N° 10
October 2012

Impact Analyses Series

ExPost

Adverse Selection Based on Observable and Unobservable Factors in Health Insurance in Rural Cambodia

Rachel Polimeni, Center of Evaluation for Global Action, University of California, Berkeley
David I. Levine, Haas School of Business, University of California, Berkeley

Contact: Stéphanie Pamies, Evaluation and Capitalisation Division, AFD

Research Department
Evaluation and Capitalisation Division
Agence Française de Développement
5, rue Roland Barthes 75012 Paris - France
www.afd.fr

AGENCE FRANÇAISE DE DÉVELOPPEMENT

Disclaimer

The analysis and conclusions of this document are those of the authors. They do not necessarily reflect the official position of the AFD or its partner institutions.

Publications Director: Dov ZERAH
Editorial Director: Laurent FONTAINE
ISSN: 2101-9657
Legal Deposit: 4th Quarter 2012

Layout: Marcelle LARNICOL

Acknowledgements

Rachel Polimeni is a researcher at the Center of Evaluation for Global Action and David Levine is a professor at Haas School of Business, both at the University of California, Berkeley.

We acknowledge the AFD, USAID, and the Coleman Fung Foundation for their generous funding. Cooperation from GRET and SKY was essential in implementing this study. We thank the staff at GRET for sharing their data and the field team at Domrei for their tireless data collection and cleaning. We worked closely with Ian Ramage of Domrei on the survey design and implementation of the SKY evaluation. Jean-David Naudet, Jocelyne Delarue, and Stephanie Pamies of the AFD gave us enormous guidance in the course of the evaluation and valuable feedback on the resulting papers. Rachel Gardner and Francine Anene provided excellent research assistance. Raj Arunachalam was an essential part of the early phases of the evaluation. We appreciate comments at seminars at USAID BASIS, UC Berkeley, CERDI and other presentations, and from Ted Miguel, Paul Gertler, and other colleagues and stakeholders.

CONTENTS

Introduction

Health insurance can increase access to health care and decrease the potentially catastrophic effects of large medical expenses. However, asymmetric information in health insurance markets may stop voluntary health insurance markets from providing protection to most consumers without substantial regulation or subsidization.

Adverse selection is one type of asymmetric information that can raise an insurer's costs and reduce coverage. The economic theory of insurance demand predicts that households that anticipate high health care costs are those that are most likely to be willing to purchase health insurance (Rothschild and Stiglitz, 1976; Akerlof, 1970). Voluntary health insurance cannot be financially sustainable if adverse selection is severe because only the most costly patients would find it worthwhile to purchase insurance, and insurers' premium levels would not be able to cover the high costs of care.

Results of empirical studies of the importance of adverse selection remain mixed. In addition, in many empirical studies the effects of adverse selection are difficult to disentangle from the effects of moral hazard. Because adverse selection and moral hazard have very different implications for insurers and policy-makers, it is important to understand which aspects of asymmetric information are important in different settings.

We utilize a randomized experimental design that allows us to separate adverse selection from moral hazard. The experiment was carried out and data were collected during the expansion of the SKY health microinsurance program in rural Cambodia. SKY partners with public health facilities and provides free care in exchange for a small (subsidized) monthly premium.

As SKY expanded, we distributed coupons for deeply discounted insurance to a random subset of potential SKY customers, inducing a sharp increase in take-up rate. Surveys from over 5,000 families provide information on health and health care utilization both before and after these households had the opportunity to purchase SKY. SKY administrative data provides further information on members' health care utilization.

We use these data to test three implications of theories of asymmetric information. Our first hypothesis (H1) is that those with high expected health care costs (based on factors observable to the econometrician such as past health care utilization) will be more likely to buy insurance. We call this selection on observables. Hypothesis 2 (H2) is that insurance buyers paying full price have higher predicted health care expenditures than buyers paying a deeply discounted price. Our final hypothesis (H3) is that among the insured, those who bought insurance at full price will have higher health care utilization after they are insured than those who bought insurance at a deeply discounted price, holding constant observable factors. We call this selection on unobservables.

Consistent with Hypothesis 1 (adverse selection based on observables), while 67% of households that declined SKY report having at least one member in poor health, an even higher share (80%) of buyers report a member in poor health. Similarly, buyers are 2.3 percentage points (33%) more likely than decliners to report major health shock in the household during the three months preceding the introduction of SKY. Contrary to hypothesis 1, households with young or old members are not more likely to buy SKY, nor are those with a stunted or wasted child.

Contrary to Hypothesis 2 (worse adverse selection on observables at higher prices), we find only weak evidence that households that paid a higher price for insurance are more likely to report a member in poor health or to have experienced a serious health shock prior to the purchase of SKY than households that paid a deeply discounted price for insurance.

Consistent with Hypothesis 3, households that paid more for health insurance have substantially higher usage of both health centers and hospitals than households that received a discounted price, even when comparing households with similar observed baseline health. This result is consistent with substantial adverse selection based on factors we did not observe at the baseline.

Tests of our hypotheses are often difficult in societies with well-established health insurance markets. First, moral hazard due to previous insurance coverage may have increased past health care utilization for some customers, so it is hard to tell whether prior health care utilization predicts purchase of insurance or if past utilization was due to being insured. In this study, no health insurance was available in these communities prior to SKY.

In addition, in other settings high insurance prices often imply higher benefits such as lower copayments. In that case, moral hazard (in addition to adverse selection) can lead to higher health care utilization for those who pay high prices. In this study, benefits were identical regardless of price.[1]

Third, in other settings insurance prices may correlate with the quality of care or with factors the insurance company observes that are not observed by the econometrician. In this study, we randomized prices so they should not be correlated systematically with such factors.

Studying health care in poor nations in particular is important because expanding voluntary health insurance is a popular policy prescription for the billions of poor people who lack affordable access to health care. Most importantly, China has largely shifted to voluntary, government-subsidized health insurance for medical care coverage in rural areas (Wang, 2007; People's Daily Online, 2008). Other developing nations such as Vietnam and Thailand also have rapidly expanded health insurance coverage (Vietnam Social Security, 2010; Antos, 2007). To understand the financial viability of voluntary insurance in these settings, it is important to study selection among poor populations.

The success of voluntary health insurance in poor nations depends on the ability to improve health and economic outcomes while staying financially sustainable, or at the least assuring donors that their money is being spent in the most efficient way possible. However, because health insurance is a relatively new product in developing countries, little is known about the risks and benefits of offering this type of insurance, or how best to design an insurance program to meet the needs of the poor. This study provides evidence as to the extent of adverse selection in a developing country.

[1] Households who pay more for insurance may want to "get their money's worth". We discuss this possibility under robustness checks.

Previous Research

George Akerlof's seminal article "The Market for 'Lemons'" (Akerlof, 1970) examines why health insurance companies do not raise their rates to match the risk of clients.[2] Akerlof theorizes that individuals who are willing to pay the highest insurance premium are those people who expect the highest expected insurance payouts. For health insurance, poor health is a primary determinant of high expected insurance payouts. Individuals seeking insurance typically have more information on their health status than an insurance company.

If an insurance company cannot identify health status, it can only charge a price based on observable characteristics. To be financially sustainable, that price must cover the expected health care costs for this group. Because healthier individuals have lower expected health care expenses, they may not be willing to pay the higher premium that is a consequence of risk pooling with less healthy individuals. Under certain assumptions, the healthiest people may drop insurance first, leading to a worse risk pool among the insured and higher prices. Higher prices, in turn, drive out the next most healthy until the market collapses altogether; this process is sometimes referred to as a "death spiral".

There is extensive empirical literature on the extent of adverse selection in insurance markets in developed countries. Studies are of various types, comparing individual characteristics by generosity of health plan, premium level and choice of whether or not to remain uninsured (see Cutler and Zeckhauser, 2000, for a review).

Most studies have found evidence of adverse selection, although estimates of the magnitude vary. For example, several studies find that people with higher expected medical expenditures are more likely to buy insurance or choose health insurance with more generous benefits than those with lower expected medical expenditures (Cutler and Zeckhauser, 1998; Cutler and Reber, 1998). However, other studies find that adverse selection in health insurance and other insurance markets is minimal (Wolfe and Goddeeris, 1991; Finkelstein and Poterba, 2004) or non-existent (Finkelstein and McGarry, 2006; Cardon and Hendel, 2001; Cawley and Philipson, 1999). In fact, there is even some recent evidence of positive selection in health insurance (Fang, Keane and Silverman, 2008).

However, existing studies of asymmetric information often have trouble distinguishing the effects of adverse selection from moral hazard. For example, an observed cross-sectional correlation between generous health insurance and high health care utilization could be because those who anticipate high health care utilization choose generous insurance (adverse selection) or because generous insurance induces more health care utilization (moral hazard).

To focus on adverse selection when choosing among health care plans, Ellis (1989) examines a change in health insurance options at a financial firm. Out-of-pocket costs in the previous year among those selecting the most generous (and highest premium) plan were 8.6 times higher than for those selecting the least generous (and lowest premium) plan, a result strongly supporting adverse selection.

Cutler and Reber (1998) examine a similar natural experiment. In 1995, Harvard greatly increased the cost to most of its employees of its more generous health plans. 3,000 of its 10,000 employees were not subject to the price increase until a year later, providing a comparison group. Many employees who faced the higher prices switched to a lower-cost plan, but there was less switching among older employees and for those with high health care utilization in the past. There was enough of this adverse selection that by the third year after the reform the more generous plan was eliminated as an option for employees because it could no longer remain profitable at a reasonable premium.

[2] This literature review draws on Polimeni (2006), and Levine, Gardner and Polimeni (2009).

The above studies are all from developed nations. There have been far fewer studies of selection in developing countries, partly because there are far fewer insurance markets in developing countries. On the one hand, the theory of adverse selection implies potential customers with higher expected health care expenses will be more likely to buy health insurance in poor nations, just as in rich ones. On the other hand, there are several reasons to believe that clients in developing countries may behave differently than what has been found in developing countries. For example, because insurance is a relatively new and unknown product, only those who are willing to take a risk on a new product may be willing to try it. More generally, Siegelman (2004) has argued there is plausibly much more adverse selection when people choose among insurance plans than when they choose whether or not to become insured.

Non-experimental studies from developing countries have found enrollment to be more common in households with chronically sick members, consistent with adverse selection (Wagstaff, Lindelow, Jun, Ling and Juncheng, 2009). These studies also typically find higher enrollment rates among wealthier households, potentially leading to positive selection because wealthier people also tend to be healthier (Wagstaff, Lindelow, Jun, Ling and Juncheng, 2009; Wagstaff and Pradhan, 2005; Jutting, 2004; Lamiraud, Booysen and Scheil-Adlung, 2005).

The current study is most similar in design to Ausubel (1999) and Karlan and Zinman (2009), who randomize offers in credit markets. Ausubel (1999) finds evidence for three hypothesized consequences of adverse selection in his study of who accepts credit card offers, each corresponding to our three focal hypotheses. First, people who accept credit card offers have worse observable credit characteristics (in his setting, credit histories) than those who reject the offers. Second, he hypothesizes that people with worse credit histories are more likely to accept offers with high interest rates than those with good credit histories. Third, after controlling for observable characteristics of clients, those who accept offers with high interest rates are more likely to default. Ausubel's finding of differences in default rates may be due to adverse selection, but it may also be due to moral hazard: households with inferior credit offers have more incentive to default.

Ausubel argues that moral hazard is an unlikely explanation because default rates remain higher among households that accept the high interest rate credit card offers even after interest rates equalize over time. Karlan and Zinman (2009) use a similar experiment but are able to disentangle the effects of adverse selection and moral hazard by giving some borrowers who accepted a high interest rate a lower actual interest rate. They find only weak evidence of adverse selection.

The research presented here adds to the literature in several ways. First, we present evidence on adverse selection in a developing country, while empirical studies have taken place for the most part in developed countries. Second, the above empirical studies have taken place in more traditional competitive markets, whereas the SKY program in Cambodia is the only health insurance option in the rural markets targeted. Because there is no plan choice, adverse selection may show up more starkly in this market, or, as Siegelman (2004) argued, there may be less adverse selection. In addition, in our setting, longitudinal data and ability to randomize prices make it possible to identify selection effects in this dataset separately from moral hazard. Finally, unlike previous studies of health insurance, we measure selection based on both observable and unobservable factors.

1. Theory and Methods

1.1. Selection on Observables

The core hypothesis of the theory of adverse selection is that customers with high expected future health care costs will be more likely to buy health insurance. We posit that expected future health care costs are higher for households with worse observable health H_{ij}. This assumption implies our first hypothesis (H1), adverse selection in joining based on observables: customers with factors observable to the econometrician that predict high future health care costs (*e.g.*, high past health care utilization) will be more likely to buy insurance than other households.

We measure observable health as a vector including a member in self-reported poor health; a recent serious health shock, defined as an illness or injury that resulted in missing 7 or more days of normal activities, a death, or an illness resulting in an expense of more than US$100; a child who is stunted or wasted; and a member who is under 6 or over 64 (groups with high health care utilization [DHS, 2005]).

Customers likely to use covered services (*i.e.*, public health facilities) have higher expected costs. Thus, we also include in H_i an interaction of having had a serious health shock and having received treatment for it in a public facility.

To test Hypothesis 1 we estimate a probit predicting SKY membership for household *i*:

$$SKY_i = F\left(\sum_j \beta_j . H_{ij} + \sum_k \beta_{D_k} . D_{ki} + \varepsilon_i\right)$$

Here F (•) is the probit function, D_{ki} is a list of demographic and other control variables, and ε_i is an error term. We analyze the effect of other variables on take-up in a companion paper (Polimeni and Levine, 2011).

1.2. Selection on Observables at High *versus* Low Price

At zero price economic theory suggests everyone takes insurance and there is no adverse selection. At a very high price, only those most likely to use expensive care will purchase insurance. More generally, Akerlof's theory of adverse selection implies (H2 : adverse selection in joining based on observables and price): the effects in Hypothesis 1 hold more strongly when insurance prices are high than when prices are low.

We test hypothesis 2 by adding interactions of observable poor health and the price of insurance (P_i) to equation 2:

$$SKY_i = F\left(\beta'_{P_0} . P_i + \sum_j \left(\beta'_j . H_{ij} + \beta'_{Pj} . P_i . H_{ij}\right) + \sum_k \beta'_{D_k} . D_{ki} + \varepsilon'_i\right)$$

Here $P_i = 1$ for consumers who were offered the full price and $P_j \bullet H_{ij}$ is the interaction of full price and poor pre-SKY health, for the several measures of pre-SKY health, H_{ij} (health prior to being offered SKY). We use fewer health measures and covariates (D_{ki}) in estimating equation (2) than in (1) because of the modest number of SKY buyers who paid full price. The implication of Hypothesis 2 is , $\beta_{Pj} > 0$ for all $j > 0$; that is, that poor health is a particularly strong predictor of SKY membership for those paying full price ($P_i = 1$).

1.3. Selection on Unobservables

While the previous tests compare the health and health care utilization prior to being offered insurance of those who later buy or decline insurance, the insurer cares about health care utilization after buying insurance. However, we cannot measure adverse selection by comparing the utilization after being offered insurance of those who join *versus* those who decline, as both moral hazard and adverse selection can lead to high utilization.

Our randomized prices provide a means to test for adverse selection, even when the selection occurs based on factors unobservable to the econometrician. As noted in Hypothesis 2, economic theory suggests adverse selection will be more severe at high insurance prices than at low prices. As long as factors observed by the econometrician do not predict insurance uptake perfectly, we have H3, adverse selection in utilization based on factors unobservable to the econometrician (but observed by households): among the insured, those who buy insurance at a high price will have higher health care utilization after they are insured than those who buy insurance at a lower price (holding constant observable factors).

The difference in utilization between insurance buyers paying full *versus* the discounted price is equal to total selection due to the difference in price. Any differences in utilization that remain after controlling for characteristics that were observable at the baseline are plausibly due to selection based on unobservables.

To test for selection on unobservables (Hypothesis 3), we run a probit regression among all SKY members in our sample, predicting post-SKY health care utilization as a function of insurance price and pre-SKY characteristics:

$$Y_i = F\left(\gamma_{P_0}.P_0 + \sum_j \left(\gamma_j.H_{ij} + \gamma_{Pj}.P_j.H_{ij}\right) + \sum \gamma_{D_k}.D_{ki} + \mu_i\right)$$

Here, Y_i is a measure of health care utilization for household *i*. We examine three measures of health care utilization in the three months following SKY purchase: an indicator variable equal to 1 for households with at least one health center visit; an indicator variable equal to 1 for households with at least one hospital visit; and the total of the list price of all services covered by SKY for all visits to public health centers and hospitals. The function F(•) is probit for predicting the indicator variables, and OLS or Tobit for predicting total costs.

Hypothesis 3 posits households that pay full price for SKY ($P = 1$) will have higher utilization of health facilities following SKY purchase than households with similar observables who paid a deeply discounted price: $\gamma P_0 > 0$.

Because buyers paying different prices have identical health coverage, moral hazard due to different out-of-pocket costs cannot explain differences in post-SKY utilization between buyers paying full and discounted prices. However, a more behavioral version of moral hazard is possible if households who purchase SKY at full price are more determined to "get their money's worth", a type of sunk cost effect (Tversky and Kahneman, 1981). Thus, if we see that buyers paying full price have higher utilization than buyers paying a deeply discounted price, it may be due to a combination of adverse selection and this behavioral moral hazard. We present evidence on the importance of behavioral moral hazard under robustness checks below.

2. The Setting

In this section we describe health care in Cambodia, SKY health insurance and our randomization procedures.

2.1. Health Care in Cambodia

Cambodia is among the world's poorest and least healthy nations. It ranks 188 out of 229 nations in GDP per capita, has the 38th highest infant mortality rate (of 224 countries with data), and the 46th lowest life expectancy (Central Intelligence Agency, 2010).

Health shocks often contribute substantially to indebtedness and loss of land. For example, one study followed 72 households with a member who had suffered dengue fever following a 2004 outbreak. A year later, half the families still had outstanding health-related debt, with interest rates between 2.5% and 15% per month. Several of the 72 families had found it necessary to sell their land to pay their debt (Van Damme, Van Leemput, Por, Hardeman and Meessen, 2004). Annear *et al.*, (2006) found similarly high levels of indebtedness due to medical expenses.

Cambodians rely on a mix of health care providers: public providers, private medical providers, private drug sellers (with and without pharmaceutical training) and traditional healers.

Public facilities consist of local health centers, which provide basic care for everyday illnesses, Operational District Referral Hospitals for illnesses requiring more involved treatment, and Provincial Hospitals for care of more severe health shocks. Public facilities are subsidized by the Cambodian government or other organizations.

However, public facilities suffer from low utilization rates. According to 2005 DHS estimates, of those who sought treatment for illness or injury, less than a quarter went to a public health facility, and instead preferred to visit private providers, including medical doctors, drug sellers, traditional healers and midwives. Private providers of varying capabilities are typically more popular than public ones, even when more expensive, because they often are more attentive to clients' needs, more available, willing to visit patients in their homes, and willing to provide more of the treatments patients prefer. They are also usually willing to extend credit (Collins, 2000; Annear, 2006). However, even when households utilize local private doctors and drug sellers for small health shocks, many visit public hospitals for surgery and other major health problems.

2.2. SKY Health Insurance

SKY health insurance was originally developed by the French NGO GRET as a response to high default rates among its microfinance borrowers due to illness.[3] Since 1998, GRET has been experimenting with microinsurance schemes by examining responses to different premiums and benefits. Historically, take-up of insurance has ranged from 2% in regions where insurance has been only recently introduced to 12% in the longest-served regions.

While the SKY program targets the poor, it also is trying to avoid financial losses and become financially sustainable (without donor support) in the long term. Thus, the policy includes several terms that limit adverse selection. For example, SKY insurance does not cover long-term care of chronic diseases. In addition, SKY does not pay for the delivery of babies within the first few months of joining. A

[3] SKY is an acronym from *Sokapheap Krousat Yeugn*, which means "health for our families" in Khmer.

government policy also reduces adverse selection: separate government programs pay for the very expensive drugs for HIV/AIDS and tuberculosis. Finally, insurance is purchased at the household level, eliminating the possibility that households would purchase insurance for only very ill or frail members.

At the time of the study, households were offered insurance at a rate ranging from US$0.50 per household per month for a single-person household to around US$2.75 per household per month for a household with eight or more members. Households sign up for a six month cycle, paying for the first month's coverage plus two reserve months up front. While a household can stop insurance payments at any time, failing to pay two consecutive months before the end of the six-month cycle results in the loss of one month of reserve. A household can join SKY at any time, but coverage will not begin until the start of the next calendar month. Households offered insurance for the first time are offered slightly lower premiums to encourage take-up. With their insurance, household members are entitled to free services and prescribed drugs at local public health centers and at public hospitals with a referral (SKY, 2009).

2.3. Randomization

Our randomized experiment was carried out as the SKY program began a major expansion at the end of 2007. When the SKY program first rolls out into a region, SKY holds a village meeting to describe the insurance product to prospective customers. The meetings are advertised ahead of time *via* loudspeaker announcements in each village. Following the meeting, SKY insurance agents visit households to sell insurance to interested families.

At the end of each meeting, SKY traditionally gives out coupons for one month of free insurance. To randomize the price of insurance, we implemented a Lucky Draw for insurance coupons. Winners of the Lucky Draw received a coupon for a deep discount off of the insurance premium, a "high" coupon, while others received the traditional one-month discount, a "low" coupon. The high coupon entitled the bearer to five months free out of the first six month insurance cycle, with the possibility of another three-month discount if the household renewed for another six months. High coupons were meant to expire two months from the date of purchase. All households have the ability to purchase SKY at any time at full price (that is, first-time buyers can pay 5 months for 6 months of coverage).[4]

[4] Details of the randomization of price are available in Appendix C.

3. Data

Our analyses use a household survey, SKY administrative and health facility utilization data and several other sources.

3.1. Household Survey

Our main data source is a survey of over 5,000 households. We rely largely on the baseline survey, which took place from two to nine months after the initial SKY marketing meetings. (Households could first start SKY coverage between one and two months after these meetings.)

The baseline survey covered demographics, wealth, self-perceived and objective health measures, health care utilization and spending, assets and asset sales, savings, debt, health risk behaviors, willingness to take financial risks, trust of health care institutions, means of paying for large health expenses, and time preference. Appendix D describes the measures used in the current analysis in more detail.[5]

SKY meetings were held in Takeo, Kandal and Kampot provinces from November 2007 to December 2008. For the baseline survey, we interviewed all Lucky Draw winners (the 20% of the village meeting attendees able to purchase SKY at a deeply discounted price) and an equal number of households offered the standard price (specifically, every fourth house on the village meeting attendance list that were offered the standard price). To increase our sample of buyers who paid full price, we also interviewed all households with a low coupon that bought insurance. In total, our sample consists of 2,537 households offered the deep discount, 2,534 households offered full price and randomly included in the sample, and 196 over-sampled households who purchased insurance at full price. A summary of sample size is presented in Appendix Table A.1.[6]

3.2. SKY Administrative and Utilization Data

For each household that joins SKY, SKY records the registration date, the date the household starts coverage, and the date the household drops out of SKY. We use this SKY administrative data to determine if and when each household from the village meeting purchased SKY insurance. To match our baseline data to the SKY database, for each village, we matched the name of household member in our survey to the names listed in the SKY database.

SKY also collects utilization data whenever a member visits a health center or hospital with which it partners, including the list price of the services provided. We use these data to measure utilization of public health centers and hospitals in the months following SKY purchase, as well as the total cost of all visits paid for by SKY.

[5] Other measures are explored in Polimeni and Levine, 2011, and described in the Appendix to that paper.

[6] A timeline of the study is presented in Levine, Polimeni and Ramage, 2011.

3.3. Other Datasets

We also used data from a second-round household survey, interviews with village leaders, and measures of health center quality as control variables in some analyses (Appendix B).

3.4. Randomization

Appendix Table A.2 shows average characteristics of high and low coupon winners at the baseline. To account for recall errors and to ensure we are looking at pre-randomization health events and behaviors, we look at health events taking place in the month before the SKY meeting and two months prior (for a total of three months).

Of the thirty variables tested, only three show a statistically significant difference between high and low coupons at the 5% confidence level. There is a statistically significant but very small difference in the percentage of households that are Khmer between the two groups (95.3% *versus* 94.6% in the full and discounted price groups, respectively). 14% of low coupon households have wealth level subjectively graded as "poor" by enumerators, while only 10% of high coupon households are rated as "poor". Similarly, low coupon households are slightly more likely to live in a house made of palm, another measure of lower wealth. Other wealth indicators did not show significant differences. We control for wealth characteristics in our regressions and keep in mind this difference when interpreting results.

4. Results

We test our three hypotheses regarding adverse selection: whether insurance buyers have worse observables prior to joining SKY, whether that adverse selection is more severe for those paying full price, and whether health care utilization is higher for those paying full price (conditional on observables). For additional evidence of selection, we compare utilization of households that drop SKY to that of households that remain insured.

4.1. Selection on Observables

In this section we compare pre-SKY characteristics of households that bought insurance to those that did not. We focus on households that bought SKY within 63 days of the introductory marketing meeting. For these households, pre-meeting health shocks are more representative of health immediately before SKY purchase.[7]

Households that purchased SKY are 2.3 percentage points more likely to have had a major health shock in the two to four months prior to the SKY meeting than households that did not join SKY (9.4 percent of buyers *versus* 7.0 percent of decliners, $P < 0.05$, Appendix Table A.3) and 12.7 percentage points more likely to report having a household member in poor health (80.0% *versus* 67.2%, $P < 0.01$). These results are consistent with Hypothesis 1, that SKY buyers are in worse health than decliners.

In contrast to Hypothesis 1, SKY buyers and decliners are not significantly different in other measures of high predicted health care expenditures such as presence of a stunted or wasted child, having children 5 and under, or having elderly household members.

To investigate selection on observables further, we perform a probit estimation of membership as a function of pre-SKY health (as in equation 1). Consistent with Hypothesis 1, having a household member in self-reported poor health has a large effect on insurance purchase. Around 22% of pooled low and high coupon households buy SKY insurance. A household is more than 12 percentage points more likely to buy SKY insurance if they have at least one household member in poor health at the baseline (Table 1, col. 1, 3, 4, 5, $P < 0.001$). A health shock pre-meeting leads to an 8.7 percentage point increase in the probability that a household will buy SKY insurance (col. 2, $P < 0.01$). This variable is somewhat collinear with having a household member in poor health (col. 3), and the result is driven by households that used a public health center rather than a private facility for care (col. 4-5, joint results not significant at traditional levels). Number of days ill and spending more than US$30 on a shock do not have a significant impact on take-up above and beyond having a major shock.

As in the comparison of means, the other measures of high expected health care costs such as having a child or elderly person in the household and having a stunted or wasted child do not predict high SKY membership. We explore the effects of other household characteristics on take-up in a companion paper (Polimeni and Levine, 2011).

[7] Results are similar for the full sample (Table A.6).

4.2. Selection on Observables by Price

Hypothesis 2 posits that households that purchase SKY at full price (that is, with a low coupon) will have higher expected health care costs on average than households that bought insurance at a deeply discounted price. Table 2 presents results of a probit regression of SKY purchase on price and price interacted with health characteristics. We include only a limited number of health characteristics due to the small sample size of purchasers paying full price.

Contrary to Hypothesis 2, having a member in poor self-reported health does not lead to a statistically significant higher increase in SKY uptake among those offered SKY at full price than among those offered a deep discount, although the point estimate is positive for households in poor health that were offered SKY at full price (β = 2.7%, SE = 3.3%, not significant, col. 1). In contrast to Hypothesis 2, there is no evidence that a major health shock prior to the SKY meeting has more effect on purchasing SKY among households offered full price than among those offered SKY at a deep discount (β = -3.8%, SE = 4.6%, not significant, col. 2).

4.3. Selection on Unobservables

We did not find purchasers paying full price had worse observable health prior to buying SKY than purchasers paying the discounted price (in contrast to Hypothesis 2). However, if some adverse selection is based on characteristics that are unobservable to the researcher at the baseline, then SKY buyers who paid full price will use health facilities more than SKY buyers who purchased at a deep discount, even when holding constant baseline characteristics (Hypothesis 3).

We examine all SKY buyers (not just early buyers) and measure health care costs based on the list cost (that is, the fees paid by uninsured patients to public facilities) of services provided to SKY members. Summary statistics are presented in Table 3.

In the regressions with no covariates, buyers paying full price are respectively 11.1 and 10.7 percentage points more likely to use a health center ($P < 0.001$) and hospital ($P < 0.01$) in the first 3 months after SKY purchase than buyers at the deeply discounted price (Table 4, col. 1 and 3).[8] Buyers who paid full price also have 60% higher health care costs than buyers at a deep discount, on average (col. 5).[9]

To test Hypothesis 3 we measure how much of this higher utilization remains after controlling for observed baseline characteristics (equation 3). Even after we condition on our pre-SKY health covariates, the price of insurance continues to predict increased likelihood of our three utilization measures (Table 4, col. 2, 4, 6). Estimated effect sizes are only slightly lower (at about 20% for hospital use and total cost) than the unadjusted effects, though none of the declines are statistically significant (controlling for observables has almost no impact on the effect size for health care use). Results are similar if we include all available covariates (Table A.8, covariates listed in Appendix D). Specifically, conditioning on baseline characteristics, buyers paying full price are 11.3 and 8.5 percentage points more likely to visit a SKY-partnered health center and hospital, respectively, in the first three months following SKY purchase (Table 4, $P < 0.001$, col. 2, and $P < 0.05$, col. 4) than are buyers at a steep discount. Totaling

[8] The summary statistics (Table 3) measure utilization in the first three months of SKY for all insured households, coding usage as zero in a month if a household dropped prior to that month (in which case we have no utilization data). In the regressions (Table 4), we include indicator variables for households that dropped in the first, second or third month of SKY.

[9] For total cost, we take the log of cost in USD. The percent increases in spending by the households paying full price are calculated by comparing the exponentiation of the coefficients on the price variable to the average cost of utilization by households that bought with a deep discount (Table 3). Results are similar when we use the Tobit functional form for columns 5 and 6 in Appendix Table A.12.

health centers and hospital visits, buyers at full price use services costing about 56% more than buyers at a deep discount (col. 6).

In short, even after controlling for all observed baseline characteristics, households that buy SKY at full price have higher utilization than those that buy at a deeply discounted price. Figure 1 summarizes our results from Table 4 graphically: Households paying more for insurance have higher utilization, and this result is barely decreased even after holding baseline variables constant. Baseline characteristics account for none of full price buyers' higher utilization of health facilities, 21.5% of their higher hospital utilization, and only 8.1% of their higher SKY-covered costs.

4.4. Dropout

Complementing selection into SKY, in this section we study self-selection among those who remain in SKY.

Dropout is highest among those who have not used SKY. For example, consider dropouts in months 6 (when the initial deep discount ran out for high-coupon buyers) to month 15 (when our dataset becomes too small to analyze). During these months, dropout averages 9% among those who have used SKY-covered services, but 15% among those that have not. (This calculation is complex because most households that remain in SKY shift into the "ever used SKY" category over time.)

A hazard analysis of dropout on utilization in the prior month demonstrates this relationship as well. We used Cox proportional hazard distribution to predict $Dropout_{it}$, the probability that a household *i* that is a member of SKY after t - 1 months of membership will drop out in the next month, as a function of whether the household visited a public hospital in the previous month or the month prior to that, visited a public health center in the previous month or the month prior to that, or never visited a SKY-funded facility.

The hazard estimates (Table 5) are that households that utilized a health facility one additional time in the one or two months prior are 28.8 ($P < 0.001$) and 2.7 (not statistically significant) percent less likely to drop SKY, respectively, in any given month. Similarly, a household that used a hospital one additional time in the one or two months prior are 48.3 and 42.1 percent less likely to drop SKY in any given month ($P < 0.001$).

The low rates of dropout among those who use SKY facilities more will lead to an increasingly costly customer base if there is autocorrelation in health care utilization. Autocorrelation in utilization is, in fact, fairly high among SKY members. For example, households that had a health center visit in months 1-3 were 36 percentage points more likely to visit a health center in months 4-6 than were households that did not use a health center in months 1-3 (results not shown, 69.8 % versus 34.1%, $P < 0.01$). Further, the total list price of all health center and hospital care in months 1-3 has a correlation of 0.41 with the cost of care starting in months 4, 5 or 6 (Appendix Table A.5, col. 5).

Autocorrelation in care coupled with the exit of households with low past utilization means SKY members' average utilization will rise over time. Figure 2 shows this trend of rising utilization in hospital utilization, but not for health center utilization. The proportion using a hospital more than doubles over the first 18 months, while the proportion using a health center falls by about a third. (Differences over time are statistically significant at the 5% level in 6 out of the first 8 months for health center utilization when comparing each month to month 18. Differences are not significant at traditional levels for hospital utilization.) When we look at costs of covered services (not shown), we find that covered health center costs decrease by 24%, from an average of US$0.48

per month in the first three months of coverage to US$0.36 per month in months 16 to 18 (P < 0.001), while hospital costs almost doubled, increasing from US$0.37 to US$0.72 in the same period (P < 0.001). Total average monthly SKY-covered costs (health center plus hospital) increased by around 27% over that period, from US$0.85 to US$1.08.

A second possible reason for increased adverse selection is that high coupon households paid only one of the first six months of insurance, but had to pay triple that price to purchase six additional months. Theory suggests this price increase will make insurance less attractive for households with modest expected health care expenditures, leading to an increase in adverse selection among buyers initially paying a deeply discounted price. Thus, renewing high coupon households will more closely resemble buyers who always paid the full price.

Figure 3 graphs utilization for all households that are still members of SKY a given number of months after initial purchase. The figure shows that while households that purchased at the full price initially have higher hospital utilization than those who purchased SKY with the deep discount, the difference tapers off by around the sixth month after SKY purchase (monthly differences are not statistically significant at the 5% level for any month). However, for health center use, households that purchased at the full price have consistently higher utilization than those that purchase at the discounted price (monthly differences are significant at P < 0.05 in months 2-5, 13, and 15).

5. Robustness Checks

In this section we present several robustness checks.

5.1. Early *versus* Late Buyers

Households can purchase SKY insurance at any time. Also, although high-value coupons were intended to expire after 2 months, we observed a number of households using the coupons after that date. Thus, it is possible that some households waited to purchase SKY until they were more in need of health care. Thus, these late buyers may exhibit high utilization immediately following SKY purchase that tapers off shortly after. If low coupon households delay more than high coupon households, perhaps because some high coupon households believe they must use their coupon immediately, higher selection among low coupon as compared to high coupon households may be due to both higher price and timing of purchase.

Some households do buy SKY after they have a health shock that needs care, and this pattern appears to be more common among households offered SKY at full price (Appendix Table A.4). Compared to early buyers at full price, late buyers who paid full price for insurance are less likely to report a member in poor health in the baseline survey (81% *versus* 86%, difference not significant at 10% level). On the one hand, this means household members in late-buying households are subjectively healthier, but on the other hand it may mean that their health care needs are less urgent, and thus these households wait to buy SKY until they experience an increase in utilization. Supporting this hypothesis, late buyers are twice as likely as early buyers to report a major health shock in the 3 months immediately following SKY purchase ($P < 0.01$). Also consistent (if less precisely), late buyers are 8 percentage points more likely to have a major health shock ($t = 1.59$), and 3 percentage points more likely to have used a hospital for care in the 3 months immediately preceding SKY purchase ($t = 0.95$).

Results are not as clear for households offered a deep discount that waited to purchase insurance. Like low coupon buyers, high coupon late buyers are somewhat less likely to have a member in poor health at the baseline. However, high coupon late buyers are not significantly more likely to have reported a major health shock in the months immediately prior to or immediately following SKY purchase. Thus, unlike low coupon households that delayed purchase of SKY, high coupon late buyers may have delayed purchase for reasons other than waiting for a health shock to arise.

5.2. Selection on Unobservables

We re-ran the tests of selection on unobservables including additional baseline covariates (Appendix Table A.8); adding an indicator variable for the timing of SKY purchase (immediately following the SKY meeting *versus* later, Appendix Table A.9), conditioning on data from 1 to 3 months pre-SKY instead of 2 to 4 months (Appendix Table A.10); only including the randomly selected sample of households who were offered the full price (dropping the over-sample of SKY buyers who were offered full price, Appendix Table A.11) and controlling for only pre-SKY shocks that caused 7 or more days of missed work (instead of also including a shock leading to a death or a US$100 expense, Appendix Table A.13). We also switched the functional form from OLS to Tobit to predict health care expenditures (Appendix Table A.12). In general, results were very similar to those presented.

5.3. Behavioral Moral Hazard

It is possible that buyers paying full price use more SKY services than those buying with a steep discount due to a sunk cost effect, or "behavioral" moral hazard, whereby households that paid a higher price for insurance utilize services more to "get their money's worth". However, we do not believe this is the case. First, recent research has not found evidence for this type of behavioral incentive to seek care. For example, in the case of treated bednets, Cohen and Dupas (2010) find no decreased usage of bednets for those receiving these nets for free, and Tarozzi *et al.* (2011) found that households randomly chosen to receive free bednets used them even more than households paying for these nets. Second, the costs of health care include several costs that are not covered by SKY, including the opportunity cost of lost time and travel costs.

Finally, while households may increase care for small illnesses at health centers to get their money's worth, we found that low coupon households also have much higher utilization of hospital services, which typically require referral from a health center.

To focus more specifically on very severe illnesses, we examined the subset of hospital visits with an overnight stay. Households purchasing SKY at full price are more than twice as likely to have an inpatient visit in the first three months after SKY purchase than households purchasing at the discounted price (approximately 12.8% *versus* 4.5%, $P < 0.015$, mean comparison not shown). The much higher utilization among those buying at full price remains after holding constant baseline characteristics ($P < 0.05$, Appendix Table A.14, col. 2).

5.4. Hazard Rates by Price

Results of the hazard rates on dropping out of SKY are similar for households paying either full or the deeply discounted price. One exception is that households that purchased at the high price are significantly more likely to drop SKY, compared to households purchasing at the discounted price, if they have never used a health center or hospital during the three months prior. Unlike households paying a lower price for insurance, households purchasing at the higher price were significantly less likely to drop based on health center utilization in the two months prior, but this result does not seem meaningful when taking into account other small differences in coefficients (results not shown).

5.5. Financial Implications of Adverse Selection

Our results indicate that while there is some adverse selection based on characteristics observable at the baseline (*e.g.*, past health care utilization and baseline health levels), most adverse selection was not based on baseline observable characteristics. How important is this observable and unobservable selection to SKY's bottom line?

The usual SKY insurance premium is set to cover the list price of services that SKY members use, what we will call "utilization costs". (We have no data on administrative or sales expenses.)

In our data, in the first six months after joining, the average utilization costs among all buyers who paid full price for SKY was US$6.94 (Table 6, col. 1). Using back-of-the-envelope calculations based on average household size and premium levels, average revenue per household was near US$9.93 during that period, US$2.99 above utilization costs per household.

Our results showed that the large discount reduced self-selection of those with high utilization. The lowered price decreases average utilization costs from US$6.94 (column 1) to US$5.29 (column 3) per person in the first 6 months of SKY. However, the steep price cut costs far more in revenue than it saves in reduced utilization, so the steep discount leaves revenue US$3.64 below utilization costs (column 3), without considering administrative or sales costs.

SKY does not wish to exclude any households from purchasing SKY, and in fact this would be counter to their mission. However, to demonstrate the size of adverse selection on SKY's bottom line, suppose SKY excluded all households that reported "poor health" at the baseline, or that reported a large health shock prior to the baseline (where large shock is defined as above). For households that paid full price for SKY (Table 6, columns 1 and 2), these baseline measures are poor predictors of future health care utilization. Among buyers who purchased SKY at the full price (column 1), selling to only households that did not have observable poor health at the baseline actually increases average utilization costs (from US$6.94 on average to US$8.33 if they did not report poor health or to US$7.58 if they reported no large health shock in the 3 months prior to purchase). This unexpected result may be due to the very low percentage of buyers at the full price that did not have a health shock or a member in poor health: only 16% of households that bought SKY at the full price did not have a member in poor health. While these households did not report poor health, our results on selection based on unobservables indicate that these buyers are apparently very adversely selected. If we look at only buyers that bought at full price soon after the Village Meeting (column 2), not selling to households with pre-SKY shocks lowers average monthly costs by around US$0.40 (US$6.23 to US$5.82).

Pre-SKY health does predict high utilization costs for households that purchased at the discounted price (columns 3 and 4). If SKY were to sell only to households that did not report a member in poor health, average costs over the first 6 months of SKY would be decreased from US$5.29 to US$3.60. However, there is little change in costs based on a major health shock in the three months pre-SKY. These results hold for the full sample (column 3) and when we look at early buyers only (column 4). Even if SKY were able to discriminate based on poor health, less than a quarter of buyers at the discounted price do not report a household member in poor health.

In short, SKY would not be able to improve its financial outcomes substantially by avoiding self-selection based on factors we observed.

Conclusion

To study who buys health insurance, we randomized the price of SKY health microinsurance and surveyed households close to the start of insurance purchase. Because insurance was not previously available, any differences in baseline characteristics such as health care utilization were not influenced by past insurance. Unlike many previous studies, the randomized price of insurance allowed us to eliminate the effect of moral hazard even when looking at utilization of the insured after insurance purchase.

We found mixed evidence of adverse selection on observable baseline characteristics (Hypothesis 1). Those who join SKY have had more past health shocks and are more likely to report a member in poor health.[10] At the same time, SKY buyers are not more likely to have very young or very old members, who we believe would be more likely to be ill.

There was little evidence that those paying the full price for insurance have more observable risk factors than those paying a deeply discounted price (Hypothesis 2).

There was strong evidence that those who pay full price for insurance have more unobservable risk factors than those who pay a deeply discounted price (Hypothesis 3). Specifically, those who bought insurance at full price have much higher health care utilization than those who paid a lower price, and this gap in utilization based on insurance price remains even when controlling for a number of observable baseline characteristics.

While some households that paid full price may have been using more health care to "get their money's worth" (a behavioral form of moral hazard), we have evidence that most of the effect is due to adverse selection on unobservables. Specifically, the effects were just as strong for hospital visits (which require a referral) – even overnight hospital visits – as for health center visits.

Ignoring past health characteristics has little impact on SKY's bottom line when we look only at households paying the full price for insurance. For these households, past health utilization and self-reported health status had little predictive power on health utilization after purchasing SKY, but this result is probably due to the low proportion of these households purchasing SKY that did not have a member in poor health. For households paying the discounted price, discriminating against households with a member in poor self-reported health lowers total costs per household by around 32% in the first 6 months of SKY, but discriminating against households' recent health shocks does little to lower health costs. The lowered price in general increases risk pooling and decreases average costs by 24%, but the steep decrease in revenue is not enough to cover these decreased costs.

Finally, insured households that have low health care utilization are more likely to drop insurance coverage. Consistent with this result, households remaining with SKY have higher average hospital utilization over time. Dropout of households with low utilization meant that by six months, average hospital utilization of remaining buyers who paid the deeply discounted price was equivalent to that of SKY members paying the full price. However, we did not find this convergence in health care utilization, and average health care utilization decreased for households who remained in SKY for longer periods of time. Future research will explore whether households with low utilization who dropped SKY overestimated their utilization needs, did not understand SKY insurance, or were dissatisfied with SKY-partnered health care services.

[10] Self-reported health was collected a few months after households joined SKY. Self-reports will under-estimate adverse selection if SKY insurance improves health but will over-estimate adverse selection if SKY insurance raises members' awareness of health problems.

In short, our results imply that a lower insurance premium leads to a more diverse risk pool with lower average utilization, consistent with the economic theory of adverse selection.

If health insurance is to address the problems of the global poor we must understand under what conditions insurance can be financially sustainable and can improve purchasers' lives. The financial viability of an insurance program depends in part on its ability to avoid adverse selection and excess utilization. SKY has a number of rules to reduce adverse selection. For example, the requirement that all members of a household must join SKY at once may explain why SKY households are not more likely to have children or elderly members than are SKY decliners.[11] At the same time, SKY policies to limit adverse selection are limited. The result is that SKY faces a high-cost population of members, which reduces both its ability to attract new members and its financial viability.

Protecting those in need is a major objective of donors and of policy-makers. Unfortunately, better protection against adverse selection, such as restrictions on coverage of pre-existing conditions, worsens insurance's ability to protect those in need. If households buying insurance are disproportionately those with the highest health care needs, subsidizing voluntary health insurance may be a cost-efficient way to increase access to health care. Ongoing donor or government support is appropriate if health insurance both increases health and has other benefits such as reducing persistent poverty due to health-related debt. A companion paper examines the health care and economic impacts of SKY (Levine, Polimeni and Ramage, 2011). At the same time, without quality health care services, increases in utilization may make little difference in health, so donors must weigh the value of increasing quality of care *versus* increasing access to care.

In addition to its policy implications, our results provide evidence of adverse selection, separated from the often confounding effects of moral hazard. More research is needed to deepen our understanding of insurance in developing countries, and to understand whether the lessons learned in our evaluation hold for insurance programs in other developing countries.

It is important to understand the other characteristics that lead households to purchase health insurance in SKY and other programs. For example, if low quality health facilities or a lack of understanding of insurance are causing some households to remain uninsured, insurers may be able to attract a more diversified risk pool by addressing these barriers. A companion paper (Polimeni and Levine, 2011) explores these topics.

It is also important to understand the dynamics of insurance. It is possible adverse selection will decline over time, for example, if insurance becomes more widespread. Longer-term studies of who buys and renews insurance policies are needed to address such questions.

[11] However, we do not find any evidence that even one-member households with an elderly member are more likely to purchase SKY (results not shown). Alternatively, households may be equally or more concerned with the care of working-aged members who contribute to household income. This possibility is explored in a related paper (Polimeni and Levine, 2011).

Tables

Table 1: Probit Regression of SKY Take-up on Baseline Characteristics

	(1)	(2)	(3)	(4)	(5)
Offered Full Price (d)	-0.362*** [0.0171]	-0.361*** [0.0172]	-0.362*** [0.0171]	-0.362*** [0.0171]	-0.369*** [0.0172]
At least one household member with poor self-reported health (d)	0.124*** [0.0125]		0.121 *** [0.0125]	0.121 *** [0.0125]	0.127*** [0.0131]
Major health shock (*) in the 2 to 4 months pre-meeting (d)		0.0866** [0.0307]	0.0634* [0.0301]	0.103 [0.108]	0.12 [0.116]
Major health shock (*) and used health center for care (0 if no shock) (d)				0.0315 [0.0563]	0.0212 [0.0578]
Major health shock (*) and used hospital for care (0 if no shock) (d)				-0.0313 [0.0496]	-0.0171 [0.0538]
Major health shock (*) and used private health care (0 if no shock) (d)				-0.027 [0.0496]	-0.0451 [0.0477
Ln of max. days ill for a major health shock (*) pre-meeting (0 if no shock)				-0.00865 [0.0196]	-0.00632 [0.0210
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)				0.0285 [0.0591]	0.0202 [0.0585]
At least one member over 64 (d)					-0.02 [0.0154]
At least one member age 5 or under (d)					-0.00207 [0.0163]
Household has a stunted or wasted child under age 6 (d)					0.0165 [0.0209]
Observations	4 701	4 701	4 701	4 701	4 387
Pseudo R-squared	0.168	0.154	0.169	0.169	0.179

*Notes: LHS variable 1 if bought SKY, 0 if declined (SKY administrative data). + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. All data are from the baseline survey. Wealth, household size and education are included in the regression but not presented. Sample is all SKY decliners and all SKY buyers who bought SKY within 63 days of the village meeting. (d) for discrete change of indicator variable from 0 to 1.*

Table 2: Probit Regression of SKY Take-up on Baseline Characteristics Interacted with Price

	(1)	(2)	(3)
Offered Full Price (d)	-0.381*** [0.0299]	-0.358*** [0.0173]	-0.381*** [0.0298]
At least one household member with poor self-reported health (d)	0.116*** [0.0163]		0.112*** [0.0162]
Full price * Poor self reported health	0.0268 [0.0325]		0.0295 [0.0326]
Major health shock (*) in the 2 to 4 months pre-meeting (d)		0.107** [0.0389]	0.0794* [0.0375]
Full price * Major health shock		-0.0379 [0.0458]	-0.0304 [0.0456]
Observations	4 701	4 701	4 701
Pseudo R-squared	0.168	0.154	0.169

*Notes: LHS variable: 1 if bought SKY, 0 if declined (SKY Admin data). + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death or an expense of over US$100. All data are from the baseline survey. Wealth, household size and education are included in the regression but not presented. Sample is all SKY decliners and all SKY buyers who bought SKY within 63 days of the village meeting. (d) for discrete change of indicator variable from 0 to 1.*

Table 3: Summary Statistics, Buyers at Full *versus* Discounted Price

	Bought at Full Price	Bought with Deep Discount	
SKY paid for health center visit in first 3 months of SKY	0.72	0.64	*
SKY paid for hospital visit in first 3 months of SKY	0.21	0.11	**
Cost of all SKY-paid care in first 3 months of SKY (USD)	4.09	2.49	+
Ln Cost of all SKY-paid care in first 3 months of SKY (USD)	1.15	0.82	**
A least one household member with poor self-reported health	0.84	0.77	**
Major Health Shock (*) in the 2 to 4 months pre-SKY	0.12	0.10	
Major health shock (*) and used health center for care (0 if no shock)	0.02	0.03	
Major health shock (*) and used hospital for care (0 if no shock)	0.03	0.02	
Major health shock (*) and used private health care (0 if no shock)	0.09	0.07	
Ln of max days ill for a major health shock (*), pre SKY start (0 if no shock)	0.36	0.32	
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock)	0.07	0.06	
Observations	243	1 262	

*Notes: + p<0.10, * p<0.05, ** p<0.001, based on t-tests clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. Post-SKY utilisation data are from SKY records; zeros averaged in utilization for months 2 and 3 for households that dropped SKY in these months. All other data is from the baseline survey. Sample is all SKY decliners and all SKY buyers.*

Table 4: Effects of Self-Selection on Utilization

	(1)	(2)	(3)	(4)	(5)	(6)
Offered Full Price (d)	0.111*** [0.0321]	0.113*** [0.0328]	0.107** [0.0342]	0.0849* [0.0359]	0.414*** [0.106]	0.329** [0.110]
At least one household member with poor self-reported health (d)		0.161 *** [0.0344]		0.0199 [0.0253]		0.250*** [0.0595]
Major health shock (*) in the 2 to 4 months pre-SKY (d)		0.0489 [0.152]		0.0372 [0.108]		-0.143 [0.354]
Major health shock (*) and used health center for care (0 if no shock) (d)		0.162* [0.0770]		0.0658 [0.0835]		0.0816 [0.191]
Major health shock (*) and used hospital for care (0 if no shock) (d)		0.115 [0.0885]		-0.0592 [0.0552]		-0.123 [0.198]
Major health shock (*) and used private health care (0 if no shock) (d)		0.0563 [0.105]		0.0286 [0.0872]		0.12 [0.221]
Ln of max. days ill for a major health shock (*) pre-meeting (0 if no shock)		-0.0452 [0.0408]		-0.0131 [0.0257]		0.0466 [0.0878]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)		0.116 [0.0787]		-0.0109 [0.0639]		0.0875 [0.210]
At least one member over 64 (d)		-0.0258 [0.0315]		0.00533 [0.0227]		-0.115+ [0.0638]
At least one member age 5 or under (d)		0.0852* [0.0347]		-0.00221 [0.0265]		0.111 [0.0774]
Household has a stunted or wasted child under age 6 (d)		0.0655 [0.0454]		-0.0375 [0.0284]		0.0853 [0.102]
Observations	**1 505**	**1 255**	**1 505**	**1 255**	**1 508**	**1 255**
Adjusted R-squared					0.024	0.045
Pseudo R-squared	0.009	0.048	0.017	0.029		

*Notes: +p<0.10, * p<0.05, **p<0.01,***p<0.001. LHS variables: Columns 1 and 2 (3 and 4): Indicator for use of a SKY-covered health center (hospital) for the first 3 months post SKY purchase; Columns 5 and 6: Ln of total cost (user fees, covered by SKY) of all SKY-covered health center and hospital visits in the first 3 months post-SKY. Columns 1-4 use probit, columns 5-0 use OLS. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. SKY status and LHS variables use SKY data. Coupon status was recorded at the village meeting. All other data are from the baseline survey. Wealth, household size and education are included in the regression but not presented. Sample is all SKY decliners and all SKYS buyers who bought SKY following the village meeting. (d) for discrete change of indicator variable from 0 to 1.*

Table 5: Cox Regression, Hazard of Dropping Coverage

SKY data: Num. HC visits 1 month prior	0.712*** (0.035)
SKY data: Num. HC visits 2 months prior	0.973 (0.023)
SKY data: Num. Hosp. visits 1 month prior	0.517*** (0.086)
SKY data: Num. Hosp. visits 2 months prior	0.579*** (0.076)
Never used HC or Hosp. through month n-3	1.226* (0.111)
Used HC but no Hosp. through month n-3	1.414*** (0.110)
Offered Full Price	0.993 (0.102)
Exponentiated coefficients	**17 314**

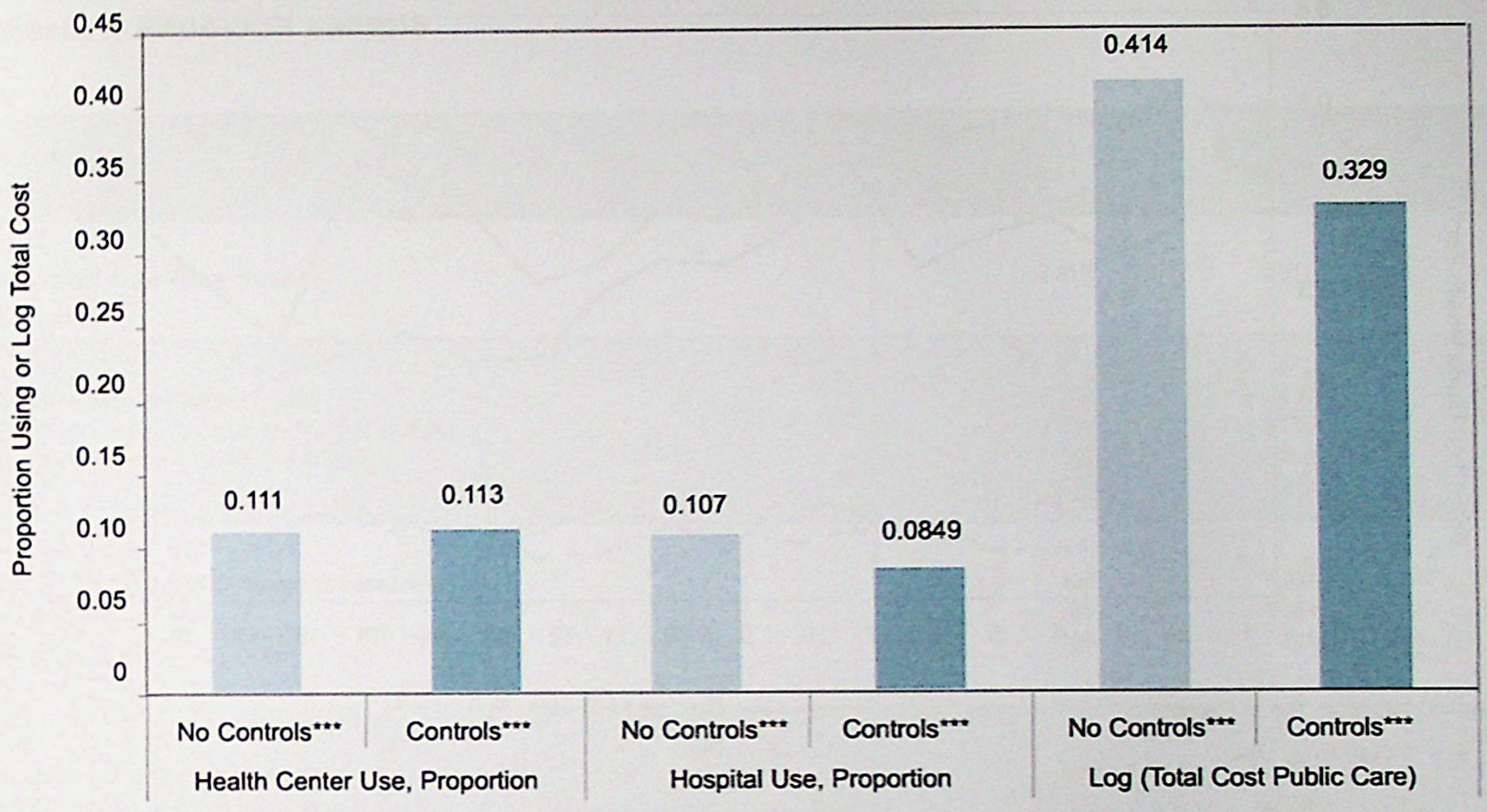
*Notes: *p<0.05. **p<0.01, *** p<0.001. LHS variable: Hazard of dropping SKY. Robust standard errors (in parentheses) clustered at the village level. Coupon status was recorded at the village meeting. All other data are from SKY records. Sample is all SKY buyers who were members anytime from Jan. 1, 2008, to Dec. 31, 2009. (d) for discrete change of indicator variable from 0 to 1.*

Table 6: Financial Implications of Selection

	(1) Full Price	(2) Full Price, Early Buyers	(3) Large Discount	(4) Large Discount, Early Buyers
All Buyers Number of Members, Month 1	359	202	1 197	856
Average Revenue, 6 months Average List Price of Services, 6 months Average Potential Margin, 6 months	$9.93 $6.94 $2.99	$9.79 $6.23 $3.56	$1.65 $5.29 -$3.64	$1.65 $5.44 -$3.79
Only households without "poor health" Number of Members, Month 1	58	27	274	183
Average Revenue, 6 months Average List Price of Services, 6 months Average Potential Margin, 6 months	$9.38 $8.33 $1.05	$9.16 $7.12 $2.05	$1.60 $3.60 -$2.00	$1.57 $3.08 -$1.50
Only households without a prior large shock Number of Members, Month 1	267	181	925	778
Average Revenue, 6 months Average List Price of Services, 6 months Average Potential Margin, 6 months	$9.91 $7.58 $2.34	$9.80 $5.82 $3.98	$1.67 $5.48 -$3.81	$1.67 $5.32 -$3.65

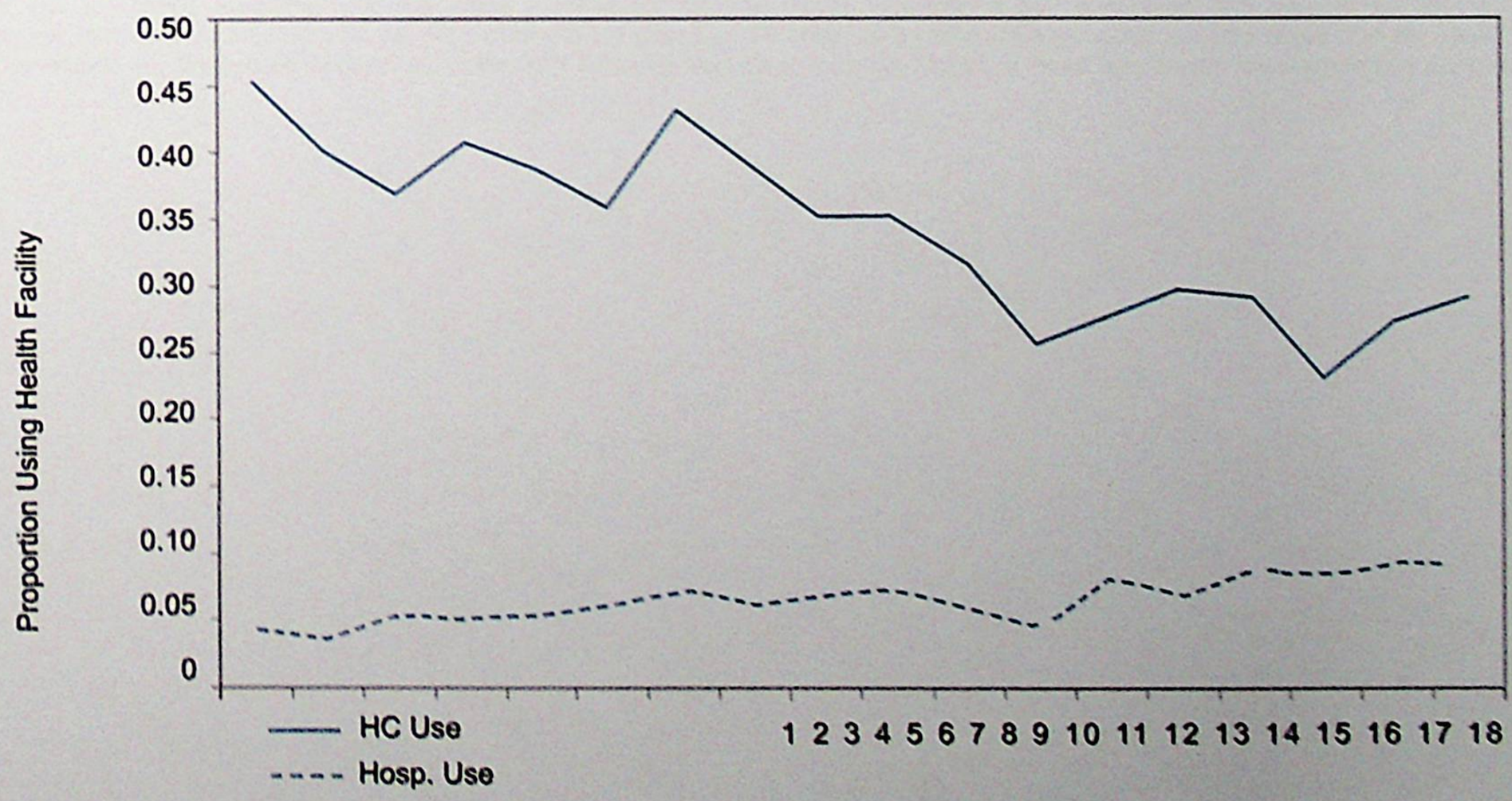
*Notes: *Averages costs per household based on SKY utilization data for the first 6 months after SKY purchase. Includes all households enrolled in SKY from January 1, 2008, to December 31, 2009. Average utilization is calculated by finding the average utilization costs in each of the first six months for any household that is a member in that month, then averaging utilization over these 6 months. Revenues are calculated by taking the average premium paid over the 6 months (taking into account household size of members in each month).*

Figure 1: Effect of Full Price (Not Steep Discount) on Utilization, with and without Baseline Controls



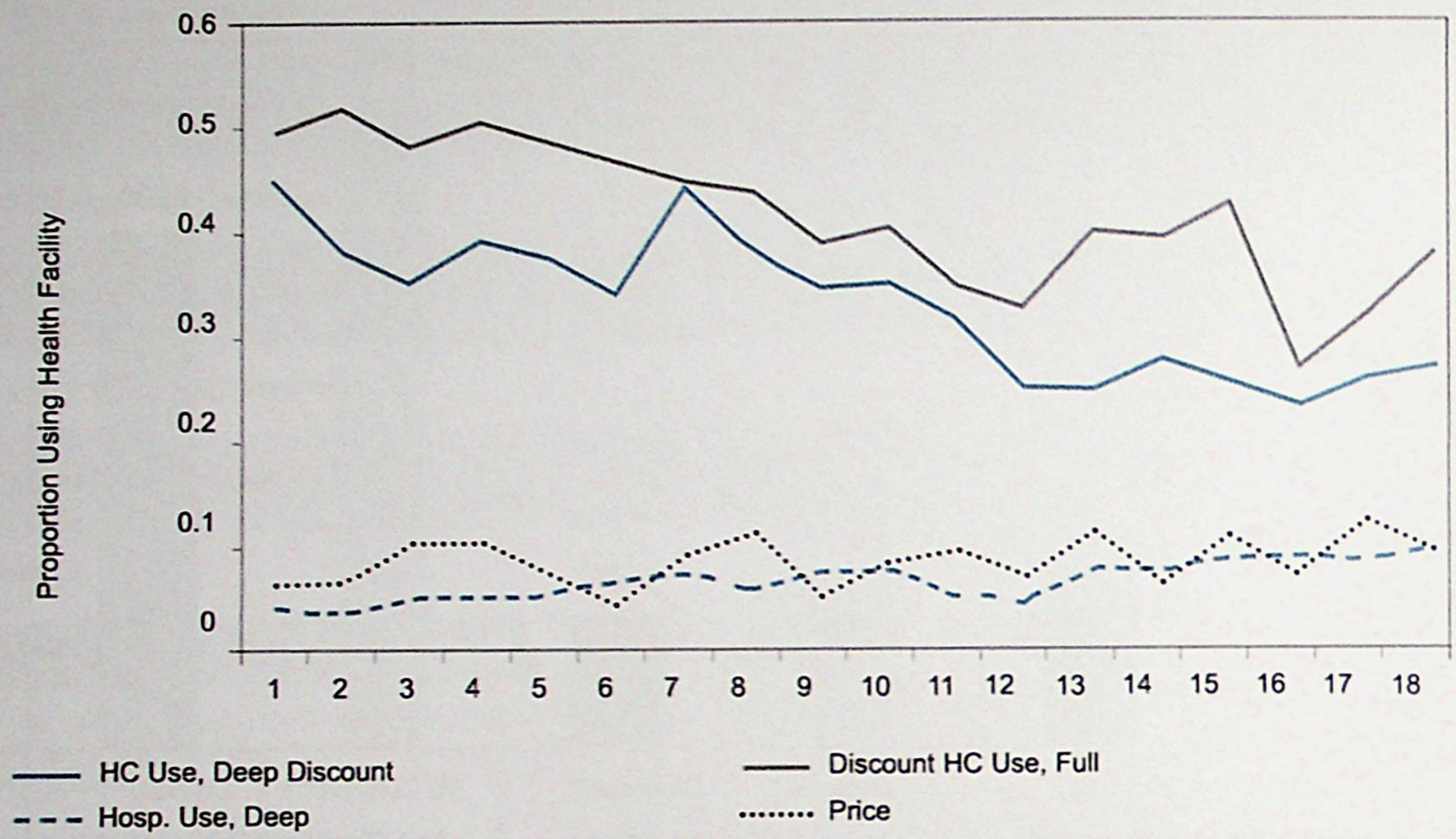
*Notes: Numbers presented are the coefficients from the regression of health center and hospital utilization, and total cost, on having been offered the full price for insurance (no large coupon). Columns labeled "No Controls" use only an indicator variable for "Full Price" on the right hand side of the regression: these coefficients represent total adverse selection. Columns labeled "Controls" use additional baseline controls: these coefficients represent hidden adverse selection (unobservable at the baseline). Stars represent significance of the coefficients in the regressions: * p < 005, ** p < -0.01, *** p < 0.001.*

Figure 2: Proportion of SKY Members Using SKY-Covered Health Facilities, by Tenure in SKY



Notes: Of households that were members for at least X months, the above chart shows the proportion that used a health center (HC) or hospital (Hosp.) in that month. Thus, the number of SKY members becomes smaller in each subsequent month. Data are collected by SKY from health centers and hospitals. The chart includes utilization of any household that was a SKY member from January 2008 to December 2009.

Figure 3: Proportion of Households Using SKY-Covered Health Facilities for Care, by Premium and Tenure with SKY



Notes: Of households that were members for at least X months, the above chart shows the proportion that used a health center (HC) or hospital (Hosp.) in that month. Thus, the number of SKY members becomes smaller in each subsequent month. Data are collected by SKY from health centers and hospitals. The chart includes utilization of any household that was a SKY member from January 2008 to December 2009.

Appendix A. Supplementary Tables

TableA.1: Research sample

	Offered Full Price	Offered Deep Discount	Total
Randomized at the Village Meeting	2 618	2 617	5 235
Complete Baseline Surveys, Total	**2 730**	**2 537**	**5 267**
Complete baseline surveys, randomized sample	2 534	2 537	5 071
Complete baseline surveys, over sample	196	N/A	196
SKY Buyers (with SKYID, complete baseline)	**438**	**1 354**	**1 792**
Randomized	243	1 354	1 597
Overs sample	195	N/A	195
SKY Buyers Registered after Meeting Date	420	1 334	1 754
As Percent of Completed Baseline	15.4%	52.6%	68.0%
Bought fewer than 63 days after SKY meeting	240	986	1 226
Bought 63 or more days after SKY meeting	180	348	528

Notes: Households "Randomized at the Village Meeting" refers to the number of households chosen for our randomized survey sample. Of those, "Completed baseline surveys, randomized sample" refers to the number that completed the baseline survey. "Complete baseline surveys, over sample" refers to additional low coupon households that purchased SKY and that completed the baseline survey. SKY "Randomized" buyers were those that are chosen for our randomized sample and also purchased SKY. SKY "Over sample" buyers are those that were not part of the randomized sample but were interviewed to increase the number of households that bought SKY at the full price. SKY buyers can also be broken down into those that purchased soon after the village meeting (within 63 days) and those that purchased sometime after the village meeting (63 days or more). Households that purchased SKY prior to the village meeting (around 43 households), according to SKY records, are not included in the analyses. SKY take-up by full price households appears larger than the usual SKY take-up rate because we include any household that ever purchase SKY following the village meeting. Not all of these households are members at once, some drop out as others join.

TableA.2: Randomization Summary Statistics

	Offered Full Price, Mean	Offered Deep Discount, Mean	Clustered t-test	
Observations	**2 534**	**2 537**		
Highest ranked wealth by enumerator	0.13	0.14	-0.98	
Lowest ranked wealth by enumerator	**0.14**	**0.10**	**3.96**	**
Answered all literacy/numeracy questions correctly	0.15	0.15	0.13	
Household Size	**5.03**	**5.02**	**0.31**	
Education of health decision-maker (years)	4.61	4.72	-1.13	
At least one household member with poor self-reported	**0.70**	**0.72**	**-1.15**	
At least one member over 64	0.25	0.26	-1.11	
No child age 5 or under	**0.55**	**0.57**	**-1.41**	
Household has a stunted or wasted child under age 6	0.16	0.15	0.88	
All vaccines fulfilled for members under 6, 0 if no under 6, pre-SKY	**0.27**	**0.25**	**0.96**	
Major health shock (*) in 2 to 4 months pre-meeting	0.07	0.07	0.07	
Major health shock (*) and used health center for care (0 if no shock)	0.01	0.02	-1.05	
Major health shock (*) and used hospital for care (0 if no shock)	0.02	0.02	0.30	
Major health shock (*) and used private health care (0 if noshock)	0.06	0.05	0.08	
Ln of max. days ill for a major health shock (*), pre-meeting (0 if no shock)	0.24	0.25	-0.46	
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock)	0.05	0.05	-0.26	
Khmer household	**0.953**	**0.946**	**2.00**	.
Ln of approximate value of animals, durables, and business (USD)	6.47	6.49	-0.64	
Ln of approximate value of animals, durables, business, cash, and gold (USD)	**6.68**	**6.74**	**-1.91**	+
Area of farm land owned by household (hectares)	0.81	0.86	-1.05	
Area of village land owned by household (hectares)	**0.14**	**0.13**	**0.90**	
Household has at least one toilet	0.26	0.26	0.34	
House made of palm	**0.04**	**0.03**	**2.23**	*
Roof made of palm	0.05	0.04	1.40	
Roof made of tin	**0.37**	**0.38**	**-0.53**	
Roof made of tile	0.51	0.52	-0.66	
House made of brick	**0.03**	**0.03**	**-0.41**	

*Notes: + p<0.10, * p <0.05, ** p<0.01, *** p<0.001, based on t-tests clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. All data is from the baseline survey. Sample is all SKY decliners and all SKY buyers in the randomized sample.*

TableA.3: Summary Statistics for Selection of Observable Characteristics

	Offered Full Price			Offered Deep Discount			Pooled Households		
	Buyers	Decliners		Buyers	Decliners		Buyers	Decliners	
Observations	**240**	**2 292**		**986**	**1 183**		**1 239**	**3 575**	
Percent Purchasing SKY[1]	9.48%			45.46%			25.74%		
Major health shock (*) in 2 to 4 months pre-meeting	0.104	0.074		0.089	0.055	**	0.094	0.070	*
Major health shock (*) and used health center for care (0 if no shock)	0.013	0.016		0.029	0.011	**	0.027	0.015	*
Major health shock (*) and used hospital for care (0 if no shock)	0.029	0.024		0.023	0.016		0.024	0.022	
Major health shock (*) and used private health care (0 if noshock)	0.083	0.055		0.064	0.045	+	0.069	0.053	*
Ln of max. days ill for a major health shock (*), pre-meeting (0 if no shock)	0.320	0.243		0.314	0.188	**	0.320	0.232	*
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock)	0.063	0.048		0.062	0.035	**	0.063	0.045	*
At least one household member with poor self-reported health	0.863	0.685	**	0.787	0.651	**	0.800	0.672	**
At least one member over 64	0.233	0.244		0.272	0.265		0.264	0.248	
No child age 5 or under	0.533	0.552		0.578	0.571		0.571	0.559	
Household has a stunted or wasted child under age 6	0.179	0.163		0.162	0.154		0.165	0.158	

*Notes: + p<0.10, * p <0.05, ** p<0.01, *** p<0.001, based on t-tests clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. Coupon status is recorded at the village meeting. All other data is from the baseline survey. Sample is all decliners plus households that bought SKY within 63 days of the meeting for which we have baseline data. (1) The low coupon percentage is the percentage including all low coupon households, which includes oversampled low coupon buyers.*

TableA.4: Summary Statistics for Early *versus* Late Buyers

	Bought with Deep Discount					Bought at Full Price				
	Late	N	Early	N		Late	N	Early	N	
PRE-SKY/Pre-Meeting										
At least one household member with poor self-reported health	0.74	348	0.79	986	+	0.81	180	0.86	240	
Major health shock (*) in the 3 months pre-SKY	0.11	134	0.11	986		0.20	106	0.12	240	
Major health shock (*) and use health center for care (0 if no shock)	0.03	134	0.04	986		0.03	106	0.03	240	
Major health shock (*) and use hospital for care (0 if no shock)	0.02	134	0.03	986		0.06	106	0.03	240	
POST-SKY										
Major health shock (*) in the 3 months post-SKY	0.19	69	0.14	679		0.33	61	0.15	175	**
Visited health center in first 3 months post-SKY, Baseline Data	0.06	69	0.06	679		0.08	61	0.06	175	
Visited hospital in first 3 months post-SKY, Baseline Data	0.07	69	0.04	679		0.07	61	0.07	175	
SKY paid for helth center visit in first 3 months of SKY	0.67	304	0.63	958		0.70	64	0.73	175	
SKY paid for hospital visit in first 3 months of SKY	0.08	304	0.12	958	+	0.28	64	0.18	175	

*Notes: + p<0.10, * p <0.05, ** p<0.01, *** p<0.001, based on t-tests clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. Coupon status is recorded at the village meeting. Pre-meeting and pre-SKY variables are from the baseline survey. Self-reported poor health is reported at the time of the baseline survey, which in some cases is after initial SKY purchase. Pre-SKY variables are for the months immediately preceding SKY purchase. Post-SKY major shock variables are from the baseline survey. All other data for post-SKY variables are from SKY data. Sample is all SKY buyers interviewed in the baseline survey. Early means bought within 63 days of the meeting. Late means all other buyers.*

Table A.5: Autocorrelation of Health Expenses

	(1)	(2)	(3)	(4)	(5)
	HC Use	Hosp. Use	Ln Cost HC, USD	Ln Cost Hosp., USD	Ln Total Cost, USD
SKY paid for health center visit in first 3 months of SKY (d)	0.359*** [0.0281]				
SKY paid for hospital visit in first 3 months of SKY (d)		0.182*** [0.0422]			
Ln Cost of SKY-paid health center visits in first 3 months of SKY (USD)			0.602*** [0.0327]		
Ln Cost of SKY-paid hospital visits in first 3 months of SKY (USD)				0.139** [0.0462]	
Ln Cost of all SKY-paid care in first 3 months of SKY (USD)					0.412*** [0.0412]
Observations	1 261	1 261	1 280	1 272	1 271
Adjusted R-squared			0.34	0.013	0.127
Pseudo R-squared	0.09	0.028			

*Notes: Left-hand side variables use SKY utilization data for the 4 to 6 months post-SKY purchase. Covariates use SKY utilization data for the 1 to 3 months post-SKY purchase. Indicator variables are included (not shown) to adjust for households that dropped SKY in months 1 through 6. Cost variables use OLS, all other regressions use probit. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. SKY status is from SKY administrative data. Coupon status recorded at the village meeting. Sample is all deep discount households that bought SKY. (d) for discrete change of indicator variable from 0 to 1.*

Table A.6: Robustness Check: Observable Selection Using Full Sample (Early and Late Buyers)

	(1)	(2)	(3)	(4)	(5)
Offered Full Price (d)	-0.379*** [0.0185]	-0.376*** [0.0188]	-0.379*** [0.0186]	-0.379*** [0.0187]	-0.384*** [0.0192]
At least one household member with poor self-reported health (d)	0.138*** [0.0140]		0.135*** [0.0141]	0.135*** [0.0141]	0.136*** [0.0149]
Major health shock (*) in the 2 to 4 months pre-meeting (d)		0.0761** [0.0286]	0.0528+ [0.0289]	0.139 [0.0968]	0.142 [0.100]
Major health shock (*) and used health center for care (0 if no shock) (d)				-0.0000219 [0.0585]	-0.00463 [0.0599]
Major health shock (*) and used hospital for care (0 if no shock) (d)				-0.00649 [0.0544]	0.00759 [0.0588]
Major health shock (*) and used private health care (0 if no shock) (d)				-0.0503 [0.0526]	-0.0664 [0.0513]
Ln of max. days ill for a major health shock (*) pre-meeting (0 if no shock)				-0.0188 [0.0190]	-0.0148 [0.0198]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)				0.0365 [0.0572]	0.0377 [0.0593]
At least one member over 64 (d)					-0.0172 [0.0177]
At least one member age 5 or under (d)					0.00114 [0.0178]
Household has a stunted or wasted child under age 6 (d)					0.0061 [0.0228]
Observations	5 229	5 229	5 229	5 229	4 871
Pseudo R-squared	0.141	0.129	0.142	0.142	0.147

*Notes: LHS variable: 1 if bought SKY, 0 if declined (SKY Admin data). + p<0.10, * p<0.05, ** p<0.01, r*** p<0.001. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. All data are from the baseline survey. Wealth, household size and education are included in the regression but not presented. Sample is all SKY decliners and buyers. (d) for discrete change of indicator variable from 0 to 1.*

Table A.7: Robustness Check: Observable Selection Using Only Health Shocks Lasting 7 or More Days

	(1)	(2)	(3)	(4)	(5)
Offered Full Price (d)	-0.362*** [0.0171]	-0.361*** [0.0172]	-0.362*** [0.0171]	-0.362*** [0.0171]	-0.368*** [0.0172]
At least one household member with poor self-reported health (d)	0.124*** [0.0125]		0.121 *** [0.0125]	0.121 *** [0.0125]	0.127*** [0.0131]
Major health shock (*) in the 2 to 4 months pre-meeting (d)		0.0985** [0.0348]	0.0756* [0.0343]	0.0915 [0.0701]	0.0837 [0.0688]
Major health shock (*) and used health center for care (0 if no shock) (d)				0.0307 [0.0544]	0.025 [0.0560]
Major health shock (*) and used hospital for care (0 if no shock) (d)				-0.0244 [0.0518]	-0.00884 [0.0556]
Major health shock (*) and used private health care (0 if no shock) (d)				-0.019 [0.0442]	-0.0301 [0.0447]
Ln of max. days ill for a major health shock (*) pre-meeting (0 if no shock)				-0.00784 [0.0170]	-0.00221 [0.0176]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)				0.0461 [0.0593]	0.0372 [0.0597]
At least one member over 64 (d)					-0.0194 [0.0155]
At least one member age 5 or under (d)					-0.00158 [0.0163]
Household has a stunted or wasted child under age 6 (d)					0.0167 [0.0208]
Observations	4 701	4 701	4 701	4 701	4 387
Pseudo R-squared	0.168	0.154	0.169	0.169	0.179

*Notes: LHS variable: 1 if bought SKY, 0 if declined (SKY Administrative data). + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. All data are from the base line survey. Wealth, household size and education are included in the regression but not presented. Sample is all SKY decliners and all SKY buyers. (d) for discrete change of indicator variable from 0 to 1.*

Table A.8: Robustness Check: Unobservable Selection Using All Observable Covariates

	(1) Use HC	(2) Use Hosp.	(3) Total Cost
Offered Full Price (d)	0.117*** [0.0344]	0.0733* [0.0357]	0.468*** [0.134]
At least one household member with poor self-reported health (d)	0.132*** [0.0364]	0.0158 [0.0256]	0.372*** [0.102]
Major health shock (*) in the 2 to 4 months pre-meeting (d)	0.0915 [0.156]	0.0334 [0.127]	-0.0993 [0.512]
Major health shock (*) and used health center for care (0 if no shock) (d)	0.167* [0.0837]	0.107 [0.0970]	0.265 [0.272]
Major health shock (*) and used hospital for care (0 if no shock) (d)	0.199** [0.0674]	-0.0391 [0.0718]	0.0987 [0.261]
Major health shock (*) and used private health care (0 if no shock) (d)	0.106 [0.109]	0.0855 [0.115]	0.429 [0.312]
Ln of max. days ill for a major health shock (*) pre-meeting (0 if no shock)	-0.0763+ [0.0454]	-0.00551 [0.0277]	0.0147 [0.123]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)	0.0977 [0.0996]	-0.0618 [0.0543]	-0.173 [0.304]
At least one member over 64 (d)	-0.0284 [0.0332]	0.00983 [0.0221]	-0.132 [0.0913]
At least one member age 5 or under (d)	0.0561 [0.0511]	0.00233 [0.0375]	0.127 [0.160]
Household has a stunted or wasted child under age 6 (d)	0.0199 [0.0507]	-0.0374 [0.0283]	0.0641 [0.146]
Observations	**1 199**	**1 199**	**1 199**
Adjusted R-squared			
Pseudo R-squared	**0.1**	**0.079**	**0.037**

*Notes: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. LHS variables: Column 1 (2): Indicator for use of a SKY-covered health center (hospital) for the first 3 months following SKY purchase; Column 3: Ln of total cost (user fees, covered by SKY) of all SKY-covered health center and hospital visits in the first 3 months post-SKY. Columns 1-2 use probit, column 3 uses OLS. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. SKY status and LHS variables use SKY data. Coupon status was recorded at the village meeting. All other data are from the baseline survey. Wealth, household size and education, and additional variables as described in the Appendix, are included in the SKY regression but not presented. Sample is all SKY decliners and all SKY buyers who bought SKY following the village meeting (d) for discrete change of indicator variable from 0 to 1.*

Table A.9: Robustness Check: Unobservable Selection Using All Covariates, and Indicator for Early Purchase

	(1) Use HC	(2) Use Hosp.	(3) Total Cost
Offered Full Price (d)	0.113** [0.0348]	0.0742* [0.0352]	0.333** [0.112]
Joined 62 or fewer days after the village meeting (d)	-0.0567 [0.0467]	0.0076 [0.0269]	-0.0243 [0.0900]
At least one household member with poor self-reported health (d)	0.136*** [0.0366]	0.0153 [0.0257]	0.216** [0.0667]
Major health shock (*) in the 2 to 4 months pre-SKY (d)	0.09 [0.155]	0.0352 [0.129]	-0.12 [0.401]
Major health shock (*) and used health center for care (0 if no shock) (d)	0.171* [0.0823]	0.106 [0.0972]	0.178 [0.213]
Major health shock (*) and used hospital for care (0 if no shock) (d)	0.204** [0.0658]	-0.0395 [0.0712]	-0.0408 [0.205]
Major health shock (*) and used private health care (0 if no shock) (d)	0.103 [0.109]	0.0865 [0.115]	0.357 [0.236]
Ln of max. days ill for a major health shock (*) pre-SKY (0 if no shock)	-0.0760+ [0.0452]	-0.00575 [0.0280]	0.0495 [0.0925]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)	0.101 [0.0987]	-0.0621 [0.0540]	-0.225 [0.245]
At least one member over 64 (d)	-0.0281 [0.0331]	0.00975 [0.0220]	-0.116+ [0.0659]
At least one member age 5 or under (d)	0.0535 [0.0513]	0.00269 [0.0376]	0.0544 [0.117]
Household has a stunted or wasted child under age 6 (d)	0.0224 [0.0507]	-0.0375 [0.0283]	0.0513 [0.113]
Observations	1 199	1 199	1 199
Adjusted R-squared			0.05
Pseudo R-squared	0.102	0.08	

*Notes: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. LHS variables: Column 1 (2): Indicator for use of a SKY covered health center (hospital) for the first 3 months post-SKY purchase; Column 3: Ln of total cost (user fees, covered by SKY) of all SKY-covered health center and hospital visits in the first 3 months post-SKY. Columns 1-2 use probit, column 3 uses OLS. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. SKY status and LHS variables use SKY data. Coupon status was recorded at the village meeting. All other data are from the baseline survey. Wealth, household size and education, and additional variables as described in the Appendix, are included in the regression but not presented. Sample is all SKY decliners and all SKY buyers who bought SKY following the village meeting (d) for discrete change of indicator variable from 0 to 1. This regression includes an indicator variable for joining within 63 days of the village meeting.*

Table A.10: Robustness Check: Unobservable Selection Controlling for Health Shocks 1-3 Months Pre-Baseline

	(1) Use HC	(2) Use Hosp.	(3) Total Cost
Offered Full Price (d)	0.114*** [0.0339]	0.0842* [0.0364]	0.312** [0.109]
At least one household member with poor self-reported health (d)	0.180*** [0.0354]	0.0173 [0.0261]	0.272*** [0.0606]
Major health shock (*) in the 3 months pre-SKY (d)	0.0216 [0.175]	-0.00395 [0.0903]	-0.0864 [0.311]
Major health shock (*) and used health center for care (0 if no shock) (d)	0.198** [0.0688]	-0.0561 [0.0371]	0.0329 [0.169]
Major health shock (*) and used hospital for care (0 if no shock) (d)	0.137+ [0.0828]	-0.0615 [0.0385]	-0.255 [0.169]
Major health shock (*) and used private health care (0 if no shock) (d)	0.173* [0.0850]	-0.02 [0.0600]	0.122 [0.208]
Ln of max. days ill for a major health shock (*) pre-SKY (0 if no shock)	-0.0521 [0.0433]	0.0287 [0.0251]	0.0531 [0.0817]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)	0.0237 [0.0873]	0.0047 [0.0627]	0.106 [0.185]
At least one member over 64 (d)	-0.0289 [0.0323]	0.00586 [0.0232]	-0.123+ [0.0646]
At least one member age 5 or under (d)	0.0875* [0.0358]	0.000386 [0.0268]	0.104 [0.0793]
Household has a stunted or wasted child under age 6 (d)	0.0601 [0.0473]	-0.0358 [0.0289]	0.0918 [0.102]
Observations	**1 215**	**1 215**	**1 215**
Adjusted R-squared			0.045
Pseudo R-squared	**0.055**	**0.028**	

*Notes: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. LHS variables: Column 1 (2): Indicator for use of a SKY-covered health center (hospital) for the first 3 months post-SKY purchase; Column 3: Ln of total cost (user fees, covered by SKY) of all SKY-covered health center and hospital visits in the first 3 months post-SKY. Columns 1-2 use probit, column 3 uses OLS. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. SKY status and LHS variables use SKY data. Coupon status was recorded at the village meeting. All other data are from the baseline survey. Wealth, household size and education, and additional variables as described in the Appendix, are included in the regression but not presented. Sample is all SKY decliners and all SKY buyers who bought SKY following the village meeting (d) for discrete change of indicator variable from 0 to 1. These regressions control for health shocks 1 to 3 months prior to joining SKY rather than 2 to 4 months prior to joining SKY.*

Table A.11: Robustness Check: Unobservable Selection, No Oversampled Households

	(1) Use HC	(2) Use HC	(3) Use Hosp.	(4) Use Hosp.	(5) Total Cost	(6) Total Cost
Offered Full Price (d)	0.0563 [0.0427]	0.109* [0.0449]	0.108* [0.0429]	0.0925+ [0.0476]	0.304** [0.114]	0.313* [0.135]
At least one household member with poor self-reported health (d)		0.155*** [0.0359]		0.0183 [0.0267]		0.252*** [0.0613]
Major health shock (*) in the 2 to 4 months pre-SKY (d)		0.00618 [0.183]		0.146 [0.174]		0.041 [0.444]
Major health shock (*) and used health center for care (0 if no shock) (d)		0.171* [0.0802]		0.0356 [0.0765]		0.0473 [0.205]
Major health shock (*) and used hospital for care (0 if no shock) (d)		0.151+ [0.0893]		0.0644 [0.0513]		0.0472 [0.222]
Major health shock (*) and used private health care (0 if no shock) (d)		0.0494 [0.114]		0.0171 [0.0867]		0.168 [0.234]
Ln of max. days ill for a major health shock (*) pre-SKY start (0 if no shock)		0.0451 [0.0441]		0.0254 [0.0281]		0.0115 [0.102]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)		0.150+ [0.0776]		0.0171 [0.0626]		0.0466 [0.231]
At least one member over 64 (d)		0.0124 [0.0332]		0.00777 [0.0227]		0.0867 [0.0623]
At least one member age 5 or under (d)		0.0926* [0.0385]		1.8E05 [0.0273]		0.101 [0.0847]
Household has a stunted or wasted child under age 6 (d)		0.0757 [0.0483]		0.0518* [0.0256]		0.0596 [0.101]
Observations	**1 380**	**1 141**	**1 380**	**1 141**	**1 383**	**1 141**
Adjusted R-squared					0.007	0.034
Pseudo R-squared	**0.001**	**0.046**	**0.01**	**0.03**		

*Notes: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. LHS variables: Columns 1 and 2 (3 and 4): Indicator for use of a SKY-covered health center (hospital) for the first 3 months post-SKY purchase; Columns 5 and 6: Ln of total cost (user fees, covered by SKY) of all SKY-covered health center and hospital visits in the first 3 months post-SKY. Columns 1-4 use probit, columns 5-6 use OLS. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. SKY status and LHS variables use SKY data. Coupon status was recorded at the village meeting. All other data are from the baseline survey. Wealth, household size and education are included in the regression. All other LHS variables use probit regressions but are not presented. Sample is all SKY decliners and all SKY buyers who bought SKY following the village meeting, only if they were part of the randomized sample (no over-sampled low coupon buyers included). (d) for discrete change of indicator variable from 0 to 1.*

Table A.12: Robustness Check: Unobservable Selection, Tobit Regression for Costs

	(1) Total Cost, Tobit	(2) Total Cost, Tobit
Offered Full Price	0.565*** [0.130]	0.458*** [0.132]
At least one household member with poor self-reported health (d)		0.447*** [0.0958]
Major health shock (*) in the 2 to 4 months pre-SKY (d)		-0.161 [0.455]
Major health shock (*) and used health center for care (0 if no shock)		0.174 [0.238]
Major health shock (*) and used hospital for care (0 if no shock)		-0.0409 [0.265]
Major health shock (*) and used private health care (0 if no shock)		0.126 [0.300]
Ln of max. days ill for a major health shock (*) pre-SKY start (0 if no shock)		0.0307 [0.113]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock)		0.209 [0.274]
At least one member over 64		-0.126 [0.0888]
At least one member age 5 or under		0.201+ [0.108]
Household has a stunted or wasted child under age 6		0.136 [0.130]
Observations	**1 508**	**1 255**
Adjusted R-squared		
Pseudo R-squared	**0.009**	**0.024**

*Notes: +p<0.10, *p<0.05, ** p<0.01, ***p<0.001. LHS variables: Ln of total cost (user fees, covered by SKY) of all SKY-covered health center and hospital visits in the first months post-SKY, using tobit regressions. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates a health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. SKY status and LHS variables use SKY data. Coupon status was recorded at the village meeting. All other data are from the baseline survey. Wealth, household size and education are included in the regression but not presented. Sample is all SKY decliners and all SKY buyers who bought SKY following the village meeting, only if they were part of the randomized sample (no over-sampled low coupon buyers included). (d) for discrete change of indicator variable from 0 to 1.*

Table A.13: Robustness Check: Unobservable Selection, Only Shocks Lasting More than 7 Days

	(1) Use HC	(2) Use HC	(3) Use Hosp.	(4) Use Hosp.	(5) Total Cost	(6) Total Cost
Offered Full Price (d)	0.111*** [0.0321]	0.114*** [0.0328]	0.107** [0.0342]	0.0860* [0.0360]	0.414*** [0.106]	0.328** [0.109]
At least one household member with poor self-reported health (d)		0.163*** [0.0344]		0.0198 [0.0252]		0.253*** [0.0592]
Missed 7 or more days of main activity due to illness, 2 to 4 months pre-SKY (d)		0.190* [0.0904]		0.0645 [0.132]		0.211 [0.337]
Major health shock (*) and used health center for care (0 if no shock) (d)		0.141+ [0.0820]		0.0624 [0.0810]		0.0186 [0.184]
Major health shock (*) and used hospital for care (0 if no shock) (d)		0.107 [0.0899]		-0.0619 [0.0541]		-0.165 [0.190]
Major health shock (*) and used private health care (0 if no shock) (d)		-0.0082 [0.111]		0.0188 [0.0877]		0.00855 [0.239]
Ln of max. days ill for a major health shock (*) pre-SKY start (0 if no shock)		-0.0761* [0.0361]		-0.0177 [0.0284]		-0.0189 [0.0841]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)		0.150* [0.0746]		0.00028 [0.0699]		0.116 [0.207]
At least one member over 64 (d)		-0.0251 [0.0316]		0.00606 [0.0228]		-0.115+ [0.0631]
At least one member age 5 or under (d)		0.0846* [0.0347]		-0.0025 [0.0264]		0.109 [0.0774]
Household has a stunted or wasted child under age 6 (d)		0.068 [0.0455]		-0.0371 [0.0284]		0.0894 [0.101]
Observations	**1 505**	**1 255**	**1 505**	**1 255**	**1 508**	**1 255**
Adjusted R-squared					0.024	0.045
Pseudo R-squared	**0.009**	**0.05**	**0.017**	**0.029**		

*Notes: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001. LHS variables: Columns 1 and 2 (3 and 4): Indicator for use of a SKY-covered health center (hospital) in the first 3 months post-SKY purchase; Columns 5 and 6: Ln of total cost (user fees, covered by SKY) of all SKY-covered health center and hospital visits in the first 3 months post-SKY. Columns 1-4 use probit, columns 5-6 use OLS. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates major health shock causing missed daily activities for 7 or more days. SKY status and LHS variables use SKY data. Coupon status was recorded at the village meeting. All other data are from the baseline survey. Wealth, household size and education are included in the regression but not presented. Sample is all SKY decliners and all SKY buyers who bought SKY following the village meeting. (d) for discrete change of indicator variable from 0 to 1.*

Table A.14: Robustness Check: Unobservable Selection, Inpatient Visits

	(1) Impatient Visit	(2) Impatient Visit
Offered Full Price	0.0780** [0.0257]	0.0571 * [0.0236]
At least one household member with poor self-reported health (d)		0.0303* [0.0120]
Major health shock (*) in the 2 to 4 months pre-SKY (d)		0.0141 [0.0669]
Major health shock (*) and used health center for care (0 if no shock) (d)		0.0264 [0.0517]
Major health shock (*) and used hospital for care (0 if no shock) (d)		-0.0281 [0.0220]
Major health shock (*) and used private health care (0 if no shock) (d)		-0.00273 [0.0398]
Ln of max. days ill for a major health shock (*) pre-SKY start (0 if no shock)		0.00131 [0.0141]
Major health shock (*) and spent 120,000 riels on care (US$30) (0 if no shock) (d)		-0.00699 [0.0292]
At least one member over 64 (d)		-0.0118 [0.0130]
At least one member age 5 or under (d)		0.0167 [0.0175]
Household has a stunted or wasted child under age 6 (d)		-0.00531 [0.0187]
Observations	**1 487**	**1 239**
Pseudo R-squared	0.027	0.055

*Notes: + p<0.10, * p<0.05, **rp<0.01, ***p<0.001. LHS variables: Indicator for an inpatient visit to a SKY-covered hospital in the first 3 months post-SKY purchase, using probit regression. Marginal effects; standard errors in brackets. Robust standard errors clustered at the village level. (*) indicates health shock causing missed daily activities for 7 or more days, a death, or an expense of over US$100. SKY status and LHS variables use SKY data. Coupon status was recorded at the village meeting. All other data are from the baseline survey. Wealth, household size and education are included in the regression but not presented. Sample is all SKY decliners and all SKY buyers who bought SKY following the village meeting. (d) for discrete change of indicator variable from 0 to 1.*

Appendix B. Other Datasets

B.1. Village Leader Interview

In each village we interviewed the village chief or another village leader in order to collect general village-level information, including the travel time and cost to get to the nearest public health center; recent village-level shocks (drought, flood, epidemics, etc.); availability of lending institutions; and price and availability of paid transportation.

Selected variables from this survey are used only as controls in the unobserved selection analysis, and are described in the Appendix to Polimeni and Levine, 2011.

B.2. Health Center Data Collection

Households may be more likely to purchase SKY if the quality of the local public health clinic with which SKY partners is good. To measure this, we administered a simple survey of health clinics in areas covered by our study.

To minimize data collection costs, the health center survey focuses on observations by SKY member facilitators. SKY hires member facilitators to be present at health facilities to facilitate treatment for SKY members and manage client complaints and questions as needed. Member facilitators typically work mornings at one particular health center. The survey consists of checklists of operating hours, drug supply, cleanliness and equipment supply. In the current paper, these variables are used only as controls in the selection on unobservables analysis. These variables are listed in the Appendix to Polimeni and Levine, 2011.

B.3. Village Meeting Data

At the end of each village meeting, our field team spoke to a village leader to collect village-level data relevant to our study. In the results presented in this paper, we use these variables only as controls in the analysis of selection on unobservables. The Appendix to Polimeni and Levine, 2011, lists variables created from responses to these interviews.

B.4. Lucky Draw Implementation

To implement the Lucky Draw, attendance was taken at the beginning of each meeting, and names of people who arrived at the meeting late were added to the roster as they arrived. We collected one name for each household represented at the meeting so that households with more than one member attending were not counted more than once. While a SKY representative conducted the meeting, staff from our field team counted out the number of high and low coupons to be distributed to the meeting attendees. The number of high coupons was set to equal 20% of households with a cap of 12 per meeting. The high coupons were put into a bag along with enough low coupons to cover all other households attending the meeting.

At the end of the meeting, the research team's field coordinator announced that there would be a raffle where the prize is a high-value coupon for insurance and explained the coupon rules. Family names were called out one by one from the roster. When a name was called out, a representative from the household came to the front of the room and pulled a coupon from the bag. High coupons were brightly colored so that everyone could see what coupon was drawn after the fact but care was taken to ensure people could not see coupons as they were drawing. As coupons were drawn, the names of households who received a high coupon were recorded so that coupons could not be traded and used by other households.

Appendix C. Description of Variables, Adverse Selection

The following variables were used on the left-hand side of the main specifications.

Table C.1: Independent Variables

Independent Variable Name	Description
Purchased SKY	1 if purchased SKY, 0 if not (observable selection regression)
Health Center Use, first 3 months	1 if any household member used a SKY-partnered health center in the first 3 months after insurance purchase, 0 if not (SKY data) (unobservable selection regression)
Hospital Use, first 3 months	1 if any household member used a SKY-partnered hospital in the first 3 months after insurance purchase, 0 if not (SKY data) (unobservable selection regression)
Cost of Health Center and Hospital Visits, first 3 months	Log of total cost of visits to a SKY-partnered health center or hospital in the first 3 months after insurance purchase, by any household member (SKY data) (unobservable selection regression)

The following is a description of the basic controls used in regressions measuring adverse selection. Additional controls used in robustness tests can be found in the Appendix to Polimeni and Levine, 2011.

Table C.2: Basic Covariates

Variable Name	Questionnaire Question	Description
Subjective poor health	How healthy is each household member? (Excellent health, good health, poor health). Primary respondent to questionnaire gives subjective response for all household members.	1 if respondent describes health of any household member as "poor"
Major health shock, 2-4 months pre-meeting. (We use shocks 1-3 months pre-meeting as a robustness check.)	Three questions: In the last year, were there any health problems in your household that made someone unable to work or go to school for one week or more? In the last year, did anyone in your household pass away? In the last year, did anyone in your household spend more than 400,000 riels (US$100) on a single health problem?	1 if respondent answers "yes" to any of these three health questions, AND the month of the health shock was 2-4 months prior to the date of the SKY meeting
Visited a health center for a major health shock, 2-4 months pre-meeting	[If household member experienced a major shock in 2-4 months pre-meeting:] Did [sick member] seek treatment for this health problem? If yes, where? [Respondent chose "health center"]	1 if, following a major health shock in the 2-4 months pre-meeting, a household member visited a public health center for first or subsequent treatment
Visited a public hospital for a major health shock, 2-4 months pre-meeting	[If household member experienced a major shock in 2-4 months pre-meeting:] Did [sick member] seek treatment for this health problem? If yes, where? [Respondent chose "public hospital"]	1 if, following a major health shock in the 2-4 months pre-meeting, a household member visited a public hospital for first or subsequent treatment
Visited a private facility for a major health shock, 2-4 months pre-meeting	[If household member experienced a major shock in 2-4 months pre-meeting:] Did [sick member] seek treatment for this health problem? If yes, where? [Respondent chose "private doctor (village or town)"]	1 if, following a major health shock in the 2-4 months pre-meeting, a household member visited a private doctor for first or subsequent treatment

Variable Name	Questionnaire Question	Description
Spent more than US$30 (120,000 riels) for a major health shock, 2-4 months pre-meeting	[If household member(s) treated for a major shock in 2-4 months pre-SKY:] What was the cost of treating the health problem (at any facility)?	Sum of treatment costs at any facility (including traditional healers, drug sellers, etc.) for all household members experiencing a major shock in the 2-4 months pre-SKY
Max. number of days ill, 2-4 months pre-meeting	[If household member experienced major shock in 2-4 months pre-meeting:] How many days was he/she unable do his/her usual activities because of this health problem?	Maximum days ill for any single health problem for any household member in the 2-4 months before the SKY meeting
Household has a stunted or wasted child	Height, age, and weight measured for all children age 5 and under	1 if household has a child that is stunted or wasted (z-score for height-for-age or weight-for-height is less than -2) according to WHO growth standards, 0 otherwise (including if household has no children age 5 or under)
Household has a member age 65 or older	Date of birth of each household member	1 if any household member is age 65 or older
Household has at least one member age 5 or under	Date of birth of each household member	1 if any household member is age 5 or under, 0 otherwise
Household size	Household roster: Names of people who usually sleep here (slept in the house 5 out of the 7 nights immediately preceding the interview)	Number of household members listed in the household roster
Poor household	Enumerator subjective wealth ranking: poorest/medium/better-off	1 if enumerator rates household as poor
Better-off household	Enumerator subjective wealth ranking: poorest/medium/better-off	1 if enumerator rates household as better-off
Respondent is literate and numerate	(Round 2 survey) Four literacy/numeracy questions: Draw a line from each picture to the correct word; Write the name of the village, commune and district where you live; Write the correct number of objects in the pictures, and what each object is; Tell me what time it is (picture of a clock shown)	1 if respondent answers all literacy and numeracy questions correctly, 0 otherwise
Education of health decision-maker (years)	Who makes the decisions about health care in your family? What is the highest grade this person completed? What is the highest grade you completed?	Education from 1 to 13 (13 = tertiary education). If respondent decides with another household member, use maximum education of the two members. Indicator variables for 0 years or 1 to 4 years used in regressions.

References

AKERLOF, G. A. (1970), "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism", *The Quarterly Journal of Economics*, 84(3), pp. 488-500.

ANNEAR, P. (2006), "Study of Financial Access to Health Services for the Poor in Cambodia. Phnom Penh", Research Report, Cambodia Ministry of Health, WHO, AusAID, RMIT University.

ANTOS, J. R. (2007), "Health Care Financing in Thailand: Modeling and Sustainability", Mission Report, World Bank.

AUSUBEL, L. (1999), "Adverse Selection in the Credit Card Market", University of Maryland.

CARDON, J.H. AND I. HENDEL (2001), "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey", *The RAND Journal of Economics*, 32(3), pp. 408-427.

CAWLEY, J. AND T. PHILIPSON (1999), "An Empirical Examination of Information Barriers to Trade in Insurance", *The American Economic Review*, 89(4), pp. 827-846.

CENTRAL INTELLIGENCE AGENCY (2010), "The CIA World Factbook", https://vTww.cia.gov/library/publications/the-worldfactbook/.

COHEN, J. AND P. DUPAS (2010), "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment", *The Quarterly Journal of Economics*, 125(1), 1-45.

COLLINS, W. (2000), "Medical Practitioners and Traditional Healers: A Study of Health Seeking Behavior in Kampong Chhnang, Cambodia", Discussion Paper.

CUTLER, D.M. AND S.J. REBER (1998), "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection", *The Quarterly Journal of Economics*, 113(2), pp. 433-466.

CUTLER, D.M. AND R.J. ZECKHAUSER (1998), "Adverse Selection in Health Insurance", *Forum for Health Economics & Policy*, 1, 2.

CUTLER, D.M. AND R.J. ZECKHAUSER (2000), "Chapter 11. The Anatomy of Health Insurance", Vol. 1, Part 1 of *Handbook of Health Economics*, pp. 563-643. Elsevier.

DHS (2005), "DHS Demographic and Health Survey, Cambodia", http://www.measuredhs.com.

ELLIS, R.P. (1989), "Employee Choice of Health Insurance", *The Review of Economics and Statistics*, 71(2), pp. 215-223.

FANG, H., M.P. KEANE AND D. SILVERMAN (2008), "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market", *Journal of Political Economy*, 116(2), pp. 303-350.

FINKELSTEIN, A. AND K. MCGARRY (2006), "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market", *The American Economic Review*, 96(4), pp. 938-958.

FINKELSTEIN, A. AND J. POTERBA (2004), "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market", *Journal of Political Economy*, 112(1), pp. 183-208.

JUTTING, J.P. (2004), "Do Community-based Health Insurance Schemes Improve Poor People's Access to Health Care? Evidence From Rural Senegal", *World Development*, 32(2), 273-288.

KARLAN, D. AND J. ZINMAN (2009), "Observing Unobservables: Identifying Information Asymmetries With a Consumer Credit Field Experiment", *Econometrica*, 77(6), 1993-2008.

LAMIRAUD, K., F. BOOYSEN AND X. SCHEIL-ADLUNG (2005), "The Impact of Social Health Protection on Access to Health Care, Health Expenditure and Impoverishment: A Case Study of South Africa", International Labour Office, Department of Social Security, Extension of Social Security N° 23, 2005.

LEVINE, D.I., R. GARDNER AND R. POLIMENI (2009), "Briefing Paper: A Literature Review on Effects of Health Insurance and Selection into Health Insurance", Discussion Paper.

LEVINE, D.I., R. POLIMENI AND I. RAMAGE (2011), "A Rigorous Evaluation of SKY Health Insurance", Discussion Paper.

PEOPLE'S DAILY ONLINE (2008), "China's New Rural Cooperative Medical Plan", http://english.peopledaily.com.cn/90002/95607/6524401.html.

POLIMENI, R. (2006), "Adverse Selection and Impacts of Health Insurance in a Developing Country: Evidence from a Randomized Experiment in Cambodia", Research Prospectus, University of California, Berkeley.

POLIMENI, R. AND D.I. LEVINE (2011), "Going Beyond Adverse Selection: Take-up of a Health Insurance Program in Rural Cambodia", Working Paper, University of California, Berkeley.

ROTHSCHILD, M. AND J. STIGLITZ (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *The Quarterly Journal of Economics*, 90(4), pp. 629- 649.

SIEGELMAN, P. (2004), "Adverse Selection in Insurance Markets: An Exaggerated Threat", *The Yale Law Journal*, 113(6), pp. 1223-1281.

SKY (2009), "SKY Health Insurance Schemes", www.sky-cambodia.org/efirstrestuls.html.

TAROZZI, A., A. MAHAJAN, B. BLACKBURN, D. KOPF, L. KRISHNAN AND J. YOONG (2011), "Micro-Loans, Insecticide-Treated Bednets and Malaria: Evidence from a Randomized Controlled Trial in Orissa (India)", Economic Research Initiatives at Duke (ERID) Working Paper N° 104.

VAN DAMME, W., L. VAN LEEMPUT, I. POR, W. HARDEMAN AND B. MEESSEN (2004), "Out-of-pocket Health Expenditure and Debt in Poor Households: Evidence from Cambodia", *Tropical Medicine and International Health*, 9(2), 273-280.

VIETNAM SOCIAL SECURITY (2010), "History Development and Expanding Program of Voluntary Health Insurance in Vietnam", http://vTww.asean-ssa.org.

WAGSTAFF, A., M. LINDELOW, G. JUN, X. LING AND Q. JUNCHENG (2009), "Extending Health Insurance to the Rural Population: An Impact Evaluation of China's New Cooperative Medical Scheme", *Journal of Health Economics*, 28(1), 1-19.

WAGSTAFF, A. AND M. PRADHAN (2005), "Health Insurance Impacts on Health and Nonmedical Consumption in a Developing Country", *Policy Research Working Paper Series* 3563, World Bank.

WANG, Y. (2007), "Development of the New Rural Cooperative Medical System in China", *China & World Economy*, 15(4), 66-77.

WOLFE, J.R. AND J.H. GODDEERIS (1991), "Adverse Selection, Moral Hazard, and Wealth Effects in the Medigap Insurance Market", *Journal of Health Economics*, 10(4), 433-459.