Adverse Selection in Health Insurance Purchasing in Cambodia: Evidence from SKY Micro-Health Insurance Program and Health Equity Fund Data

Hongyu (Karen) Zhang

May 1st 2012

Advisor: Professor David I. Levine

Department of Economics

University of California, Berkeley

Acknowledgements: I would like to thank Professor David Levine for his guidance, support, and encouragement through every step of the research and writing process. I am infinitely grateful for his mentoring in the past two years. I would also like to thank the staff at GRET and SKY sharing their data and Domrei for data collection. Thanks also to Rachel Polimeni and Xuam-Mein for their great assistance in clarifying the data.

Abstract

The economic theory of adverse selection predicts that those who purchase insurance are those who anticipate high future health care costs. However, few studies of the selection problem have been conducted in developing countries, and empirical evidence for adverse selection is mixed. A recent study based on the SKY Macro-Health Insurance Program and Health Equity Fund (HEF) in Cambodia shows that voluntary SKY insurance purchasers have higher health care utilization than HEF members, who are pre-selected and are fully subsided by a pooled fund from the Cambodian Government and NGOs. We are interested in the composition of the persistent utilization gap, especially the contribution from adverse-selected voluntary members who pay a higher premium. The randomized coupon distribution scheme in SKY provides us with a useful tool to detect adverse selection among voluntary members and to separate the selection problem from other possible causes of the utilization gap. With data provided by SKY administration and HEF, this paper will test why HEF members may use health care services less than voluntary insurance members. Important variables include remoteness, poverty, district and voluntary member self-selection. This research will provide a platform upon which further studies of cost-effective methods of improving health care in developing nations can be built.

1. Introduction

Health insurance can increase access to health care and can protect people against financial burdens resulting from potentially catastrophic illness or injury. In developing countries, where the economy relies heavily on labor-intensive activities, the health benefit of an insurance policy can be even more pronounced. It can break the vicious circle in which a sick person is denied access to appropriate health care, causing their health stock to deteriorate, and thus further limiting their financial capability to purchase medical treatment. Therefore, promoting health insurance in developing countries has a potentially high marginal benefit in the sense that it can not only protect the poor against unnecessary health-stock and financial loss, but also has the potential to break the cycle of poverty in the long term.

Adverse selection in insurance purchasing nevertheless can cause inefficiency in allocating benefits, and can even make insurance schemes financially unsustainable. The economic theory of insurance demand predicts that households that expect high health care expenditures are more likely to take up insurance. Hence, by offering a steep premium discount, insurance companies are able to encourage healthier people to join the insurance pool and thereby to reduce overall adverse selection (George Akerlof, 1970). Based on SKY Macro-Health Insurance program in Cambodia, a randomized experiment was carried out in 2008. Survey data, utilization data and membership data has been collected as SKY expands. One advantage of using SKY data is that randomization of the distribution of insurance premium coupons allows researchers to identify adverse selection from other effects such as moral hazard. Strong evidence is found that adverse selection exists among SKY Voluntary Members (VMs) (Polimeni and Levine, 2011). In particular, SKY members paying full price for insurance exhibit higher utilization than those who pay at steep discounts.

A joint study of the Health Equity Fund also finds that HEF members, who are fully subsidized for health insurance, also have lower health care utilization rates than VMs. At first glance, the utilization gap seems to be a result of adverse selection in the VMs. Because HEF members subscribe insurance for free, they are theoretically the least adversely selected subscribers. Thus it is reasonable for HEF members to use less health service than the more adversely selected VMs. However, we should be careful about drawing any conclusions for the following reasons. First, economic theories about adverse selection primarily assume joining the insurance pool is done on a voluntary basis, whereas HEF members are selected by Ministry of Health, which disqualifies this assumption. Second, HEF is exclusively operated in Kampot, while SKY insurance is available in Kampot, Kandal and Takeo. Comparing utilization of all SKY members with that of HEF members tends to overlook factors such as regional differences and fixed effects. Third, VMs can afford health insurance regardless of whether or not they obtain premium discounts, while HEF members are incapable of purchasing insurance even at discounted rates. The wealth effect on HEF members, who are granted free insurance subscription, and the wealth effect on those VMs that pay full price but could have paid at a lower rate are therefore very different. These effects can drive the utilization gap in opposite directions, and we need at least to know the dominant driving force behind the utilization gap in order to apply theories to explain that gap.

In addition to the wealth effect puzzle, another study based on high-frequency logbook data finds no support of adverse selection among newly joined SKY members (Levine and Zhang, 2011), a result contrasting with that of Polimeni and Levine. According to Levine and Zhang, households that recently join SKY are not at higher medical risk than existing members. If high utilization is not induced by medical risk and other unobservable factors, it is worth investigating the cause of the utilization gap, which may shed light on solutions that will narrow the gap.

In this paper, we define a baseline test that a household's decision on health care utilization is a function of HEF membership. The coefficient of HEF in the base line function reflects the utilization gap between HEF members and all VMs. We then break down the utilization gap by adding covariates and examining the changes of the HEF coefficient. We find that the gap of average utilization between HEF members and VMs are firstly driven by the HEF members' inferior financial conditions, which makes them unable to afford nonmedical expenditures. Evidence also supports the hypothesis that HEF members live in more remote villages, and that transportation inconvenience generates barriers to accessing public care. Moreover, the utilization gap is narrower if we restrict our sample to the insured in Kampot district, which suggests VMs in Kampot are less adversely selected than in other districts. Finally, we find strong evidence that self-selection is the main driving force creating the utilization gap.

2. Previous Research

In article "The Market for Lemons" (Akerlof 1970) George Akerlof examines why health insurance companies do not raise their rate to match the risk of clines. Akerlof pointed out that individuals who are willing to pay insurance premium are those who expect high insurance payouts. Transferring this theory to health insurance, insurance purchasers are individuals who are more knowledge about their health conditions and expect a higher medical expenditure than insurance premium.

Many studies have found evidence of adverse selection in health insurance purchasing. Cutler and Zeckhauser show that people with higher expected medical expenditure are more likely to take up insurance, and they are also more attracted to policy that offers generous benefits comparing to those with lower expected medical expenditure. Ellis (Ellis 1989) studies a case where health insurance options changes at a financial firm. Among those who chose the most generous plan, their out-of-pocket costs in the previous year were 8.6 times higher than that of those who choose the least generous plan. In a natural experiment conducted by Culter and Reber (Culter and Reber 1998), employees are found switching to cheaper insurance plan when their current plans increase premium payment. However, older employees as well as those who have high utilization in the past are found to have low switching rate. In short, evidence of adverse selection is found among people who are at high medical risks and people who expect high medical expenditure.

Although adverse selection theory has been long established and developed, few studies are carried out from developing countries. This may due to the limited size of insurance market in these regions, where the industry is not yet mature and consumers' behavior is less predictable than that in developed countries. We expect mixed results of adverse selection studies in developing countries because On one hand, adverse selection theory implies potential customers with higher expected health care expenses will be more likely to buy insurance; one the other hand, cultural barriers, poverty and administrative misconducts can all possibly lead to other type of market failure, which may cause more concerns than adverse selection problem.

Based on SKY household survey data, Polimeni and Levine (Polimani and Levine, 2011) find strong evidence of adverse selection. Households who join SKY had more past health shocks. Moreover, households who paid full price premium have much higher health care utilization than those who paid discounted prices. The study also finds that Insured households with low health care utilization are more likely to drop insurance coverage, and the remaining households have higher hospital utilization over time.

As another part of SKY evaluation, Levine and Zhang (Levine and Zhang, 2011) use high frequent logbook data to study the dynamic of SKY members joining and leaving insurance coverage. In contrast to Polimeni and Levine, Levine and Zhang find no evidence of adverse selection. In particular, new members are not disproportionately those who had severe health shock. The disagreement between two studies attracts our attentions to the behavior of Cambodian insurance subscribers and the effectiveness of SKY program. Our research is a continuation of previous SKY-based studies and the results present in this paper will add to the existing literature.

3. Hypotheses and Methods

Previous studies on SKY insurance program suggest several possibilities explaining the large gap in health care utilization between HEF members and VMs. We are interested in if any covariates (poverty, remoteness, etc.) reduce the usage gap. We add one variable at a time, building to the complete model that has all the control variables.

The based line regression functions defines health care utilization as a function of HEF membership:

$$U_i = \beta_0 + \beta_1 HEF_i + \varepsilon_i \tag{1}$$

Here U_i is the HC/hospital utilization of household i measured in a) percentage of months with any positive usage b) number of visit to health facilities per month c) monthly cost occurred at health facilities. Dummy HEF_i=1 for households enrolled in HEF.

Utilization is measured by number of Hospital Center (HC) visits per household-month, number of hospital visits per household-month, monthly cost at HC per household-month, and monthly cost at hospital per household-month. Both HCs and hospitals have many zero total visits per month. These zeros are mostly because household members are not sick and therefore do not need to use health care service. Zero total monthly visits can be predicted by a dummy that measures probability if any utilization occurred. We use -OLS- model to estimate monthly visits, and use -dprobit- model to estimate the likelihood of any positive usage. We also use zero-inflated negative binomial regression to analyze the number of monthly HC/hospital visit, where the positive usage dummies are used as zero inflators.

Both costs at HCs and at Hospital have large positive outliers. Thus, we focus our main parametric analysis on the natural logs of (monthly costs at HCs + 1000[1]) and of (total health care costs + $1000) and compress them by compressing at 99% percentile. This transformed variable has neither positive nor negative outliers. discourages

3.1 Poverty as a barrier to usage

Because HEF targets households who cannot afford health insurance, this assumption implies our first hypothesis H[1A]: The inferior wealth condition of HEF members discourages them from utilizing health care.

Poor households reduce HC/hospital visits to avoid non-medical expenditure that will incur. We compare the poorest voluntary SKY members to HEF members by restricting the pooled sample to a common support group (wealth score 1-9, see section 5, VMs in common support groups have wealth scores ranging from the 3rd percentile to the 49th percentile). Change of β_1 in (1) and (2) will illustrate how much of the utilization gap is caused by wealthier VMs who can afford non-medical expenditure associated with HC/hospital usage.

$$U_i = \beta_0 + \beta_1 HEF_i + \varepsilon_i \tag{2}$$

[1] In riel, Cambodian currency; 4000 riel= 1 USD

Among those who are equally poor, we are interested if wealth condition can affect households' decision on health care usage. Our second hypothesis is H[1B]: poorer households tend to use more public care than wealthier households. We use regression (3) to test this hypothesis, where W_i is the i^{th} categorical dummy if wealth index (wealth score)=*i*. However, because even poor VMs can afford the other associated costs, effects of wealth conditions for the VMs may not be observable. In addition, as we are unable to separate moral hazard from wealth effect on HEF members, γ_i measures the net effect of moral hazard and wealth effect by wealth categories.

$$U_i = \beta_0 + \beta_1 HEF_i + \gamma_j \sum_{j=1}^{n} W_{i,j} + \varepsilon_i \qquad (3)$$

Notice that regression (3) also cannot capture the effects of opportunity costs. For example, where no one is available to replace HEF members at work or to take care of the children.

3.2 HEF members are more remote than VMs

Distance to local HCs may create barriers to access public care service, which implies our second hypothesis H[2A]: households who live farther from HCs have less utilization than those who live closer. We compare usage of health services of those who live in the same village to condition on remoteness across villages. In regression (4), T_i measures transportation time (by hour) to a local HC by motorcycle. These tests ignore distance within a village, and it cannot capture whether the availability of transportation (potentially a significant issue for poor people who live far from main road) becomes a barrier.

$$U_i = \beta_0 + \beta_1 HEF_i + \gamma_{i,j} \sum_{j=1}^{n} W_{i,j} + \beta_2 T_i + \varepsilon_i \qquad (4)$$

Due to lack of the distance information, regressions on hospital visits do not include the variable of transportation cost. However, exclusion of distance variable should not significantly change the results because people only go to hospitals if they are referred by local HCs. Thus distance is less of an important factor to hospital visits than to HC visits. Distance variable will not be included in hospital usage regressions.

It is plausible that H[2B]: remoteness is more of a barrier to the poor, as they cannot pay for travel as easily. If transport costs are an especially large burden for the poor, then the term $T_i \times HEF_i$ in regression (5) should capture the interaction between distance and wealth conditions.

$$U_i = \beta_0 + \beta_1 HEF_i + \gamma_{i,j} \sum_{j=1}^{n} W_{i,j} + \beta_2 T_i + \beta_3 (T_i \times HEF_i) + \varepsilon_i \tag{5}$$

3.3 Utilization gap is narrower in Kampot

As aforementioned, variation in SKY's performance in different districts can also cause overestimation of the utilization gap. Out third hypothesis is therefore H[3]: the utilization gap is overestimated by higher utilization of VMs who do not live in Kampot overestimate. In regression (6), district dummy K_i=1 if household *i* is in Kampot district.

$$U_i = \beta_0 + \beta_1 HEF_i + \gamma_{i,j} \sum_{j=1}^{n} W_{i,j} + \beta_2 T_i + \beta_3 (T_i \times HEF_i) + \beta_4 K_i + \varepsilon_i \tag{6}$$

3.4 Self-selection among voluntary insurance members

Previous studies on SKY insurance program reflect strong self-selection among VMs who pay higher premium. If VMs in the common support group receive more treatment than HEF members, poor VMs could be even more adversely selected. Thus, it is worth testing H[4] that:

self-selection among those paying full price causes the gap between HEF and voluntary members. In regression (7a), discount dummy D_i =1 if a VM pays premium at steep discounts. For illustration, we replace D_i by N_i in (7b), where dummy N_i =1if a VM pays at normal prices.

$$U_i = \beta_0 + \beta_1 HEF_i + \gamma_{i,j} \sum_{j=1}^{n} W_{i,j} + \beta_2 T_i + \beta_3 (T_i \times HEF_i) + \beta_4 K_i + \beta_5 N_i + \varepsilon_i \tag{7a}$$

$$U_i = \beta_0 + \beta_1 HEF_i + \gamma_{i,j} \sum_{j=1}^{n} W_{i,j} + \beta_2 T_i + \beta_3 (T_i \times HEF_i) + \beta_4 K_i + \beta_5 N_i + \varepsilon_i \tag{7b}$$

4 The Setting

In this section we discuss SKY Macro-Health Insurance program, its linkage model with HEF and randomization procedures.

4.1 SKY Health Insurance

SKY is a non-profit micro health insurance scheme active in Cambodia since 2000. SKY offers broad health coverage at affordable prices and plays a role of a third party purchaser. The target population of SKY Health Insurance scheme has been defined as “the near-poor,” whereby members are eligible for this definition given that they must be in a secure position to afford the regular premiums for the health insurance of the family. Every eligible household can join SKY on a voluntary basis at anytime, but coverage will not begin until the start of the next month. Households sign up for a six-month cycle, paying for the first month’s coverage plus two reserve months up front. Premiums are paid on a monthly basis with prices varying by family size and geographical zone. With their insurance, household members are entitled to free services and prescribed drugs at local HCs and at public hospitals with a referral.

4.2 Health Equity Fund Linkage

In Kampot Health District SKY has developed a linkage model with an HEF since May 2008. Through pooled funding, the Cambodian government and donors purchase premiums with the SKY scheme on behalf of the poor. The premium covers medical expenses only. Once identified as "poor" by the Ministry of Planning (MOP), households received a SKY Health Insurance booklet. Pre-identified households received the same booklet and therefore the same comprehensive health benefits package in the same health facilities as SKY VMs.

4.3 Randomization

Previous SKY evaluation designs utilized the randomization of coupons for premium reductions to isolate the impact of health insurance from all other factors that might affect a household's decision to take up insurance. Originally introduced as a marketing technique, premiums discounts are distributed randomly as SKY visits villages and sell insurance at village meetings. Of households that attend the village meeting, 20% of households receive a coupon for 5-months free insurance in the first-month cycle.

5 Data

We create a long-formatted dataset using Rachel's SKY membership data, second round SKY household survey, and SKY HC/hospital utilization data, Nicolas' HEF membership data, HEF-SKY utilization data and household survey data, and distance information provided by Kempf's survey. We collect the sample by merging SKY membership data with HEF membership data, expanding household level data to household-month, and linking utilization data and wealth condition to each household-month.

Although SKY members in Rachel's data are also listed in Nicolas' membership data, the utilization data from of two sources do not agree well. Moreover, some households have double-entered membership status. The following sections describe in details about each dataset we use and the treatment we applied in constructing the pooled data.

5.1 Data Sources Description

The two data sources we use are Rachel's datasets by SKY and Nicolas-ADF' datasets by Domrei Research and Consulting[2]. We draw information from SKY membership data, SKY members' HC/hospital utilization, SKY household survey from Rachel, and HEF membership data, HEF utilization, and HEF wealth information is from Nicolas.

Nicolas' data has monthly utilization records of 18945 households (685710 household-month), dating from November 2007 to December 2010. Among the 18945 household data, 984 households are neither HEF members nor have SKY membership status. Among the rest (17061 households), 5300 are HEF members, 11761 are VMs.

Rachel's data is from SKY household survey round 2. It consists monthly utilization records of 2143/6993 households[3], dating from November 2007 to December 2010. Among these 2143 households, 42 are HEF members and 2098 are VMs.

5.2 SKY Membership

SKY membership data is obtained from SKY administration data. However, Nicolas' HEF membership data also provide households' SKY membership status. We find that all households with SKY membership status in Rachel's data are covered in Nicolas's data. However, 1645

[2] Domrei Research and Consulting is an economic consulting group based in Cambodia. It is in charge of distributing premium coupons, conducting household surveys, and monitoring logbooks data for SKY health insurance program.

[3] Rachel's SKY administration data and utilization data both show that 4850 households are assigned SKY IDs but lack monthly membership status and utilization data. Those households are excludes from our study.

households have different monthly SKY membership[4]. After exanimating those error data, we convert SKY membership status to that in Rachel's data.

Due to data size difference and recording errors, SKY membership status is found missing in many cases. In Nicolas' data, 108 household-month entries have no SKY-membership status, and we recoded them as non-SKY members according to Rachel's SKY membership data. As for missing SKY membership status in Rachel's SKY membership data, 184,262 household-months (4,849 households) overlap in Nicolas' data, where their SKY membership status is known.[5]

After adjusting SKY membership status, the merged data consists of household-month data for 18,048 households (685,818 household-month). Among these households, 17,073 households are enrolled in SKY either as HEF members or as VMs (350,200 enrolled household-month). 975 households are never SKY members (37,004 household-month), and will be excluded from our sample regardless of their utilization.

5.2.1 Membership Type: HEF Members and Voluntary Members (VMs)

We divide households that have SKY/HEF membership status into two categories—voluntary members (VM) and HEF members. We index SKY membership status and HEF membership and the pooled membership data are structured as Table. 1[6].

[4] Among these unmatched data, 23536 (1472 households) are SKY members in Rachel's data, but not in Nicolas' data. Because the raw data of HEF membership in Nicolas' data is recorded by month, we suspect that the original form HEF membership recording is possibly the cause. However, none of the 1472 households are ever HEF members, those mismatches are simply errors and we convert membership status in Nicolas' data to that in Rachel's data. The other 832 household-month SKY membership data (173 households) are cases such that household-months are SKY members in Nicolas' data, but not members in Rachel's data. We divide these 832 entries by HEF membership status, and find 335 are associated with HEF membership, while 497 are not. Because our study treat HEF membership by household, and HEF member are automatically enrolled in SKY program, we adapt SKY membership status in Rachel's data to that in Nicolas' data.

[5] There are 499 household data with unknown SKY membership status in two membership data. Each of these households has only one household-month entry in the combined data. We drop those data from our pooled dataset.

[6] In indexing SKY membership and HEF membership, we adopt the following rules: 1) HEF membership status is by household. Once been a HEF member, households cannot withdraw HEF membership. A HEF household is indexed as a HEF member in every month from November 2007 to December 2010. 2) SKY membership status is

Table. 1 Membership composition

	HEF member	Ever a VM[b]	Out of SKY	Total
Nicolas' Data	5,300 (201,400)	11,773 (151,227)	972 (36,936)	18,045 (685,710)
Rachel's Data	42[c] (1,512)	2,098[d] (53,768)	3 (108)	2143 (77,148)
Overlap	42 (1,512)	2098 (53,768)	0 (0)	2140 (77,040)
Total	5,300 (201,400)	11,773 (151,227)	975 (37,004)	18,048 (685,818)

Note:

a) Unit = household; household-month in parenthesis.

b) Numbers in parenthesis of column "Ever a VM" count the actual household-month enrolled in SKY. Household-months before joining SKY and that after withdrawing from SKY are excluded.

c) 42 overlapped HEF members in Rachel's data are restricted to households with filled information including monthly SKY membership status, HC/hospital utilization, etc. Besides these 42 households, there are 1,627 HEF members whose SKY IDs are also listed in Rachel's data. However, Rachel's SKY administration data and utilization data have no information about these households other than their SKY IDs.

d) 2098 overlapped HEF members in Rachel's data are restricted to households with filled information including monthly SKY membership status, HC/hospital utilization, etc. Besides these 2098 households, there are 3,221 households whose SKY IDs are also listed in Rachel's data and qualify for VM based on Nicolas' data of SKY membership. Nevertheless, Rachel's SKY administration data and utilization data have no information about these households other than their SKY IDs.

5.3 Utilization Data

We use SKY HC/hospital utilization data from Rachel and health utilization data from Nicolas for our study. The utilization data according to Nicolas are originally long-shaped and utilization is recorded by each visit. SKY membership data, HEF membership data and utilization data from Rachel are wide-shaped. Each household has membership dummies associated with each month since November 2007 to December 2010. Rachel's utilization data is also recorded by

---

by household-month. Data starts in November 2007 and ends in December 2010. We refer to Rachel's data if any conflict arises. 3) Membership type of a VM is indicated by household-month. A household is a VM if household is never a HEF member and it is a SKY member in the corresponding month. 4) Household-month that after household temperately quits SKY but rejoins later is still considered as a VM. There are 2427 household-months qualify for this criterion.

household-month. Month with zero utilization is coded as missing if a household has no HC/hospital utilization occurring since November 2007, and is coded as zero otherwise. We collapse Nicolas' data by month and merge with two membership datasets and Rachel's utilization data. Merging automatically expands household in Nicolas' utilization data into 38 household-months. We code household-month with no positive utilization as zeros.

Households with utilization records in Rachel's data are covered by Nicolas' data. However, many of the data do not agree in the two data sets. For data with positive HC utilization, 3081 household-month entries (802 households) have different HC visits and 2505 household-month entries (633 households) have different HC costs in two data sets. For data with positive hospital utilization, 982 household-month entries (527 households) have different records of hospital visits and 452 household-month entries (247 households) have different records of hospital costs. Correlations of conflict utilization data under each category are positive but all below 0.9[7].

In addition to errors of positive entries, we observe many utilization data that are clearly recorded in one data set, but are missing values in another. HC utilization data of 5329 households (191105 household-month) is missing in one of two datasets, among which 95% are VMs[8]. For hospital utilization, 4905 households (174896 household-months) have missing data in one of two datasets, and 65% of these households are VMs[9].

---

[7] Correlation of conflict utilization data of HC visit, HC expenditure, hospital visits and hospital expenditure are 0.75, 0.57, 0.75 and 0.87, respectively.

[8] For HC utilization, 249 households (8283 household-months) have full records in Rachel's utilization data, but have no information in Nicolas' utilization data. 246/249 say it in words households according to Nicolas' data are households that once are VMs (the other three are out of SKY program). In the other direction, 5080 households (182822 household-months) have utilization data in Nicolas' data, but no information in Rachel's utilization data. 3414 of these 5080 households are once VMs. In addition, 4810/5080 households are households without SKY membership information; the remaining 270 are those with membership status but no utilization record.

[9] For hospital utilization, 1038 households (26096 household-months) have full records in Rachel's utilization data, but have no information in Nicolas' utilization data. 1038/1041 households according to Nicolas' data are households that once are VMs (the other 3 are not in SKY program). In the other direction, 3864 households (139028 household-months) have utilization data in Nicolas' data, but no information in Rachel's utilization data.

We are very concerned about the data inconsistency and errors. If we focus on household with both SKY membership status and HC/hospital utilization listed in Rachel's data, we observe that 1956 nonempty household-month entries (487 households) are different in Nicolas' data to Rachel's data in term of percentage-of-any-visit-to-HC. We also worry about the 27607 household-months (519 households) that are only recorded in one of our two dataset. These worrisome entries count for about 23% of Rachel's household-month non-missing data (Table. 2).

Table. 2 Unmatched HC visit data

	Rachel's Data			
Nicolas' Data	No HC visit	Any HC visit	Info. missing	Total
No HC visit	45,197	2	7,388	52,587
Any HC visit	1,954	10,109	2,274	14,337
Info. Missing	8,283	0	1,941	10,224
Total HH-months	55,434	10,111	11,603	77,148

Errors in hospital utilization are similar to those in HC utilization data. Table 3 shows that 120 household-month data with positive entries are recorded differently in our two utilization data sources. Although the proportion of conflict entries is small, 54% households-months with SKY membership have missing hospital utilization data (42184 household-months, 1219 households), and 70% of those missing data are from VMs.

Table. 3 Unmatched hospital visit data

	Rachel's Data			
Nicolas' Data	No hospital visit	Any hospital visit	Info. Unknown	Total
No hospital visit	26,451	5	5,925	32,381
Any hospital visit	115	1853	407	2,395
Info. Unknown	35,852	0	6,504	42,372
Total	62,418	1,878	12,836	77,148

2195 of these 3864 households are once VMs. In addition, 3686/3864 households have no SKY membership information and the remaining 270 are those with membership status but no utilization record.

As shown in Table. 3, missing values are misleading and can underestimate the utilization of VMs in Rachel's data and overestimate that from Nicolas' data. We thus adjust conflict utilization data and some descriptive statistics are presented in the table below[10] (Table. 4). Row 2 shows that the utilization of VMs in Rachel's data is substantially higher than other groups. This is because besides the 2143 households with valid utilization data, there are 3221 VMs also listed in Rachel's data[11], but with missing utilization information. Nevertheless, Nicolas' data tells us that 44.35% of these VMs have zero visits to HCs and 87.64% have zero visits to hospitals. That is, by not including low usage households, Rachel's data tends to overestimate the utilization of VMs and thus is not a desirable sample if it stands alone. We thus only use Rachel's data for SKY membership status, discount distribution, and supplementing the pooled data if utilization data are available.

Table. 4 Descriptive statistics after utilization data adjustment

	Membership Type	Data Source	[a,b] # of households-month	% with any HC visit	Health center visits / month	Hospital visits / month
1	HEF members	Nicolas'	201400 (5300)	7.61%	0.17	0.03
2	Voluntary members	Nicolas'	151008 (11761)	36.11%	1.18	0.17
3	Voluntary members	Rachel's data	53768 (2098)	22.39%	0.72	0.07
4	Voluntary members paying at normal price	Rachel's Data	16496 (627)	26.03%	0.91	0.09

[10] In adjusting conflict/missing utilization data, we adopt the following rules: a) If both Rachel's data and Nicolas's data have non-missing household month utilization (HC/hospital), refer to Rachel's utilization data. b) If household-month utilization is missing in one dataset but not in the other, refer to the non-missing value and use that to replace corresponding entry in the other data source. c) If household-month utilization is missing in both datasets, but the household is a HEF member, code utilization as zero in both data sets. d) If household-month utilization is missing in both datasets, but the household has ever been a VM, in both data sets code utilization as zero of the month enrolled in SKY, and leave it as missing value of the month out of SKY.

[11] 3220 are VMs paying at normal prices and one is VM paying at discounts, 49409 household-months in total.

5	Voluntary members paying a deep discount	Rachel's Data	37272 (1471)	20.77%	0.64	0.06
6	Voluntary members in Kampot paying a deep discount	Rachel's Survey + Kempf	2137 (77)	19.65%	0.53	0.09

Note:

a) Unit = household-month; N. households in parenthesis

b) Although we adjust HC and hospital utilization data by referring across two data sources, we restrict out sample size to household-month that i) households are HEF members or they are enrolled in SKY member, ii) household that are listed with HEF or SKY membership status in the original data sets.

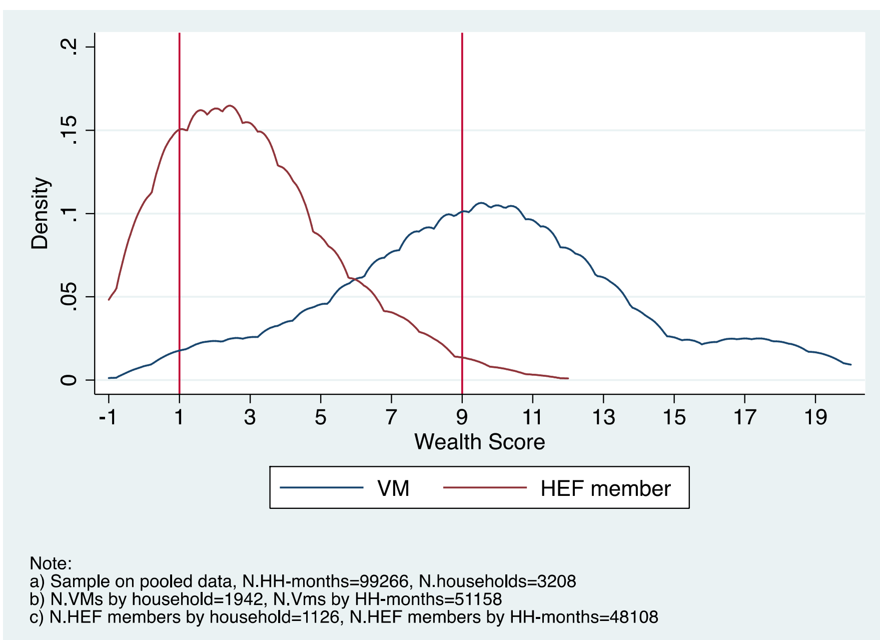
5.3 Wealth Index

Wealth conditions are available from SKY household survey from Rachel, and HEF survey from Nicolas. Although questionnaires are drafted in different forms, we find common measures such as raising stocks and owning debts that can help us build near-identical sets of wealth measures in Rachel's and Nicolas' datasets. Referring to the wealth score scheme used in *Assessment of HEF beneficiary identification in three operational health districts in Kampong Cham*, we create wealth index by summing scores what are assigned to each of households' property. Rachel's data provides wealth information of 1837/1874 households who have associated SKY IDs (the total number of household in Rachel's health data is 5394). Nicolas' data provide wealth information of 1913 households.

There are 26 households overlapped in both datasets. For these households, we calculate the correlation between wealth scores obtained from Rachel's data and those obtained from Nicola's data. Unfortunately, we see a negative correlation (corr. = -0.1439). Ideally our wealth score are weighted scores of property dummies. However, by comparing wealth measures from both datasets, we find many property dummies of the same household do not agree in two data sets. Moreover, errors reflected by wealth scores can still be underestimated because the actually differences in property are fairly large. We suspect that some of these errors are caused by change of households' property over time, and records in two surveys simply reflect such changes. In our study of wealth effect on HC/hospital utilization, we take the mean of wealth scores for each household.

Figures 7A-B show the distribution of wealth scores among VMs and HEF members, respectively. Most HEF members are associated with wealth scores between 0 and 10, while those of VMs are more widely distributed (from 1 to 20). We define *common support group* as households whose wealth scores range between 1 and 9, inclusively. The common support group cuts HEF member wealth score from the 31st percentile to the 99th percentile, and it cuts VMs wealth score from the 3rd percentile to the 48th percentile. Among the common support groups, we round down fraction scores by taking the largest integer values that are less than or equal to them[12].

Figure. 7A Distribution of wealth scores among HEF members and among VMs



[12] Fraction wealth scores are generated by taking average of unmatched wealth scores of the 26 overlapped households

Figure. 7B Distribution of wealth score among HEF members and among VMs

Wealth score	HEF		Voluntary SKY members	
	# of households	% of total HEF members that have wealth scores	# of households	% of total VMs that have wealth scores
<0	30	2%	4	0%
0-0.5	168	13%	8	0%
1-1.5	195	15%	56	2%
2-2.5	209	17%	61	2%
3-3.5	217	17%	55	2%
4-4.5	153	12%	69	2%
5-5.5	115	9%	102	4%
6-6.5	58	5%	107	4%
7-7.5	65	5%	160	6%
8-8.5	26	2%	171	6%
9-9.5	17	1%	184	7%
10-14.5	13	1%	739	27%
>=15	0	0%	1057	38%

5.4 Remoteness measures

Information regarding distance to HC is obtained from SKY village leader survey, village data, SKY household survey both from Rachel, and survey data from Nicolas. We also have supplementary distance data from Kemft data.

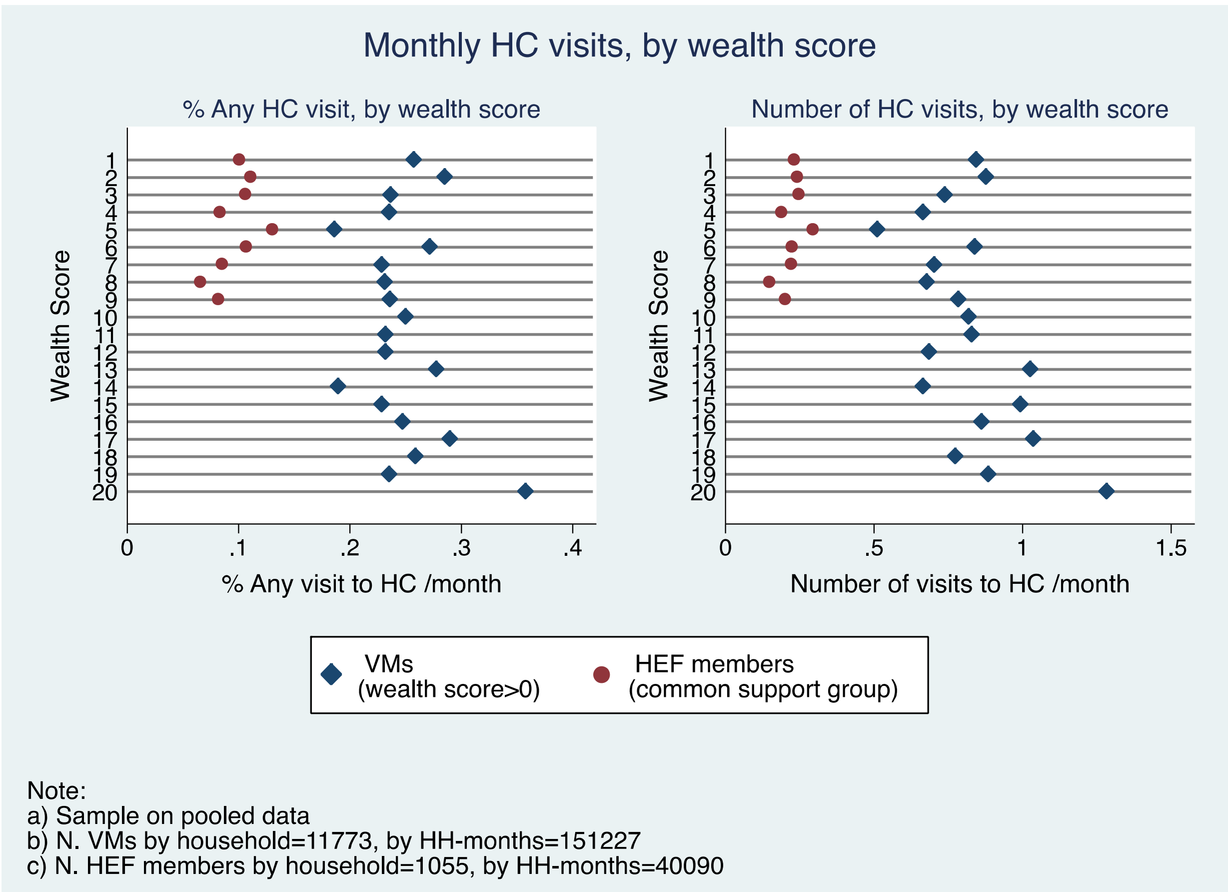
SKY village data and Nicolas's survey data are identical and they provide distance to HC and referred hospital in kilometer and at household level. However, these data only cover 4573 households in 76 villages (542 village in total). SKY village leader survey and household survey are at village level and give approximate time needed to walk to HC or transport to HC by a motorcycle. Using this distance measure about 50% of total villages are covered (transportation-time-by-walking is available for 248 villages and transportation-time-by-motorcycle is available for 275 village). Another advantage of using transportation time as distance measure is that it also contains transportation cost information. Although the correlation between distance-in-kilometer and distance-by-transportation-time is small (correlation= 0.3), we suspect that the distance-in-kilometer only measures that straight-line distance from village to HC, which can be very different from the distance people have to travel by walking or by motorcycle. Our study will focus on the effect of transportation-time-by-motorcycle on health facility usage.

6 Results

6.1 Non-parametric analysis

This section presents summary statistical results regarding health care utilization on condition of common support group, membership types and wealth scores. Statistics about HC visits, hospital visits, HC costs and hospital costs are compared within and across conditions.

Figure 9. Who visit HC more often, by wealth scores?

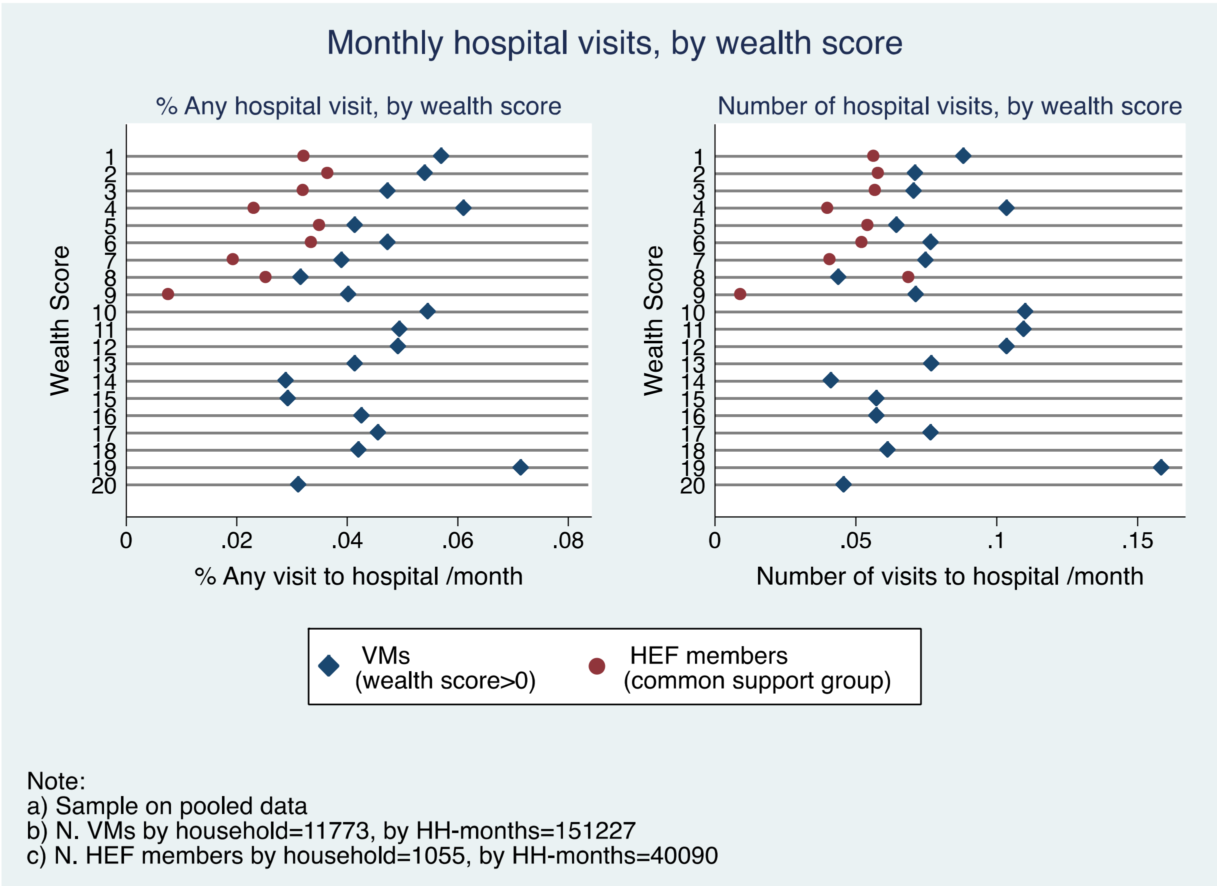


6.1.1 Who pay more visits to HC/hospital?

Figure 9 summaries the HC visits by wealth score category in the common support group. Restricted to common support group (wealth score 1-9), VMs on average are more likely to visit HCs than HEF members in each wealth score group. Beyond the common support group, VMs remains high likeliness of visiting HCs, and the number of monthly visits has an increasing trend from low score groups to high score groups. In the common support group, 10% of HEF members have a positive usage by month, a percentage less than half of VMs' (24%). On average, HEF members are 14% less as likely to visit HCs as VMs. The gap is smaller among low scored households and is wider among high scored households. Nevertheless, no clear pattern is found between wealth scores and household HC visits both among HEF members and VMs.

HEF members in common support group have an average of 0.23 HC visits per household-month, comparing to VMs' average of 0.73. The low HC-visits-per-month of HEF members pervade in all wealth subcategories, whereas the gap between HEF members and VMs are the largest among households who score 1, 2 and 9. The median difference by wealth score is as large as 0.60, in which case VMs have three times as many monthly HC visits as HEF members. We also observe that up to wealth score group 5, poorer VMs tend to visit HC more than wealthier members, while the trend reverts among households who score 5 or higher.

Figure 10. Who visit hospital more often, by wealth scores?



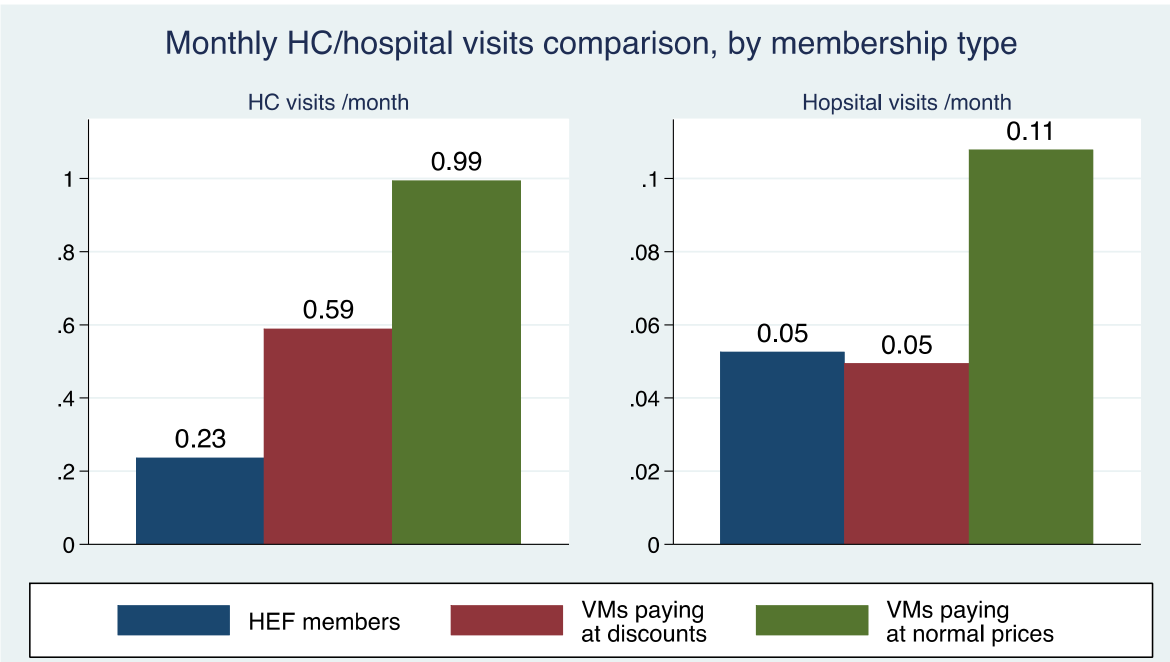
By wealth score, HEF members remain to have lower rate of hospital visit and fewer hospital visits than VMs, but gaps are narrow. 3% of HEF household-months have a hospital visit, whereas the percentage for VMs is 4%. The utilization gaps among wealth groups average at 2%, and the difference reaches the largest in wealth groups 1 and 9. Group average declines as wealth score improves.

The counts of monthly hospital visits are consistent with the rate of any visiting by household-month. On average HEF members pay 0.05 visits to hospitals per month, while the number for VMs is 0.07. By wealth score, the averages of HEF members remain low (except wealth score group 7). Both HEF members and VMs experience usage declines as households' wealth scores increase. It is noticeable that the two widest utilization gaps are in wealth groups 1 and 9, the same groups where the largest HC utilization gaps persist.

6.1.2 Do HEF members have higher utilization than VMs who paid at discounts?

Figure 11 summarizes the average HC and hospital utilization by dividing membership into three categories: HEF members, VMs paying at discounts and VMs paying at normal price. We find that HEF members have the lowest average HC visits among all three groups. The median of HEF members (avg.= 0.23) is less than half of VMs paying at discounts (avg.= 0.59), and is only 1/5 of VMs paying at normal prices (avg.= 0.99).

Figure. 11 Do HEF members have more HC/hospital visits than VMs?



Note:
a) Sample on pooled data and restried to common support group
b) N. HEF members by household=1055, by HH-months=40090
c) N. VMs paying at discounts: by household=578, by HH-months=16199
d) N. VMs paying at normal prices: by household=384, by HH-months=8336

Findings about hospital visits are slightly different. Consistent with H[4], VMs paying at normal prices have on average twice as many as the other two types of members. However, HEF members have a slightly higher average than VMs who paid at discounts (0.052 vs. 0.049). We found that almost 90% of HEF members never visited HCs, and that number is 80% among VMs with discounts, and is only 68% among VMs without a coupon. The proportions of zero hospital visits are 97%, 97% and 93.5% for HEF members, VMs with discounts and VMs with no discounts, respectively.

Figure 12. Do HEF members spend more at HC/hospital than VMs?

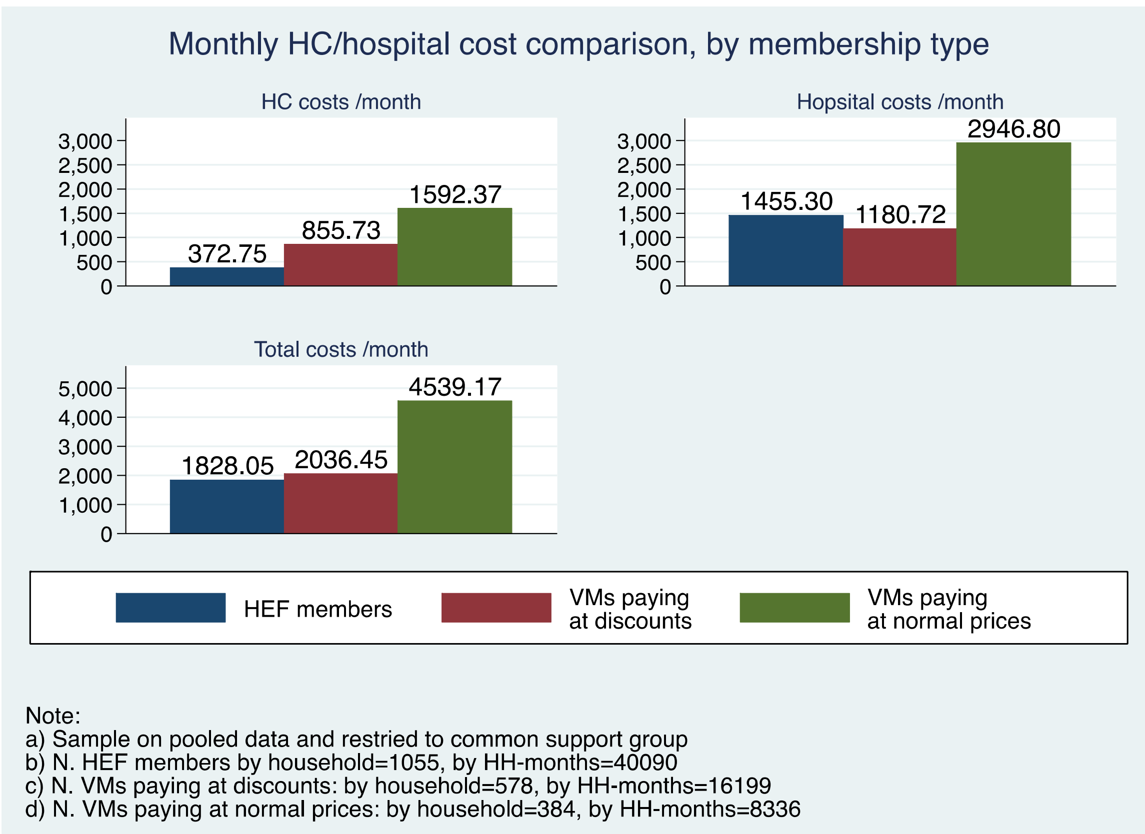


Figure 12 shows the averages of HC cost, hospital cost and total cost by membership type. For HC costs, HEF members have the lowest average among all, and VMs paying at normal prices remain to have the highest averages (more than three times more than HEF members, and twice as much as VMs with discounts). High utilization of VMs also persists in hospital cost (2946.80). Although HEF members have a higher average in hospital spending than VMs paying without discounts (1456.30 vs. 1180.72), the gap is not significant. Graphs of costs are consistent with H[4], which claimed that VMs paying at normal price are adversely selected and thus cause a fairly large part of utilization gap between HEF members and VMs.

Figure. 13 Average HC visits by membership types and wealth score

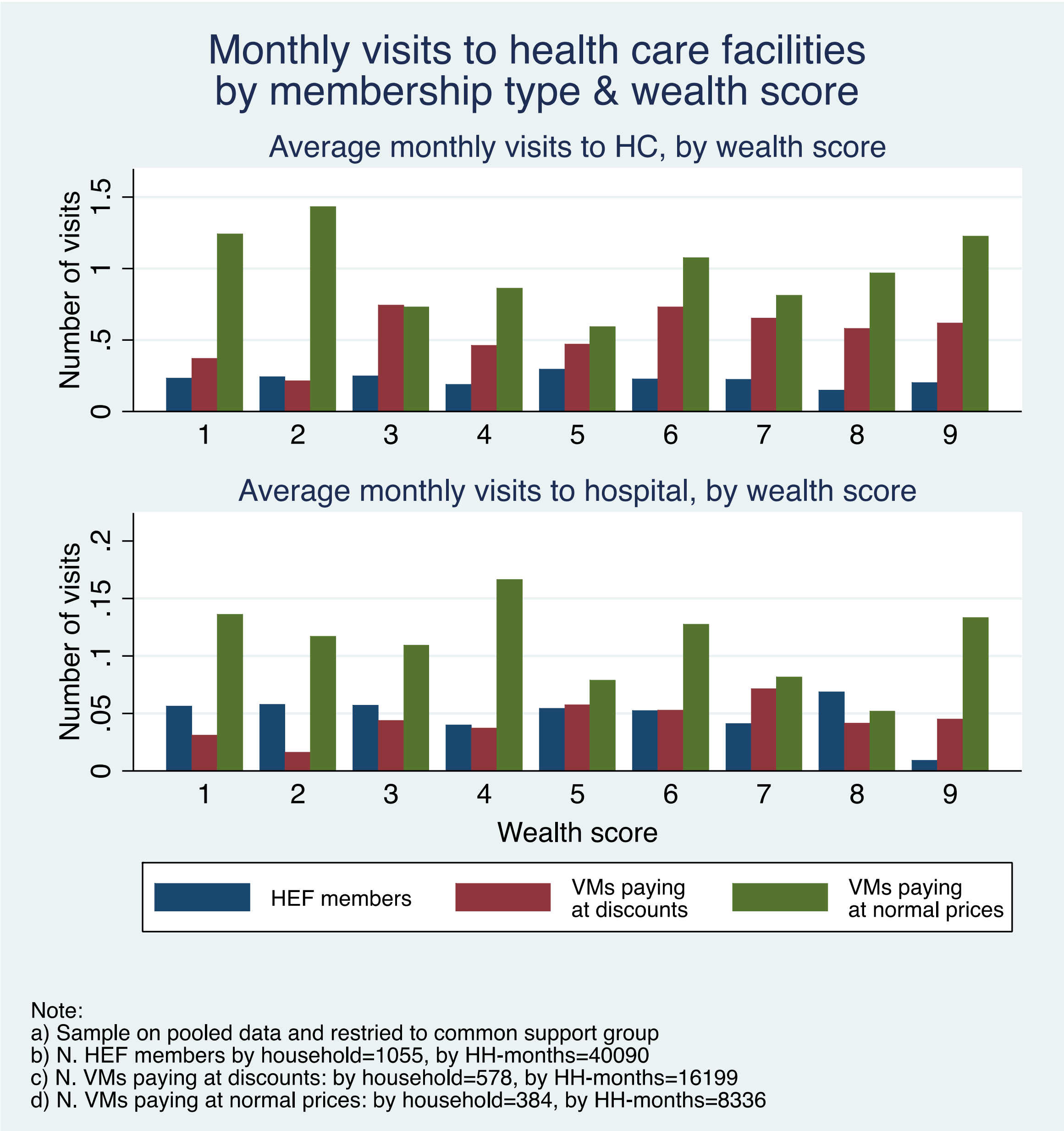


Figure 13 compares the average HC and hospital visits by membership type against wealth score. HEF members have the lowest average HC visit across wealth subcategories. VMs paying at normal prices have the overall highest averages. Especially for those with scores 1 and 2, their averages HC visits are substantively high, both in absolute terms and relative terms comparing to HEF members and VMs with discounts.

We have similar discoveries about the average hospital visits per month. Consistent with H[4], VMs paying at normal prices spike hospital cost in almost all wealth subcategories. Of the rest two groups, HEF members have higher averages in wealth scores 1 and 2, but the difference diminishes in the remaining wealth subcategories. Table 7B tell us that about 38% of HEF members concentrate in the bottom two wealth subcategories, while the poorest only consist less than 7% of VMs who paid at discounts. This composition signifies moral hazard among the poorest HEF members, which results a higher average hospital costs of HEF members than VMs without discounts.

6.2 Regression analysis

We apply regression function (1)-(7) to test our three HC utilization measures: number of HC visits, percentage of any-HC-visit per month, natural logarithm of (monthly HC costs+1000[13]). Due to lack of distance information of hospitals, we are unable to test the impact of remoteness on hospital usage. That is, we will only test hypothesis (1), (3) and (4), using regression function (1)—(3), (6), (7) to test on number of hospital visits, percentage of any-hospital-visit per month, and natural logarithm of (hospital cost+1000). For additional information, we run regression on the natural log (monthly HC cost + hospital costs).

6.2.1 Poverty as a barrier to usage

In this section we compare the utilization gap between HEF members and all VMs, and the gap between HEF members and poor VMs.

---

[13] 1000 riel (Cambodian currency) = 0.25 USD

Under –dprobit– model, the coefficient of being a HEF member drops from -1.07 to -0.54 (Table 14 col. 1, 2). It is impossible for HEF members to use public facilities as voluntary members in baseline. But if we only compare with poor VMs (wealth score< 40^{th} percentile), HEF members are only 54% as likely to visit HC as VMs. Restricting VMs sample also changes the probability an HEF visits a hospital. HEF members are 62% as likely to visit hospital, and the gap drops to 14% after controlling on poor VMs (Table 15, col. 1, 2).

HEF members on average have fewer visits to HCs and to hospitals. The differences however decrease substantively after restricting sample to common support group (Appendix. Table. 18, 19, col. 1, 2) These results are consistent with hypothesis [1A] that utilization gap between HEF members and VMs are largely induced by the inferior wealth condition of HEF members.

We perform an OLS estimation of monthly HC visits as a function of HEF membership. Being a HEF member has a large effect on HC usage. In base line regression an HEF member have one fewer HC visit than VMs, however the gap drops to 0.5 fewer visit per month (Appendix. Table. 18, col. 1, 2) in column (2). Similarly HEF members have 0.13 fewer hospital visit per month, and the difference changes to 0.02 fewer visits per month (Appendix. Table. 19, col, 1, 2). These results show that although HEF members have much lower health care utilization, the gap is driven by wealthier VMs who are more capable of affording non-medical cost associate with health care usage. The evidence supports hypothesis [1A].

Because many observations of the common support group have zero visits, we repeat the analysis of HC visits using negative binomial model. Dummies of any-HC-visit-per-month are used as zero inflation predictors. The results are consistent with OLS model (Appendix, Table. 20), although the relative change of HEF coefficient is smaller than that in leaner regression (relative change in -zinb-:14.3% vs. relative change in -OLS-:50%).

Evidence of hypothesis [1A] is also found in HC costs and hospital costs. Using –tobit– model and top-coded at 99% of natural log (monthly HC costs+1000), the difference between HEF members and VMs drops from -0.46 percentage point to -0.21 percentage point (Appendix. Table 21, col 1, 2).

Hypothesis [1B] suggests that HC/hospital usage varies contingent on wealth condition. We test this hypothesis by adding wealth scores as categorical variables to baseline regression (column (3)). Evidence of wealth effect on a household's usage decision are mixed. As for HC visits and HC costs, all wealth indices have negative coefficients but they are nevertheless insignificant (Table. 18). In analyzing hospital visits and hospital costs, we find most coefficients of wealth indices positive and the effect of wealth group (1)-(3) on hospital cost are significant at 5% level. However, F-test shows that joint results are not significant at traditional levels. We also notice that the coefficient of HEF also decreases significantly in absolute term. This suggests that for hospital costs, wealth conditions such as income and assets have a large effect on a households' decision of hospital expenditures. However, wealth indices of the poorest lose their significance after adding district variable Kampot, and thus the correlation is not strong enough to support hypothesis H[1B].

6.2.2 HEF members are more remote than VMs

To test hypothesis H[2A], we employ distance measure. In Column (4) of Table 14, coefficient of HEF separates the effect of living in a remote village from other effects associated to HEF membership. Utilization gap is still significant. HEF members are now 36% (t=-6.53) less likely to visit HC as VMs (Table. 14, col. 4), comparing to 57% less without distance control. On average, an additional hour of transportation time to HC will reduce the likelihood of visiting HC by 97% (t=-4.78), a result consistent with OLS estimation of monthly HC visits (coefficient of transportation cost=-0.8, t=-5.14, Appendix. Table 18). Therefore H[2A] is accepted.

Hypothesis [2B] posits that transportation costs could create extra burden for HEF members. We test this hypothesis by introducing dummy $T_i \times HEF_i$ (Table. 14, col.5), an interaction between transport costs and HEF membership. Controlling on distance, the -dprobit- model (Table. 14) predicts that regardless of the inferior wealth conditions, HEF members are in fact are 9.82% more likely to visit HCs, however, the coefficient of interaction is not significant (t=1.25). The -OLS- model also shows an HEF members pay 0.36 more visits to HCs if we remove casual effect of distance and wealth condition (coefficients of distance and interaction are both significant, SE=-4.98, and SE=2.94). Suspicious about the OLS results, we use negative binomial model again to test the interaction between HEF membership and distance and its effect on HC visits (any-HC-visit -per -month as zero predictor, Appendix. Table. 20). The coefficient of interaction term remains positive but not significant (t=0.39). Therefore, we cannot conclude whether transportation cost worsens the burden of HEF members. Hypothesis H[2B] is not conclusive.

6.2.3 Utilization gap is narrower in Kampot

Hypothesis H[3] proposes that the utilization gap can be overestimated if the utilization of VMs from other districts are substantially higher than VMs from Kampot, where HEF is operated exclusively. We test this hypothesis by adding Kampot dummy K_i. -dprobit- model (Table. 14) indicates that removing the district effect of Kampot, the HEF members are only 46.4% less likely to use HCs than VMs, a gap that is slightly larger than that without district control at 44.7%. Regardless of HEF membership, households in Kampot are 16.67% more likely to use HC service than households from other districts. OLS model also estimates 0.035 more HC visits per month for households in Kampot than the rest; nevertheless the magnitude of change is very small and coefficient is not statistically significant (Appendix. Table 18). In contrast to the large effect on HC usage decision, district effect contribute little to the hospital utilization gap. The coefficient of HEF is almost identical before and after adding district control (Table. 15, col.5). Regardless of HEF membership, households in Kampot are 33% (t=7.04) more likely to visit hospitals, and they have an average of 0.044 (t=4.96) more visits per month than households elsewhere. However we should notice that although the coefficient of Kampot dummy is alone significant, it also capture the distance effect among villages and the fixed effect of each village in Kampot (recall that we lack access of distance information from hospital to villages).

6.2.4 Self-selection among voluntary insurance members

Previous studies reflect that among VMs, households who paid at normal prices are more adversely selected than those who paid at normal prices. In our study, VMs with steep discounts have higher utilization than HEF members. Regression Table. 14 predicts that on average HEF members are 61.6% less as likely to use HC service as VMs who paid at normal prices, a much larger gap comparing to 46.3% less in column (6). In addition, HEF members are 38% less as likely to visit hospital as VMs who paid at normal prices, whereas the gap is 23.8% if not exclude VMs who paid discounted rate. Moreover, F-test shows that holding other variable constant, hospital utilization of HEF members and that of VMs with steep discounts are essentially the same (F-test: p=0.0574, Table. 15). A closer look at OLS regression on hospital visits tells us that monthly hospital visits of HEF members and that of VMs who paid discounted premium are not significant different (Table 16, col.5, F-test: p=0.548). This observation implies that gap in hospital utilization is largely induced by self-selection of VMs paying full prices. Nevertheless, the HC utilization gap persists despite the self-selection effect from VMs paying full prices. On average HEF members are 15% less as likely to visit any HC as VMs who paid at steep discount (F-test p. = 0.00000242), and 61% as less likely to visit any HC as VMs who paid at full prices (Table. 16, col. 7). Speaking of monthly HC visits, HEF members on average have 0.36 less hospital visit than VMs who paid discounted premium, and have 0.61 less hospital visit than VMs who paid at full prices. Monthly hospital visits are significantly different between HEF members and VMs with discounts. This result suggests that some other factors also contribute to the HC utilization gap, especially the gap between HEF members and VMs who paid at discounts. Unfortunately we are unable to include these factors in our model.

Because hospital cost data have many zero and large outliers, we focus on the nature log of total expenditure, that is ln(HC costs + hospital costs+1000). Table. 17 suggests that HEF members have about 0.8 percentage points fewer costs than VMs who paid discounted prices, and 1.8 percentage points fewer costs than VMs who paid full price. Both differences are significant. Again this result is consistent with Polimeni and Levine (Polimeni and Levine, 2011).

7 Conclusion and Implication

To study the cause of utilization gap between HEF members and VMs, we define our baseline regression on a household's health care utilization as a function of HEF membership. By adding one variable at a time, we are able to approximate the contribution of the new added variable to the total utilization gap by comparing the change of HEF coefficient. Our main hypotheses posit that utilization gap can be generated by poverty, remoteness, district difference, and self-selection among VMs.

We find strong evidence supporting Hypothesis [1A] that a large proportion of utilization gap is due to HEF members' overall inferior financial condition. Although HEF members are entitled to health care service at zero cost, it is possible that non-medical expenditure associated with HC/hospital usage discourages them from receiving treatments. At the same time, there is little evidence that within HEF members and poor VMs, worse wealth condition predicts higher HC/hospital utilization (Hypothesis [1B]).

There is strong evidence that living in a remote village limits HEF member's access to health care utilization (Hypothesis [2A]). Nevertheless, we find no support that transportation cost worsens HEF members' financial burdens and thus restricts their health care utilization (Hypothesis [2B])

Controlling on district, we find that households in Kampot on average have higher health care utilization regardless of membership types. Consequently, the utilization gap is smaller if we only compare HEF members with VMs who live in Kampot. However, we do not observe significant change of the utilization gap after adding district variable. Thus Hypothesis [3] is not supported.

We find mixed evidence of self-selection for Hypothesis H[4]. VMs paying at normal prices have higher HC/hospital visits and expenditure. We find that being an HEF member predicts essentially the same hospital expenditure as being a VM with a coupon. This observation implies that the hospital utilization gap is mainly driven by the self-selection of VMs who paid full price. Results regarding HC utilization is mixed. Although evidence does suggest adverse selection among VMs and support the Hypothesis H[4], the difference between HEF members and VMs paying at discounts remains significant. We look forward to future studies that may shed light on other factors.

In short, our results imply that poverty and remoteness are barriers for HEF members to access public health care. While completely removing these barriers is infeasible, appropriate policies such as building infrastructure can help limit these obstacles. Self-selection among VMs also is the major part of the utilization gap. Although identifying adverse selected new subscribers is less practical, insurance providers should design insurance schemes in order to avoid adverse selection and excess utilization. Doing so also improves the financial sustainability of insurance institutes and will help build a system that efficiently distributes benefits and improves social equality.

Reference

Annear, P.L, D. Wilkinson, et al. 2006. "Study of Financial Access to Health Services for the Poor in Cambodia, Phnom Penh, Cambodia Ministry of Health, WHO, AusAID, RMIT University.

Currie, J. and J. Gruber. 1996. "Health Insurance Eligibility, Utilization of Medical Care, and Child Health." The Quarterly Journal of Economics 111(2): 431-66.

Cardon, J. H., and I. Hendel, 2011. "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. Philopson. 1990. "An Emperical Examination of Information Barriers to Trade in Insurance," *The American Economics Review,* 89)4), pp. 827-846.

Cutler, D. M., and S. J. Reber. 1998 "Adverse Selection in Health Insurance, " *The Quarterly Journal of Economics,* 113(2), pp. 433-466.

Gottret, P. and G. Schieber. 2006. Health Financing Revisited: A Practitioner's Guide. International Bank for Reconstruction and Development/The World Bank.

Karlan, D., and J. Ziman. 2009. "Observing Unobservable: Identifying Asymmetric With a Consumer Credit Field Experiment," *Econometrica,* 77(6), 1993-2008

Kenjiro, Y. 2005. "Why Illness Causes More Serious Economic Damage than Crop Failure in Rural Cambodia." Development and Change 36(4): 759-83.

Levine, David I., Rachel Polimeni and Ian Ramage. 2011. "Insuring Health or Insuring Wealth? An experimental Evaluation of Health Insurance in Cambodia. 2012. Working paper, University of California, Berkeley.

Levine, David I., Rachel Polimeni and Ian Ramage. 2011a. "Adverse Selection based on Observable and Unobservable Factors in Health Insurance." Working paper, University of California, Berkeley.

Levine, David I., Nhong Hema, and Ian Ramage. 2007. "*Insuring Health: Testing the Effectiveness of Micro-health Insurance to Promote Economic Wellbeing for the Poor.*" Working paper, University of California, Berkeley.

Levine, David I., Karen Zhang. 2011. "Adverse Selection in Purchasing Health Insurance: Evidence from High-Frequency Logbook Data." Working paper, University of California, Berkeley.

Manning, W.Get al. 1987."Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." The American Economic Review 77(3): 251-77.

Methieu Nirhomme, Bruno Messen, Fred Friffiths, PorJr, Bart Jacobs, Rasoka Thor, Batr Criel and Wim Van Damme, “Improving access to hospital care for the poor: comparative analysis of four health equity funds in Cambodia’ Health Policy Plan. (2007) 22 (4): 246-262

Ir, et al. 2010. “Assessment of HEF beneficiary identification in three operational health districts in Kampong Cham”

SKY website. 2006. <http://www.sky-cambodia.org>

Wagstaff, A. and M. Pradhan. 2005. “Health Insurance Impacts on Health and Nonmedical Consumption in a DevelopingCountry.” World Bank Policy ResearchWorking Paper 3563.World Health Organization. 2007. Cambodia

Table 14. Regression on % Any HC visit, -dprobit- model

	(1) HH-months either HEF members or VMs, pooled data	(2) HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, pooled data	(3)	(4) HH-months that in common support group (wealth scores 1-9) that are either HEF members or VMs and transportation data available, pooled data	(5)	(6)	(7)
HEF member	-1.074472***	-0.544178***	-0.569557***	-0.354642***	-0.423136***	-0.463785***	-0.616239***
	(-76.51)	(-15.86)	(-13.56)	(-6.35)	(-5.25)	(-5.64)	(-6.97)
Transp. Cost				-0.971233***	-1.154932***	-1.163932***	-1.121524***
				(-4.78)	(-4.57)	(-4.62)	(-4.39)
Transp. Cost*HEF					0.515950	0.553889	0.497780
					(1.25)	(1.38)	(1.23)
In Kampot District						0.166728**	0.153639**
						(2.90)	(2.67)
VM discount							-0.255874***
							(-4.32)
Constant	-0.356934***	-0.719528***	-0.728738***	-0.653048***	-0.633894***	-0.639287***	-0.461192***
	(-37.31)	(-27.41)	(-12.51)	(-10.01)	(-9.36)	(-9.45)	(-5.70)
p. (VM discount=HEF)							0.0000242
N. wealth dummies	9	9	9	9	9	9	9
N. HH-month	352627	64625	64625	38751	38751	38751	38751
N. Household	17073	2017	2017	1249	1249	1249	1249

t statistics in parentheses, SEs adjust for clustering at the household level
Wealth dummies omitted
p. (VM discount=HEF): Probability of VMs paying at deep discounts = HEF members
Transportation Cost: Hours to transport to HC by a moto
Fraction wealth scores rounded down to nearest integer
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table. 15 Regression on % Any hospital visit, -dprobit- model

	(1)	(2)	(3)	(4)	(5)
	HH-months either HEF members or VMs, pooled data	HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, pooled data			
HEF member	-0.628653***	-0.143468***	-0.232589***	-0.238016***	-0.388160***
	(-36.20)	(-3.72)	(-5.30)	(-5.49)	(-8.07)
Wealth score=1			0.186034*	0.124848	0.0867674
			(2.27)	(1.49)	(1.05)
Wealth score=2			0.227363**	0.146706	0.110575
			(2.63)	(1.70)	(1.29)
Wealth score=3			0.167858*	0.0800289	0.0537312
			(2.00)	(0.92)	(0.63)
Wealth score=4			0.106304	0.0451863	0.00445794
			(1.23)	(0.51)	(0.05)
Wealth score=5			0.150945	0.0802642	0.0627969
			(1.75)	(0.98)	(0.77)
Wealth score=6			0.153187	0.127946	0.113249
			(1.77)	(1.48)	(1.32)
Wealth score=7			0.0146622	-0.00117858	-0.0100321
			(0.18)	(-0.01)	(-0.12)
Wealth score=8			-0.0433326	-0.0506875	-0.0358140
			(-0.52)	(-0.61)	(-0.43)
Wealth score=9			Omitted	Omitted	Omitted
In Kampot District				0.357697***	0.333705***
				(7.51)	(7.04)
VM discount					-0.292513***
					(-5.68)
Constant	-1.433579***	-1.726332***	-1.795319***	-1.820048***	-1.633581***
	(-142.43)	(-66.77)	(-30.45)	(-31.06)	(-24.15)
p. (VM discount=HEF)					0.0574
N. wealth dummies	9	9	9	9	9
N. HH-month	352627	64625	64625	64625	64625
N. Household	17073	2017	2017	2017	2017

t statistics in parentheses, SEs adjust for clustering at the household level * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
p. (VM discount=HEF): Probability of VMs paying at deep discounts = HEF members
Fraction wealth scores rounded down to nearest integer

Table 16. Regression on number of hospital visit, OLS model

	(1)	(2)	(3)	(4)	(5)
	HH-months either HEF members or VMs, pooled data	HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, pooled data			
HEF member	-0.134439***	-0.0166663**	-0.0251593***	-0.0254780***	-0.0568876***
	(-27.71)	(-2.61)	(-3.42)	(-3.48)	(-5.88)
In Kampot District				0.0476338***	0.0442291***
				(4.96)	(4.64)
VM discount					-0.0518600***
					(-4.88)
Constant	0.169117***	0.0692480***	0.0807424***	0.0683867***	0.100592***
	(36.58)	(13.81)	(7.76)	(6.22)	(7.87)
p. (VM discount=HEF)					0.548
N. wealth dummies	9	9	9	9	9
N. HH-month	352627	64625	64625	64625	64625
N. Household	17073	2017	2017	2017	2017

t statistics in parentheses
SEs adjust for clustering at the household level
p. (VM discount=HEF): Probability of VMs paying at deep discounts = HEF members
Fraction wealth scores rounded down to nearest integer
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 17.Regression on ln(total HC & hospital costs), OLS model

	(1)	(2)	(3)	(4)	(5)
	HH-months either HEF members or VMs, pooled data	HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, pooled data			
HEF member	-2.587965***	-1.110255***	-1.185482***	-1.189358***	-1.798241***
	(-74.35)	(-13.88)	(-12.38)	(-12.49)	(-12.80)
In Kampot District				0.579064***	0.513064***
				(5.79)	(5.18)
VM discount					-1.005323***
					(-6.39)
Constant	3.284803***	2.080240***	2.168278***	2.018075***	2.642374***
	(102.45)	(29.28)	(17.40)	(15.99)	(16.22)
p. (VM discount=HEF)					6.83e-14
N. wealth dummies	9	9	9	9	9
N. HH-month	352627	64625	64625	64625	64625
N. Household	17073	2017	2017	2017	2017

t statistics in parentheses
SEs adjust for clustering at the household level
p. (VM discount=HEF): Probability of VMs paying at deep discount = HEF members
Transportation Cost: Hours cost to transport to HC by a moto
Fraction wealth scores rounded down to nearest integer
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table. 18 OLS Regression on number of HC visit

	(1) HH-months either HEF members or VMs, pooled data	(2) HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, pooled data	(3)	(4) HH-months that in common support group, enrolled in HEF or SKY, distance info. available, pooled data	(5)	(6)	(7a)	(7b)
HEF member	-1.004442***	-0.490682***	-0.494330***	-0.333419***	-0.447124***	-0.455133***	-0.613301***	-0.361793***
	(-52.41)	(-12.49)	(-9.83)	(-6.34)	(-6.27)	(-6.37)	(-6.97)	(-4.66)
Wealth score=1			-0.0345444	-0.168162	-0.161502	-0.165634	-0.185745	-0.185745
			(-0.33)	(-1.66)	(-1.60)	(-1.63)	(-1.82)	(-1.82)
Wealth score=2			-0.0238437	-0.118322	-0.115126	-0.119527	-0.139648	-0.139648
			(-0.24)	(-1.12)	(-1.10)	(-1.12)	(-1.32)	(-1.32)
Wealth score=3			-0.0364240	-0.0884152	-0.0802569	-0.0860726	-0.100509	-0.100509
			(-0.37)	(-0.73)	(-0.67)	(-0.70)	(-0.81)	(-0.81)
Wealth score=4			-0.0990323	-0.149620	-0.140776	-0.145227	-0.172604	-0.172604
			(-1.03)	(-1.34)	(-1.26)	(-1.29)	(-1.52)	(-1.52)
Wealth score=5			-0.0924472	-0.154530	-0.157477	-0.162030	-0.168879	-0.168879
			(-0.94)	(-1.41)	(-1.44)	(-1.46)	(-1.52)	(-1.52)
Wealth score=6			0.00706121	0.0300505	0.0329296	0.0321059	0.0262677	0.0262677
			(0.07)	(0.25)	(0.28)	(0.27)	(0.22)	(0.22)
Wealth score=7			-0.0706786	-0.0521741	-0.0543527	-0.0564852	-0.0658136	-0.0658136
			(-0.69)	(-0.47)	(-0.49)	(-0.50)	(-0.59)	(-0.59)
Wealth score=8			-0.106894	-0.0772502	-0.0751036	-0.0759487	-0.0701484	-0.0701484
			(-1.05)	(-0.73)	(-0.71)	(-0.72)	(-0.67)	(-0.67)
Wealth score=9			0	0	0	0	0	0
			(.)	(.)	(.)	(.)	(.)	(.)
Transportation Cost				-0.800400***	-1.139363***	-1.140616***	-1.098475***	-1.098475***
				(-5.14)	(-4.98)	(-5.00)	(-4.77)	(-4.77)
Transp. cost*HEF					0.809718**	0.817565**	0.765604**	0.765604**
					(2.94)	(2.99)	(2.78)	(2.78)
In Kampot District						0.0351946	0.0251107	0.0251107
						(0.71)	(0.50)	(0.50)
VM discount							-0.251508**	
							(-3.05)	
VM normal								0.251508**
								(3.05)
Constant	1.178513***	0.725453***	0.778742***	0.819420***	0.857434***	0.856360***	1.036105***	0.784597***
	(63.28)	(19.48)	(9.21)	(8.86)	(8.82)	(8.84)	(8.49)	(8.34)
p. (VM discount=HEF)							0.00000357	
p. (VM normal=HEF)								5.13e-12
# of HH-months	352627	64625	64625	38751	38751	38751	38751	38751
# of Households	17073	2017	2017	1249	1249	1249	1249	1249

t statistics in parentheses, SEs adjust for clustering at the household level * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
p. (VM discount=HEF): Probability of VMs paying at deep discount = HEF members
Transportation Cost: Hours cost to transport to HC by a moto
Fraction wealth scores rounded down to nearest integer

Table. 19 OLS Regression on number of hospital visit

	(1) HH-months either HEF members or VMs, pooled data	(2) HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, pooled data	(3)	(6) HH-months that in common support group, enrolled in HEF or SKY, distance info. available, pooled data	(7a)	(7b)
HEF member	-0.134439***	-0.0166663**	-0.0251593***	-0.0254780***	-0.0568876***	-0.00502753
	(-27.71)	(-2.61)	(-3.42)	(-3.48)	(-5.88)	(-0.60)
Wealth score=1			0.0166237	0.00989720	0.00413328	0.00413328
			(1.23)	(0.72)	(0.30)	(0.30)
Wealth score=2			0.0155525	0.00733971	0.00173292	0.00173292
			(1.11)	(0.54)	(0.13)	(0.13)
Wealth score=3			0.0148336	0.00493785	0.000399053	0.000399053
			(1.07)	(0.35)	(0.03)	(0.03)
Wealth score=4			0.00780532	0.000482667	-0.00549565	-0.00549565
			(0.58)	(0.03)	(-0.40)	(-0.40)
Wealth score=5			0.00792917	0.000493809	-0.00290017	-0.00290017
			(0.60)	(0.04)	(-0.22)	(-0.22)
Wealth score=6			0.0110598	0.00808352	0.00529276	0.00529276
			(0.78)	(0.57)	(0.37)	(0.37)
Wealth score=7			0.00566648	0.00389056	0.00136457	0.00136457
			(0.35)	(0.24)	(0.08)	(0.08)
Wealth score=8			-0.0132281	-0.0137523	-0.0127306	-0.0127306
			(-0.97)	(-1.01)	(-0.95)	(-0.95)
Wealth score=9			0	0	0	0
			(.)	(.)	(.)	(.)
In Kampot District				0.0476338***	0.0442291***	0.0442291***
				(4.96)	(4.64)	(4.64)
VM discount					-0.0518600***	
					(-4.88)	
VM normal						0.0518600***
						(4.88)
Constant	0.169117***	0.0692480***	0.0659088***	0.0634489***	0.100192***	0.0483324***
	(36.58)	(13.81)	(6.52)	(6.31)	(7.24)	(5.10)
p. (VM discount=HEF)					0.548	
p. (VM normal=HEF)						4.80e-09
# of HH-months	352627	64625	64625	64625	64625	64625
# of Households	17073	2017	2017	2017	2017	2017

t statistics in parentheses
SEs adjust for clustering at the household level
p. (VM discount=HEF): Probability of VMs paying at deep discount = HEF members
Transportation Cost: Hours cost to transport to HC by a moto
Fraction wealth scores rounded down to nearest integer
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table. 20 Regression on number of HC visit, zero-inflated negative binomial model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	HH-months either HEF members or VMs, pooled data	HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, pooled data		HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, distance information available pooled data			
HC Visit							
HEF member	-0.357651***	-0.301099***	-0.282493***	-0.278271***	-0.291895***	-0.256846***	-0.270587***
	(-22.88)	(-8.03)	(-5.03)	(-4.97)	(-4.08)	(-3.64)	(-3.81)
Wealth score=1			-0.0740627	-0.187060*	-0.185806*	-0.162974	-0.166205
			(-0.64)	(-2.14)	(-2.13)	(-1.79)	(-1.81)
Wealth score=2			-0.154327	-0.188837*	-0.188111*	-0.165613	-0.170956
			(-1.65)	(-2.21)	(-2.21)	(-1.87)	(-1.90)
Wealth score=3			-0.0807752	-0.0667723	-0.0656129	-0.0490106	-0.0496441
			(-0.83)	(-0.51)	(-0.50)	(-0.37)	(-0.38)
Wealth score=4			-0.151217	-0.112187	-0.110656	-0.101306	-0.105694
			(-1.82)	(-1.14)	(-1.12)	(-1.02)	(-1.05)
Wealth score=5			-0.156339	-0.154236	-0.154861	-0.148405	-0.147598
			(-1.86)	(-1.71)	(-1.71)	(-1.62)	(-1.62)
Wealth score=6			-0.116853	-0.0813056	-0.0810259	-0.0770021	-0.0762829
			(-1.52)	(-0.97)	(-0.97)	(-0.91)	(-0.91)
Wealth score=7			-0.0763871	-0.0648289	-0.0645425	-0.0582538	-0.0570386
			(-0.94)	(-0.76)	(-0.76)	(-0.68)	(-0.67)
Wealth score=8			-0.162531*	-0.155977*	-0.155651*	-0.150674*	-0.148763*
			(-2.29)	(-2.19)	(-2.19)	(-2.06)	(-2.05)
Wealth score=9			0	0	0	0	0
			(.)	(.)	(.)	(.)	(.)
Transp. Cost				-0.356827*	-0.386164*	-0.372665*	-0.367407*
				(-2.41)	(-2.07)	(-2.00)	(-1.97)
Transp. cost*HEF					0.107373	0.0975700	0.0895847
					(0.39)	(0.35)	(0.32)
In Kampot District						-0.129440*	-0.131603*
						(-2.18)	(-2.17)
Transp. cost*HEF							0
							(.)
VMs paying at deep discount							-0.0280601
							(-0.52)
Constant	1.184321***	1.123356***	1.219979***	1.222249***	1.224695***	1.229075***	1.246296***
	(130.03)	(40.75)	(21.09)	(19.30)	(18.93)	(18.61)	(16.19)
inflate							
Any HC	-59.60413***	-54.65360***	-54.40450***	-53.36041***	-53.36046***	-53.36239***	-53.36211***
	(-3064.21)	(-1082.14)	(-1077.16)	(-848.25)	(-848.17)	(-848.29)	(-848.03)
Constant	30.76188***	27.33691***	27.21282***	26.90863***	26.90868***	26.91061***	26.91033***
	(2984.84)	(1136.87)	(1130.88)	(893.25)	(893.41)	(893.36)	(893.11)
lnalpha							
Constant	-1.405171***	-1.707445***	-1.721102***	-1.867264***	-1.867318***	-1.876769***	-1.877673***
	(-55.07)	(-17.55)	(-17.88)	(-18.55)	(-18.56)	(-18.59)	(-18.56)
p. (VM discount=HEF)							0.00265
# HH-month	352627	64625	64625	38751	38751	38751	38751
# Household	17073	2017	2017	1249	1249	1249	1249

t statistics in parentheses
SEs adjust for clustering at the household level
p. (VM discount=HEF): Probability of VMs paying at deep discount = HEF members
Transportation Cost: Hours cost to transport to HC by a moto
Fraction wealth scores rounded down to nearest integer
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table. 21 Regression on ln(total HC cost), top-coded by 99%

	(1)	(2)	(3)	(4)	(5)	(6)	(7a)	(7b)
	HH-months either HEF members or VMs, pooled data	HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, pooled data		HH-months that in common support group (wealth scores 1-9) and are either HEF members or VMs, , distance information available, pooled data				
HEF member	-0.465274***	-0.209002***	-0.215991***	-0.140529***	-0.185995***	-0.201821***	-0.278173***	-0.156800***
	(-64.01)	(-13.51)	(-11.82)	(-6.64)	(-6.30)	(-6.82)	(-7.45)	(-5.08)
Wealth score=1			0.000434921	-0.0524284	-0.0499848	-0.0585202	-0.0677797	-0.0677797
			(0.01)	(-1.23)	(-1.18)	(-1.39)	(-1.62)	(-1.62)
Wealth score=2			0.00493745	-0.0292141	-0.0279256	-0.0365291	-0.0457975	-0.0457975
			(0.13)	(-0.66)	(-0.64)	(-0.83)	(-1.06)	(-1.06)
Wealth score=3			-0.00285209	-0.0402085	-0.0373256	-0.0493542	-0.0562316	-0.0562316
			(-0.08)	(-0.91)	(-0.85)	(-1.11)	(-1.26)	(-1.26)
Wealth score=4			-0.0298312	-0.0552664	-0.0515107	-0.0600125	-0.0722329	-0.0722329
			(-0.78)	(-1.23)	(-1.15)	(-1.32)	(-1.60)	(-1.60)
Wealth score=5			-0.0281278	-0.0700452	-0.0713053	-0.0804692	-0.0828583	-0.0828583
			(-0.74)	(-1.65)	(-1.68)	(-1.88)	(-1.95)	(-1.95)
Wealth score=6			0.0102336	0.0109247	0.0120748	0.0103659	0.00775414	0.00775414
			(0.25)	(0.24)	(0.26)	(0.23)	(0.17)	(0.17)
Wealth score=7			-0.0318840	-0.0291068	-0.0300315	-0.0342972	-0.0384121	-0.0384121
			(-0.80)	(-0.67)	(-0.69)	(-0.79)	(-0.89)	(-0.89)
Wealth score=8			-0.0356099	-0.0210107	-0.0204500	-0.0221522	-0.0188922	-0.0188922
			(-0.86)	(-0.48)	(-0.47)	(-0.51)	(-0.44)	(-0.44)
Wealth score=9			0	0	0	0	0	0
			(.)	(.)	(.)	(.)	(.)	(.)
Transportation Cost				-0.348204***	-0.483804***	-0.486323***	-0.465779***	-0.465779***
				(-4.77)	(-4.72)	(-4.83)	(-4.58)	(-4.58)
Transp. cost*HEF					0.323941*	0.339590**	0.314190*	0.314190*
					(2.39)	(2.62)	(2.41)	(2.41)
In Kampot District						0.0698790**	0.0649558**	0.0649558**
						(3.01)	(2.80)	(2.80)
VM discount							-0.121373***	
							(-3.66)	
VM normal								0.121373***
								(3.66)
Constant	7.474192***	7.251971***	7.268223***	7.288430***	7.303725***	7.301616***	7.387959***	7.266587***
	(1080.06)	(510.29)	(222.33)	(201.13)	(191.85)	(192.87)	(155.89)	(196.35)
sigma								
Constant	0.632498***	0.553171***	0.552929***	0.588215***	0.587915***	0.587334***	0.585820***	0.585820***
	(161.64)	(57.25)	(57.30)	(49.23)	(49.33)	(49.34)	(49.68)	(49.68)
p. (VM discount=HEF)							0.000000371	
p. (VM normal=HEF)								9.62e-14
# HH-month	352627	64625	64625	38751	38751	38751	38751	38751
# Household	17073	2017	2017	1249	1249	1249	1249	1249

t statistics in parentheses, SEs adjust for clustering at the household level, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

p. (VM discount=HEF): Probability of VMs paying at deep discount = HEF members, p. (VM normal=HEF)= Probability of VMs paying at normal price = HEF members)

Transportation Cost: Hours cost to transport to HC by a moto, Fraction wealth scores rounded down to nearest integer