressions use $POST \times SENIOR$ as the instrument and control for the main effects of $POST$, $SENIOR$, and a quadratic age polynomial.

Table A4: Health Expenditures Shares by Expenditure Type, APIS and FIES

	2014 APIS	2015 FIES	2016 APIS
Inpatient expenditure share	0.07	0.08	0.06
	(0.22)	(0.24)	(0.22)
Outpatient expenditure share	0.09	0.11	0.09
	(0.21)	(0.21)	(0.22)
General outpatient		0.07	
		(0.17)	
Other outpatient		0.04	
		(0.13)	
Medical product expenditure share	0.65	0.81	0.85
	(0.48)	(0.31)	(0.30)
Pharmaceutical products		0.78	
		(0.32)	
Other products		0.03	
		(0.11)	

Notes: Sample includes all households in the 2014 APIS, 2015 FIES, and 2016 APIS. General outpatient services include consultations, physical check-ups, and laboratory services. Other outpatient services include specialized medical services (analysis and interpretation of X-rays, etc.), dental services, and paramedical services (freelance acupuncturists, optometrists, etc.). Other medical products include thermometers, bandages, corrective eye glasses, dentures, etc.

Table A5: IV Estimates: Effect of Health Insurance on Expenditures (APIS), Restricted Sample

	Log Health Expenditures		Health Expenditure Share		Log Inpatient Expenditures		Log Outpatient Expenditures		Log Medical Product Exp.	
Bandwidth	5	10	5	10	5	10	5	10	5	10
Enrolled in PHIC	1.44	1.13	0.054	0.073	-0.020	0.48	1.60	1.73	1.19	1.13
	(0.87)*	(0.59)*	(0.022)**	(0.017)***	(0.56)	(0.39)	(0.67)**	(0.48)***	(0.80)	(0.55)**
	[0.21]	[0.09]	[0.07]	[0.00]	[0.96]	[0.24]	[0.01]	[0.00]	[0.26]	[0.08]
Weak identification F	37.4	84.9	37.4	84.9	37.4	84.9	37.4	84.9	37.4	84.9
Underidentification F	37.0	83.9	37.0	83.9	37.0	83.9	37.0	83.9	37.0	83.9
Polynomial Order	1	2	1	2	1	2	1	2	1	2
Mean of Dep. Var.	4.60	4.53	0.027	0.026	0.38	0.37	0.73	0.75	4.53	4.46
N	5584	10860	5584	10860	5584	10860	5584	10860	5584	10860

Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. Each cell represents a different regression, defined by the specified dependent variable, bandwidth, and polynomial order. We report the results of instrumental variables regressions, with 1(Enrolled in PHIC) as the endogenous variable of interest, and $POST \times SENIOR$ as our instrument. All regressions control for the main effects of $POST$, $SENIOR$, and a flexible polynomial for age that varies above and below the cutoff. All regressions exclude individuals from households whose total health expenditures exceed the sum of their inpatient, outpatient, and medical product expenditures.

Table A6: Quantile IV Estimates: Effect of Health Insurance on Log Expenditure Deciles

	Decile								
	1	2	3	4	5	6	7	8	9
Enrolled in PHIC	-0.135	0.377	0.639	0.875	1.101	1.324	1.572*	1.860*	2.236*
	(1.591)	(1.209)	(1.040)	(0.918)	(0.842)	(0.817)	(0.851)	(0.964)	(1.192)
Polynomial Order	2	2	2	2	2	2	2	2	2
Mean of Dep. Var.	4.94	4.94	4.94	4.94	4.94	4.94	4.94	4.94	4.94
N	13116	13116	13116	13116	13116	13116	13116	13116	13116

Notes: Standard errors reported in parentheses ($^{***}p < 0.01,^{**} p < 0.05,^{*} p < 0.1$). All regressions use the APIS. We report the results of quantile instrumental variables regressions on log per capita medical expenditures, with 1(Enrolled in PHIC) as the endogenous variable of interest, and $POST \times SENIOR$ as our instrument. All regressions control for the main effects of $POST$, $SENIOR$, and a quadratic age polynomial.

Table A7: Reduced Form Estimates: Effect of Policy Change on Expenditures (APIS)

	Log Health Expenditures		Health Expenditure Share		Log Inpatient Expenditures		Log Outpatient Expenditures		Log Medical Product Exp.	
Bandwidth	5	10	5	10	5	10	5	10	5	10
A. Baseline Specification (Age Polynomials)										
Senior x Post	0.22	0.22	0.0042	0.0081	-0.082	0.019	0.27	0.27	0.24	0.32
	(0.12)*	(0.092)**	(0.0040)	(0.0033)**	(0.10)	(0.084)	(0.11)***	(0.084)***	(0.11)**	(0.083)***
	[0.22]	[0.06]	[0.39]	[0.03]	[0.47]	[0.83]	[0.01]	[0.00]	[0.19]	[0.02]
Polynomial Order	1	2	1	2	1	2	1	2	1	2
B. Age Fixed Effects										
Senior x Post	0.23	0.23	0.0044	0.0082	-0.086	0.015	0.28	0.28	0.25	0.33
	(0.12)**	(0.092)**	(0.0040)	(0.0033)**	(0.11)	(0.084)	(0.11)***	(0.085)***	(0.11)**	(0.083)***
	[0.22]	[0.06]	[0.38]	[0.03]	[0.47]	[0.87]	[0.01]	[0.00]	[0.18]	[0.02]
Mean of Dep. Var.	4.99	4.94	0.036	0.036	0.62	0.62	1.23	1.25	3.77	3.69
N	6707	13116	6707	13116	6707	13116	6707	13116	6707	13116

Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. Each cell represents a different regression, defined by the specified dependent variable, bandwidth, and polynomial order. All regressions control for $POST$ and a data quality indicator equal to one for individuals from households whose total health expenditures exceed the sum of their inpatient, outpatient, and medical product expenditures. Panel A additionally controls for $SENIOR$ and a flexible polynomial for age that varies above and below the cutoff, while Panel B controls for age fixed effects.

Table A8: Reduced Form Estimates: Effect of Policy Change on Utilization (DHS)

	Hospital Stay Last Year		Health Visit Last Month		Hospital Stay Last Month		
Bandwidth	5	10	5	10	5	10	15
A. Baseline Specification (Age Polynomials)							
Senior x Post	0.00082	0.00081	0.0075	0.0031	0.026	0.031	0.041
	(0.0088)	(0.0064)	(0.012)	(0.0087)	(0.035)	(0.026)	(0.021)*
	[0.91]	[0.88]	[0.38]	[0.74]	[0.15]	[0.25]	[0.05]
Polynomial Order	2	2	2	2	2	2	2
B. Age Fixed Effects							
Senior x Post	0.0011	0.00076	0.0077	0.0028	0.025	0.029	0.041
	(0.0089)	(0.0064)	(0.012)	(0.0087)	(0.035)	(0.026)	(0.022)*
	[0.91]	[0.88]	[0.38]	[0.74]	[0.15]	[0.25]	[0.05]
Mean of Dep. Var.	0.060	0.056	0.11	0.10	0.13	0.14	0.14
N	13650	27356	13650	27356	1469	2827	4081

Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. Each cell represents a different regression, defined by the specified dependent variable, bandwidth, and polynomial order. All regressions control for $POST$. Panel A additionally controls for $SENIOR$ and a flexible polynomial for age that varies above and below the cutoff, while Panel B controls for age fixed effects. The Hospital Stay Last Month variable is only available for those with Health Visit Last Month equal to one.

Table A9: IV Estimates: Effect of Health Insurance on Expenditures (APIS), Age Fixed Effects

	Log Health Expenditures		Health Expenditure Share		Log Inpatient Expenditures		Log Outpatient Expenditures		Log Medical Product Exp.	
Bandwidth	5	10	5	10	5	10	5	10	5	10
Enrolled in PHIC	1.50	1.27	0.028	0.046	-0.55	0.084	1.77	1.56	1.62	1.86
	(0.82)*	(0.50)**	(0.023)	(0.017)***	(0.61)	(0.49)	(0.55)***	(0.33)***	(0.73)**	(0.44)***
	[0.16]	[0.04]	[0.30]	[0.02]	[0.46]	[0.87]	[0.01]	[0.00]	[0.11]	[0.01]
Weak identification F	22.1	48.2	22.1	48.2	22.1	48.2	22.1	48.2	22.1	48.2
Underidentification F	7.55	15.3	7.55	15.3	7.55	15.3	7.55	15.3	7.55	15.3
Mean of Dep. Var.	4.99	4.94	0.036	0.036	0.62	0.62	1.23	1.25	3.77	3.69
N	6707	13116	6707	13116	6707	13116	6707	13116	6707	13116

Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. Each cell represents a different regression, defined by the specified dependent variable and bandwidth. We report the results of instrumental variables regressions, with 1(Enrolled in PHIC) as the endogenous variable of interest, and $POST \times SENIOR$ as our instrument. All regressions control for $POST$, age fixed effects, and data quality indicator equal to one for individuals from households whose total health expenditures exceed the sum of their inpatient, outpatient, and medical product expenditures.

Table A10: IV Estimates: Effect of Health Insurance on Utilization (DHS), Age Fixed Effects

	Hospital Stay Last Year		Health Visit Last Month		Hospital Stay Last Month		
Bandwidth	5	10	5	10	5	10	15
Enrolled in PHIC	0.0075	0.0041	0.052	0.015	0.24	0.18	0.24
	(0.039)	(0.025)	(0.052)	(0.045)	(0.17)	(0.15)	(0.12)**
	[0.87]	[0.88]	[0.39]	[0.76]	[0.13]	[0.29]	[0.05]
Weak identification F	48.7	76.8	48.7	76.8	4.28	21.6	40.2
Underidentification F	8.51	16.4	8.51	16.4	3.23	10.9	18.0
Mean of Dep. Var.	0.060	0.056	0.11	0.10	0.13	0.14	0.14
N	13650	27356	13650	27356	1469	2827	4081

Notes: Standard errors, clustered at the household level, reported in parentheses ($^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$). Wild cluster bootstrap p-values from clustering by age in square brackets. Each cell represents a different regression, defined by the specified dependent variable and bandwidth. We report the results of instrumental variables regressions, with 1(Enrolled in PHIC) as the endogenous variable of interest, and $POST \times SENIOR$ as our instrument. All regressions control for $POST$ and age fixed effects. The Hospital Stay Last Month variable is only available for those with Health Visit Last Month equal to one.

A.2 Average Complier Characteristics Derivations

We derive the formulas given in equations (4) and (5). We will exploit the fact that $Z = D$ for $U \in [p_0, p_1]$. In addition, we will assume that $Z \perp (U, X)$. First, we note that

$$
\begin{aligned}
&\mu_X(0) = \\
&E\left[X|D = 0, p_0 \leq U < p_1\right] = \\
&E\left[X|Z = 0, p_0 \leq U < p_1\right] = \\
&\frac{1}{p_1 - p_0}\left[(1 - p_0)E[X|p_0 \leq U \leq 1] - (1 - p_1)E[X|p_1 \leq U \leq 1]\right] = \\
&\frac{1}{p_1 - p_0}\left[(1 - p_0)E[X|Z = 0, p_0 \leq U \leq 1] - (1 - p_1)E[X|Z = 1, p_1 \leq U \leq 1]\right] = \\
&\frac{1}{p_1 - p_0}\left[(1 - p_0)E[X|D = 0, Z = 0] - (1 - p_1)E[X|D = 0, Z = 1]\right].
\end{aligned}
$$

Per Figure 1, the second term nets out the contribution of the never takers from the composite of the untreated compliers and never takers. Similar arguments deliver that

$$
\begin{aligned}
&\mu_X(1) = \\
&E\left[X|D = 1, p_0 \leq U < p_1\right] = \\
&E\left[X|Z = 1, p_0 \leq U < p_1\right] = \\
&\frac{1}{p_1 - p_0}\left[p_1 E[X|0 \leq U < p_1] - p_0 E[X|0 \leq U < p_0]\right] = \\
&\frac{1}{p_1 - p_0}\left[p_1 E[X|Z = 1, 0 \leq U < p_1] - p_0 E[X|Z = 0, 0 \leq U < p_0]\right] = \\
&\frac{1}{p_1 - p_0}\left[p_1 E[X|D = 1, Z = 1] - p_0 E[X|D = 1, Z = 0]\right]
\end{aligned}
$$

where, similar to above, the second term nets out the effects of the always takers from the first term which is a composite of the treated compliers and always takers (per Figure 1).

A.3 Quantile IV Regressions

In this section, we investigate the effects of PHIC membership on medical expenditures at various points in its distribution, using a quantile IV estimator developed by Chernozhukov and Hansen (2008). We index the quantile of the medical expenditure distribution with τ. This procedure delivers $\sqrt{N}$-consistent estimates of the parameters of the model

$$S_Y(\tau|D, X) = \alpha(\tau)D + X'\beta(\tau) + f_\tau(a - 60)$$

where $X = [SENIOR, POST]'$ and the function $S_Y(\tau|D, X)$ is what Chernozhukov and Hansen (2008) refer to as the *structural quantile function* or SQF.

The interpretation of the SQF is that it describes the quantile function of a latent variable $Y_d = \alpha(U)d + X\beta(U)$ where the treatment is fixed at $D = d$ and U is sampled from a uniform on $[0, 1]$ conditional on X. The treatment variable, which is PHIC membership in our case, is potentially correlated with U. Identification of $\alpha(\tau)$ requires an IV that impacts PHIC membership, is independent of U, but satisfies an exclusion restriction. Accordingly, we re-purpose the same IV that we used for estimation of equation (3), $Z_{at} = POST \times SENIOR$, for the identification of the parameters of the SQF.

Quantile IV estimation is useful because it can shed light on whether the impacts of the PHIC were at the extensive or intensive margins. If insurance is changing the intensive but not the extensive margin of utilization, we should see movements in the upper end of the expenditure distribution, not at the lower end. That is, we should see higher spending as a consequence of PHIC membership among individuals who were already spending on healthcare as opposed to individuals who were not spending at all. That said, we also note that movement at the upper end of the expenditure distribution and not the lower end could also be indicative of individuals going from very low to very high spending as a result of PHIC.

Related, quantile regression is also useful because it provides another way of addressing the presence of many zeros in medical expenditure data. Recall that in Figure A1, close to 10% of the medical expenditure observations were zeros. Hence, estimation of the model for $\tau = 0.1$ is informative of the effects of the PHIC at the extensive margins, whereas estimation for $\tau > 0.1$ is informative of the intensive margins. Finally, note that the commonly used two-part model in health economics discussed in Mullahy (1998) cannot be modified to handle endogenous regressors without relatively stringent parametric assumptions.

In Table A6, we present the estimations of $\alpha(\tau)$ from the SQF for $\tau \in \{0.1, 0.2, ..., 0.9\}$, estimated using the methods proposed by Machado and Santos Silva (2018). In Figure A5, we also plot the estimates from the table along with their 95% confidence intervals. It is clear that the effect of insurance is increasing across deciles, with magnitudes higher than unity between the 5th to 9th deciles. In fact, the effects are only significant (at the 10% level) for the top 3 deciles of the expenditures distribution, implying more movement at the upper end of the distribution.

A.4 Robustness

We begin by reporting reduced form regressions, in which we directly regress each outcome on $SENIOR \times POST$, our instrument and policy variable of interest. These regressions provide estimates of the effect of the policy, which expanded eligibility for seniors, on our outcomes of interest. Consistent with our IV estimates, we find the policy increased OOP health expenditures by 22% (with larger increases for outpatient and medical products), and significantly increased the likelihood of a hospital stay conditional on having a health visit.

Next, we replace the age polynomials in our original specification with age fixed effects (which results in a standard generalized difference-in-differences framework) and repeat our IV regressions. Results, reported in Tables A9 and A10, are unchanged. Reduced form results using this specification are reported in Panel B of Tables A7 and A8.

One important question relevant to the interpretation of our coefficients is whether people choose to get insurance, and then make their expenditures and utilization decisions, or whether they make their utilization decisions and get enrolled (at point-of-service) as a result. In the most extreme case, suppose seniors who know they are covered due to the ESCA amendment increase their utilization because of this, and only get insured when they go to the doctor. We would see a one-to-one relationship between insurance and utilization and would not be able to interpret our IV coefficients as a LATE. There are two main reasons why we think this situation is unlikely. First, as mentioned above, point-of-service enrollment is very uncommon (for seniors and in general). Second, we do not find any effects of insurance coverage on extensive margin utilization (see Table 4). This means that people are choosing to enroll even if they are not planning to utilize care, suggesting that there are two separate decisions (as modeled in section 4.1): the decision to sign up for insurance, and the decision to utilize care.

Even if this two-step decision sequence does not apply to everyone in our sample, the reduced form estimates in Tables A7 and A8 still provide valid estimates of the effect of the policy change. That is, regardless of whether insurance enrollment comes before utilization, our reduced form results demonstrate that becoming eligible for health insurance led seniors to spend more on healthcare, specifically outpatient care and drugs.

A.5 Marginal Treatment Effect: Technical Details

Following Brinch et al. (2017), we express the expected values of Y_j (conditional on $U = p$) as the sum of a systematic component (denoted by μ_j) and the expectation of an unobserved component (denoted by $k_j(p)$), that is: $\mu_j + k_j(p) \equiv E[Y_j|U = p]$ for $j \in \{0, 1\}$. The MTE can then be expressed as $MTE(p) = (\mu_1 - \mu_0) + (k_1(p) - k_0(p))$. We assume that the expected unobserved components of the outcomes, $k_j(p)$, are linear in p.[30] Specifically, we assume that

$$k_j(p) = \alpha_j p - \frac{1}{2}\alpha_j \text{ for } j \in \{1, 0\}.$$

Note that this formulation guarantees that the idiosyncratic components of Y_1 and Y_0 have mean zero. The key insight of Brinch et al. (2017) is that if $MTE(p) = \mu_1 - \mu_0$ and therefore it is constant in p, then there is no treatment effect heterogeneity. The LATE will then be equally applicable to the always takers and never takers. This requires $\alpha_1 = \alpha_0$. Therefore, given our linearity assumption, the LATE will be externally valid if and only if $k_1(p_1) - k_1(p_0) = k_0(p_1) - k_0(p_0)$.[31] This, in turn, implies that

$$E[Y_1|U \le p_0] - E[Y_0|U > p_0] = E[Y_1|U \le p_1] - E[Y_0|U > p_1]. \tag{7}$$

Note that all four of these expectations can be estimated in the data.[32] Accordingly, we can re-write equation (7) as

$$E[Y|D = 1, Z = 0] - E[Y|D = 0, Z = 0] = E[Y|D = 1, Z = 1] - E[Y|D = 0, Z = 1]. \tag{8}$$

To conduct this test, we estimate the following difference-in-difference regression, using total OOP expenditures as our outcome:[33]

$$Y_{iat} = \phi_0 + \phi_1 D_{iat} + \phi_2 Z_{iat} + \phi_3 D_{iat} \times Z_{iat} + \upsilon_{iat}.$$

Under the null hypothesis of no treatment effect heterogeneity, $\phi_3 = 0$.

[30] Note that with only a binary IV, a model that is quadratic in p is not identified. Identification of the quadratic model requires additional information. Brinch et al. (2017) show that the addition of a binary X achieves identification of the quadratic model. We do not explore a quadratic model in this paper for the sake of parsimony. We leave this for future explorations.

[31] We note that much of this is not obvious *prima facie* and, so we encourage interested readers to carefully read Section III of Brinch et al. (2017).

[32] Figure 1 shows how the expectations in (7) can be mapped to the expectations in (8).

[33] Recall that D_{iat} represents treatment, which is insurance coverage in our context, and Z_{at} represents the instrument, which is the $SENIOR_a \times POST_t$ interaction term. We estimate this regression and conduct the analysis that follows after partialling out our control variables (linear age functions that vary above and below the cutoff, a year dummy, a senior dummy, and a flag for inconsistent expenditure totals).

In addition to testing for the existence of treatment effect heterogeneity, this regression also allows us to identify the parameters α_j and μ_j for $j \in \{0,1\}$. Following Brinch et al. (2017), we note that given the linearity assumption on $k_j(p)$ for $j \in \{T, U\}$, we will have that

$$
\begin{aligned}
E\left(Y|P(Z)=p, D=0\right) &= \mu_0 + \frac{1}{2}\alpha_0 p \\
E\left(Y|P(Z)=p, D=1\right) &= \mu_1 + \frac{1}{2}\alpha_1 (p-1).
\end{aligned}
$$

Matching these moments with the coefficients from the difference-in-differences regression, we obtain

$$
\begin{aligned}
E\left(Y|Z=0, D=0\right) &= \phi_0 = \mu_0 + \frac{1}{2}\alpha_0 p_0 \\
E\left(Y|Z=1, D=0\right) &= \phi_0 + \phi_2 = \mu_0 + \frac{1}{2}\alpha_0 p_1 \\
E\left(Y|Z=0, D=1\right) &= \phi_0 + \phi_1 = \mu_1 + \frac{1}{2}\alpha_1 (p_0 - 1) \\
E\left(Y|Z=1, D=1\right) &= \phi_0 + \phi_1 + \phi_2 + \phi_3 = \mu_1 + \frac{1}{2}\alpha_1 (p_1 - 1).
\end{aligned}
$$

These four moment conditions identify μ_j and α_j for $j \in \{0,1\}$ as follows:

$$
\begin{aligned}
\alpha_0 &= \frac{2\phi_2}{p_1 - p_0} & \mu_0 &= \phi_0 - \frac{p_0 \phi_2}{p_1 - p_0} \\
\alpha_1 &= \frac{2(\phi_2 + \phi_3)}{p_1 - p_0} & \mu_1 &= \phi_0 + \phi_1 - \frac{(\phi_2 + \phi_3)(p_0 - 1)}{p_1 - p_0}.
\end{aligned}
$$

These estimates allow us to identify the treatment effects for the always takers, compliers, and never takers as

$$
\begin{aligned}
E\left[\Delta|0 \leq U \leq p_0\right] &= \mu_1 - \mu_0 + (\alpha_1 - \alpha_0)\left(\frac{p_0 - 1}{2}\right) && \text{Always Takers} \\
E\left[\Delta|p_0 \leq U \leq p_1\right] &= \mu_1 - \mu_0 + (\alpha_1 - \alpha_0)\left(\frac{p_0 + p_1 - 1}{2}\right) && \text{Compliers} \\
E\left[\Delta|p_1 \leq U \leq 1\right] &= \mu_1 - \mu_0 + (\alpha_1 - \alpha_0)\left(\frac{p_1}{2}\right). && \text{Never Takers}
\end{aligned}
$$

Accordingly, the linearity assumption on the marginal treatment effect allows us to estimate all relevant treatment effects using a simple difference-in-difference model.

References

Brinch, C. N., Mogstad, M., and Wiswall, M. (2017). Beyond late with a discrete instrument. *Journal of Political Economy*, 125(4):985–1039.

Chernozhukov, V. and Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics*, 142(1):379–398.

Frandsen, B. R. (2017). Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete. In *Regression discontinuity designs*. Emerald Publishing Limited.

Machado, J. A. and Santos Silva, J. (2018). Quantiles via moments. *Journal of Econometrics*.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Mullahy, J. (1998). Much ado about two: reconsidering retransformation and the two-part model in health econometrics. *Journal of health economics*, 17(3):247–281.