博　士　論　文

Developing and validating regression models for predicting household consumption for the health insurance in Cambodia

（カンボジアの医療保険のための家計消費推計モデルの構築と検証）

中　村　治　代

Developing and validating regression models for predicting household consumption for the health insurance in Cambodia

（カンボジアの医療保険のための家計消費推計モデルの構築と検証）

In partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

A thesis presented
by
Haruyo Nakamura
(41-187104)
Admission in April 2018
Supervisor: Professor Masahiro Hashizume
to
Department of Global Health Policy
School of International Health
Graduate School of Medicine
The University of Tokyo

Tokyo, Japan
2021

Developing and validating regression models for predicting household consumption for the health insurance in Cambodia

（カンボジアの医療保険のための家計消費推計モデルの構築と検証）

Abstract

Background: Health financial protection is a challenge for low- and middle-income countries (LMICs), where the general tax revenue is limited. Although contributory insurance allows LMICs to effectively mobilize domestic financial resources, it requires compulsory participation realized through capacity-based contribution, that is another challenge for LMICs where most of the population is engaged in the informal economy. This study developed and validated a household consumption predictive model for Cambodia to extend the contributory health insurance to the general population.

Methods: This study used data from nationally representative surveys involving 38,472 households. Four alternative linear models were developed. Out-of-sample cross-validation was performed, and the prediction performance was evaluated using statistical measurements. Subsequently, ten options of hypothetical insurance contribution were estimated based on the insurance fee schedule and the healthcare

utilization assumptions. Finally, equity and progressivity of each option were evaluated using the Kakwani index.

Results: A linearly positive relationship was found between observed and predicted household consumptions. The model with a backward-selection technique within a stepwise analytical framework was found to best suit Cambodia. The prediction-based household contribution was comparable to that on an observation basis. Capacity-based contribution was suggested to increase relative amount of insurance revenue, reduce burden on the poor, and eventually mitigate regressivity of the current healthcare financing in Cambodia.

Conclusions: This study suggested the possibility of developing a regression model using population survey data to make reasonable prediction of household consumption that may support the policy discussion for capacity-based health insurance system in Cambodia. The model should be tested in real settings, re-evaluated and improved periodically.

Key words: Health financing, Health insurance, Contribution, Equity, Cambodia

Table of Contents

Acknowledgements

I would like to thank my supervisors Professor Masahiro Hashizume and Professor Kenji Shibuya for their support throughout the years. They provided me with valuable guidance and taught me importance of persistent work towards research questions. I would like to thank my primary supervisor, Dr. Shuhei Nomura for his continuous support starting from the very beginning with designing the research to the last moment of completing the dissertation. I would also like to thank Assistant Professor, Md. Mizanur Rahman for his input to improve the contents of the study with his broad knowledge on universal health coverage in various countries. I also owe my accomplishment to my colleague, Dr. Floriano Amimo, who supported me with his abundant knowledge and experience in research.

Outside the university, I would like to express my appreciation to the Japan International Cooperation Agency that provided me with an opportunity to boldly tackle the grand challenge of finding a strategy for Cambodia to move towards UHC goals. I am also grateful for my Cambodian colleagues, the officers in Ministries and development partners who collaboratively worked with me on the hard tasks. I also thank Professor Kenji Shimazaki of the International University of Health and Welfare Graduate School,

who taught me the history and overall concept of the Japanese universal health coverage, Dr. Ayako Honda of the Sofia University, who gave me encouragement to pursue this research question, and Dr. Ryota Nakamura of the Hitotsubashi University, who gave me valuable suggestions to make my research meaningful to the society.

I would also like to thank my fellow friends and colleagues at the Department of Global Health Policy who supported me and helped me resolve the issues which raised in the course of the PhD program. I would also like to thank my family, colleagues at the firm, and friends, both in Japan and in Cambodia, for supporting my PhD research work technically and psychologically. Without their support, I wouldn't be able to complete the course work.

Finally, I would like to dedicate my thesis to the memory of Madam Sokha, the former chair of the Executive Committee, the National Social Protection Council in Cambodia, who has laid the foundation of Cambodian social protection system, and Dr. Bart Jacobs, the former GIZ health advisor and the P4HC+ coordinator, who has established a harmonious network of development partners in Cambodia. I will never forget their heroism and would like to take over them together with other partners.

List of Figures

List of Tables

List of Abbreviations

ASEAN	Association of Southeast Asian Nations
CHE	Current Health Expenditure
CI	Confidence Interval
CPA	Complementary Package of Activities
CSES	Cambodia Socio-Economic Survey
CTP	Capacity to Pay
EA	Enumeration Area
EU	European Union
GDP	Gross Domestic Product
GGE	General Government Expenditure
HEF	Health Equity Fund
IQR	Interquartile Range
KI	Kakwani index
LASSO	Least Absolute Shrinkage and Selection Operator
LMIC	Low- and middle-income country
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MEF	Ministry of Economy and Finance
MOH	Ministry of Health
MPA	Minimum Package of Activities
NGO	Non-governmental Organization
NHI	National Health Insurance
NIS	National Institute of Statistics

NSSF	National Social Security Fund
OLS	Ordinary Least Squares
OOP	Out-of-pocket Payment
PSU	Primary Sampling Unit
Q1-5	Quintiles 1-5
RMSE	Root Mean Squared Error
UHC	Universal Health Coverage
UNICEF	United Nations Children's Fund
WHO	World Health Organization

1. INTRODUCTION

1.1. Background: universal health coverage

The World Health Assembly endorsed Resolution WHA 58.33 in 2005, urging its member states to ensure access to affordable healthcare services for prevention, promotion, treatment, and rehabilitation.[1] A decade later, universal health coverage (UHC) became a target (3.8) of the 2030 Sustainable Development Goal 3.[2] Progress towards UHC has been assessed by monitoring coverage of essential health services and financial protection in each nation.[3, 4]

Overall global coverage of essential health services reportedly increased from 2000 to 2017, albeit at a slow pace.[4] On the other hand, financial protection deteriorated. The proportion of the population with catastrophic health expenditure, defined as out-of-pocket (OOP) spending exceeding 10% or 25% of their household total consumption or income, rose from 9.4% to 12.7% at the 10% threshold, and from 1.7% to 2.9% at 25% threshold between 2000 and 2015. Furthermore, based on a relative poverty line defined as 60% of median daily per capita consumption or income, the percentage of the population impoverished by out-of-pocket health spending increased from 1.8% to 2.5% during the same period.[4] In 2016, over 3.6 billion people, roughly half of the

world population, did not receive the essential health services they needed due to service unavailability or unaffordability.[3] Between 2000 and 2015, the largest concentration of the world population with catastrophic health spending shifted from low-income countries to middle-income countries, while around 70% of the population with catastrophic health spending was persistently concentrated in Asia.[5]

A comprehensive assessment of UHC in 111 countries suggested that strong UHC performance is correlated with the share of a country's health budget channeled through government and social health insurance schemes.[6] Other evidence suggests that financial protection can only be universally available if it is backed by funds from prepaid and pooled sources with subsidies for the indigent.[7] The entitlement to guaranteed services should also be universal, not linked to employment status.[7] Meanwhile, it was also suggested that an important factor for UHC achievement is the equitable revenue collection.[8]

1.2. Health financing and financial protection

In health economics, it is generally assumed that the market mechanism does not function in health because of its peculiarities, including (1) information asymmetry

between patients and physicians, (2) the unpredictable and urgent nature of diseases, and (3) incomplete competitive-markets of health institutions which are often geographically not equitably distributed across a country.[9-11] Additionally, an experiment in the United States in the 1970s suggested that health of the poor and those with specific medical conditions would deteriorate without financial protection.[12] Therefore, some form of public intervention is necessary in health financing, and financial protection should be provided for the vulnerable population.

There are two distinct public health financing approaches. One is through taxation, known as the Beveridge model [13], and the other is a contributory insurance, primarily known as the Bismarck model [14], in which people pay a fee to an insurance fund that in turn pays for healthcare activities provided by a state-owned or a private institution.[15] Tax-based financing can improve efficiency [16] as the government is the healthcare provider as well as purchaser. Whereas, a contributory insurance can secure the contribution revenue for the healthcare services.[16, 17] It also allows a country to maintain financial discipline by establishing a contribution level to balance revenues and expenditures.[18, 19] Furthermore, the contractual relationship established between the insured and the insurer in a contributory insurance can encourage healthcare

providers to improve their services.[19] Today, the former approach is adopted in the United Kingdom, the Scandinavian countries, and New Zealand, and the latter in Germany, France, and the Netherlands. Most other countries' healthcare systems have elements of either or both models.[15]

1.3. Challenges for low- and middle-income countries

In 2019, the World Bank warned that most of the low- and middle-income countries (LMICs) would fail to achieve their targets for UHC unless they took urgent steps to strengthen their health financing.[7] However, it is financially challenging for LMICs to establish a health system by employing the Beveridge model due to limited tax revenue.[20] It was reported that government efforts to raise taxes consistently fell short of 15 percent of gross domestic product (GDP), a threshold that the International Monetary Fund has identified as critical to endanger sustained, inclusive growth in nearly half of LMICs.[7] Moreover, a study that explored revenue-raising potentials through different types of taxation in Benin, Mali, Mozambique and Togo concluded that such efforts could raise limited amount of revenue.[20]

On the other hand, a contributory insurance, i.e., the Bismarck model, effectively mobilizes domestic financing resources.[21] The Bismarck model was originally

introduced in Germany in 1883 to provide cures to the workers and their families [14], and it came to be adopted in other countries primarily by enrolling the workers engaged in the formal sector, such as civil servants and private sector employees, since insurance contribution can be deducted from their salaries. In this regard, this approach is another challenge for LMICs, where nearly 70% of the labor force is engaged in the informal economy [22], which is, in law or practice, not covered or insufficiently covered by formal arrangements.[23]

Nevertheless, some LMICs still pursue adopting a contributory health insurance to cover their general populations by mobilizing domestic resources. Such countries usually introduce the contributory insurance in a voluntary basis and set a flat contribution rate at a level that the lowest-income group can afford, with subsidy by the government or development partners. Some countries set a simple sliding scale that categorizes households into two to three economic groups, as seen in the Philippines and Rwanda.[24, 25] However, these practices often lack equity [26] and place LMICs at risk of reducing the insurance service coverage, endangering financial sustainability of the insurance fund, or imposing a heavy burden of subsidy on the government, whose fiscal resources are already limited.[21] The National Health Insurance Authority in Ghana

regulates the contribution to be paid according to the member's self-reported income. Nonetheless, there is not a clear guidance regarding how much is to be paid according to a given level of income or information on the registered members' income. Consequently, there is a chance that equal contribution is not collected from members of equal income, and the contribution collection might impede enrollment of the new members or even impoverish the poor.[8] Therefore, there is an urgent need for LMICs to find an alternative approach to make progress towards achieving UHC targets.

1.4. Capacity-based contribution collection from the general population

Some countries, namely Japan, Korea, and Taiwan, have achieved UHC by introducing a contributory health insurance [16], along with the employees' health insurance and social assistance for the indigent, when their income level was relatively low.[18, 27] Among them, Japan was the first country that adopted the contributory insurance for the general population, including the informal sector. Shimazaki pointed out that one of the key factors, that enabled Japan to successfully introduce the National Health Insurance (NHI) for the general population, was presence of the local resident taxation systems [18], which coincided in Korea and Taiwan.[27, 28] It is assumed that the system has played an important role to promote the implementation process, by allowing

contribution collection according to household capacity to pay (CTP), and thus encourage lower income households to be enrolled in the insurance.[19, 27-29]

Unlike the present NHI in which approximately 50% of the revenue is comprised of public fund [19], contribution was the major revenue source when the NHI was first enacted in Japan.[30] Figure 1 shows revenue breakdowns of the NHI insurers in 1947 and today.

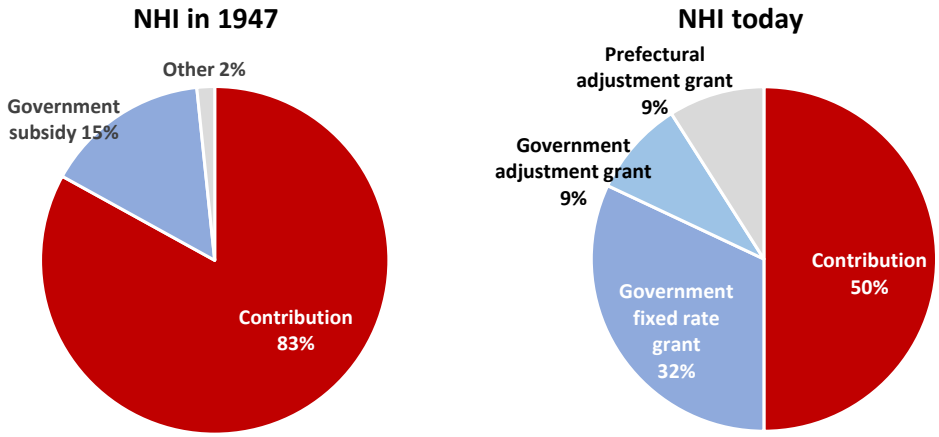


Figure 1: Revenues of the Japanese National Health Insurance associations in 1947 and today

Source: The government subsidy applications submitted by the health insurance associations across the country for nine months in 1947 [30] and the National Health Insurance Act, Article 70 and 72.

Under the NHI Act of 1938, the government subsidy was stipulated to be 0.5 yen per person per year in principle, that was equivalent to 10% of the medical expenses, and the government or a municipality could provide additional subsidy to health insurance associations in a state of emergency.[31] It should be noted that around 50% of co-

payment was additionally required at the point of service then.[30] It is reported that the National Health Insurance was initially managed by associations on a voluntary basis because the Japanese government was reluctant to expand the scope of publicly funded medical services to the general population due to difficulty of securing tax revenue sources.[32]

Table 1 shows an example of contribution schedule in the NHI health insurance association.[29] Although contribution collection method was left to each autonomous association to decide, the majority of the associations followed the official notice of the Ministry of Health and Welfare that recommended contribution collection based on household tax payment in about ten grades.[32] The insurance association in the table was suggested to have collected a total of 11,094 yen from 840 households per year.

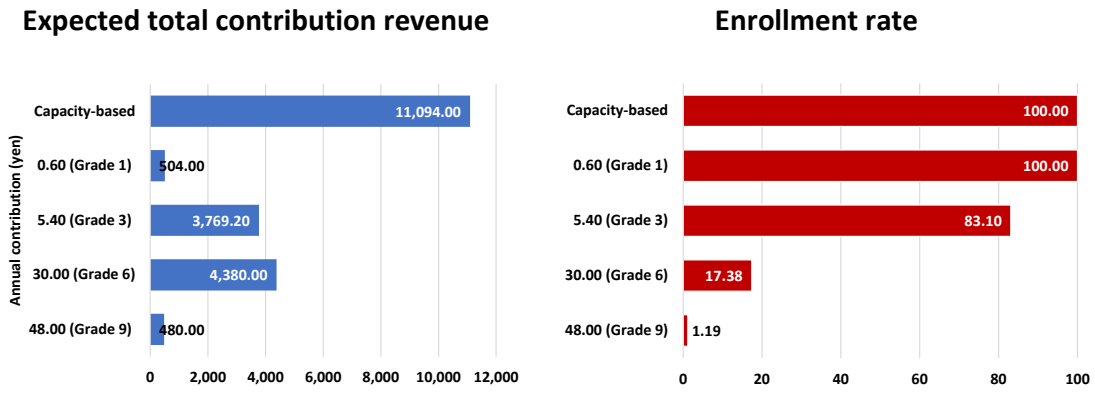
Table 1: Suggested contribution schedule of a health insurance association in Japan

	Payment of household tax (yen)	Number of households	Monthly contribution (yen)	Annual contribution (yen)	Total contribution revenue (yen)
1	0	30	0.05	0.60	18.00
2	<2	112	0.20	2.40	268.80
3	≥2 & <5	188	0.45	5.40	1,015.20
4	≥5 & <10	186	0.70	8.40	1,562.40
5	≥10 & <20	178	1.60	19.20	3,417.60
6	≥20 & <50	112	2.50	30.00	3,360.00
7	≥50 & <100	12	3.00	36.00	432.00
8	≥100 & <200	12	3.50	42.00	504.00
9	≥200 & <400	7	4.00	48.00	336.00
10	≥400	3	5.00	60.00	180.00
	Total	840			11,094.00

Note: This contribution schedule was submitted to the 73rd Imperial Diet as an example along with the National Health Insurance Bill in 1937.

Source: Kawamura H, Ishihara T, Kizuki M [29]

Figure 2 shows expected total contribution revenues and enrollment rates when the above-mentioned association collected flat-rate contributions in different settings in comparison with the capacity-based contribution collection.



Source: Kawamura H, Ishihara T, Kizuki M [29]

Figure 2: Expected total contribution revenues and enrollment rates when flat-rate contributions were collected in comparison with the capacity-based contribution collection in the health Insurance association

When a flat rate was set at 0.60 yen per year, i.e. the rate of Grade 1, all households could join the insurance, but the annual contribution revenue would be only 504.00 yen. When a flat rate was set at 5.40 yen per year, i.e. the rate of Grade 3, the annual contribution revenue would increase to 3,769.20 yen, but the enrollment rate would be decreased to 83% of what could have been collected through capacity-based contribution, assuming that the households in Grades 1 and 2 could not afford the amount of contribution. When the flat rate of Grade 6, i.e. 30.00 yen per year, was applied, the annual contribution revenue would further increase to 4,380.00 yen, but the enrollment rate would even decrease to 17%. Application of the Grade 9 rate, i.e. 48.00 yen per year, would result in reduction of both contribution revenue and enrollment rate to 480.00 yen and 1%, respectively. These suggest that capacity-based

contribution collection has ultimately increased the amount of contribution revenue as well as enrollment rate in the contributory insurance.

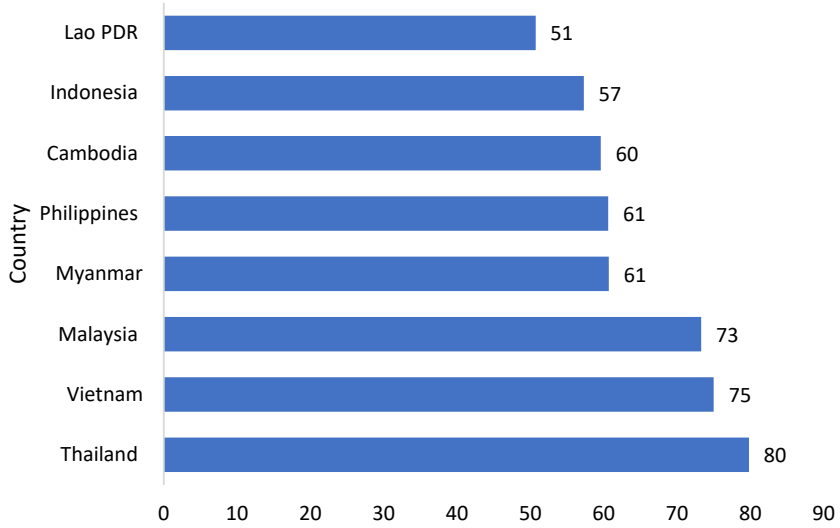
There is also a chance that capacity-based contribution collection has motivated the population to join the health insurance, not only from an economic point of view, but also psychologically. The survey undertaken by the World Health Organization (WHO), conducted for 1,007 respondents from over 100 countries, indicated that more than 70% of the WHO staff and more than 60% of the general public preferred a health system in which everyone contributes an equal share of CTP, regardless of the household health status or use of the health system.[33] The Ghanaian National Health Insurance also set the contribution rates on a capacity basis because it was believed that inequitable contributions by people in the same income group could displace members and hence reduce membership.[8]

Nonetheless, it is assumed that capacity-based contribution was realized in Japan, Kore and Taiwan because they had household taxation system.[18, 27, 28] The challenge for LMICs, where household taxation is not practiced, is how to assess household CTP without tax information.

1.5. Health financing and financial protection in Cambodia

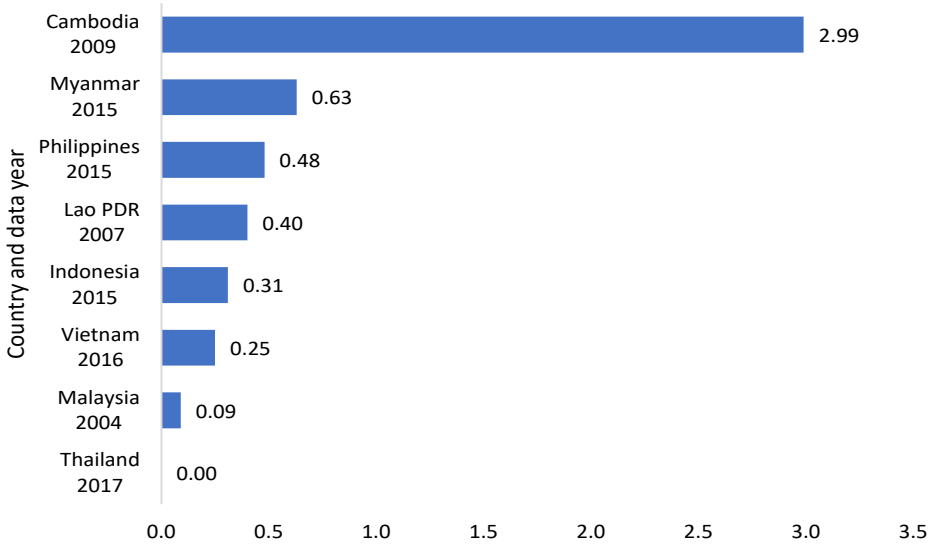
Cambodia is a lower-middle-income country [34] in Southeast Asia with a total population of 15.29 million [35] and GDP per capita of 1,643.12 current US Dollars in 2019.[36] Figure 3 shows the country's progress towards achieving universal health coverage (UHC) from the perspectives of health service coverage and financial protection in comparison with other middle-income countries of the Association of Southeast Asian Nations (ASEAN). While UHC service coverage index that measures achievement of essential health service coverage was higher in Cambodia than Laos and Indonesia, and comparative to the Philippines and Myanmar in 2019, the proportion of households that experienced catastrophic health expenditure in Cambodia was pronouncedly higher than other ASEAN middle-income countries.[4]

Universal Health Coverage Service Coverage Index (UHC SCI) 2019



Note: The UHC SCI is constructed from geometric means of the latest available 14 tracer indicators in the four essential health service areas collected from January to June 2019. The UHC SCI provides a strong signal on the coverage of health services needed by most populations across sociodemographic settings on a scale of 0 to 100, with higher scores indicating better performance. The UHC SCI tracers include (1) reproductive, maternal, newborn and child health (family planning, 4+ visits of antenatal care, diphtheria-tetanus-pertussis immunization and care-seeking for suspected pneumonia), (2) infectious diseases (tuberculosis effective treatment, Human Immunodeficiency Virus (HIV) treatment, insecticide-treated nets and water, sanitation & hygiene), (3) noncommunicable diseases (normal blood pressure, mean fasting plasma glucose, and tobacco nonsmoking), and (4) service capacity and access (hospital bed density, health worker density, and the International Health Regulations core capacity index).[4]

Incidence of catastrophic health expenditure (%)



Note: The threshold is at 10% of household total consumption or income.
Source: World Health Organization, the World Bank [4]

Figure 3: Universal Health Coverage Service Coverage Index (UHC SCI) and incidence of catastrophic health expenditure in the middle-income countries of the Association of Southeast Asian Nations (ASEAN)

Table 2 presents key health financing indicators of Cambodia and the lower-middle-income countries in 2017.[37] Although per capita current health expenditure (CHE) as a proportion of GDP in Cambodia was 0.9 percentage points higher than that in lower-middle-income countries, domestic general government expenditure (GGE) for health in the country was 20.2% percentage points lower than the group average. In addition, coverage of voluntary health insurance was merely 0.6% in Cambodia. As a result, the Cambodian share of OOP payment in CHE was 21.4 percentage points higher than that in the lower-middle-income countries: 60.4% and 39.0%, respectively. It was also reported that the health share in GGE decreased by 2.1 percentage points, despite the fact that overall government expenditure increased by 6.4 percentage points between 2000 and 2017 in Cambodia.[38] These evidences suggest that the Cambodian health system has been highly reliant on OOP payment, and extra health budget is unlikely to be allocated, particularly in the short term.

Table 2: Key health expenditure indicators of Cambodia and lower-middle-income countries (2017)

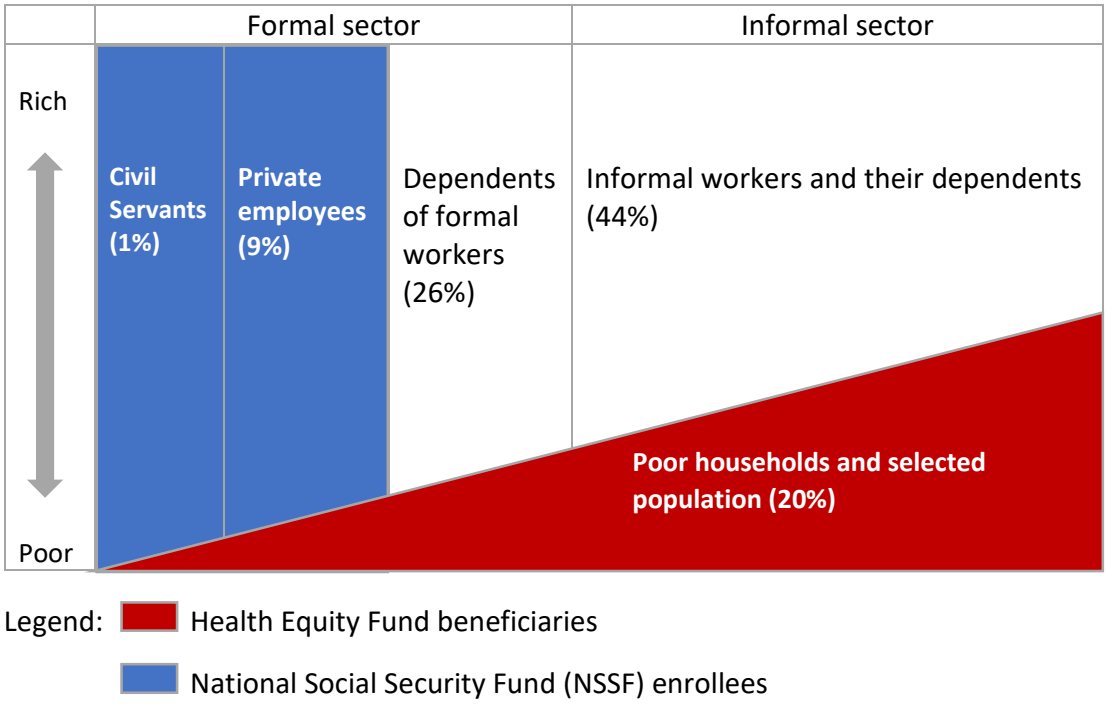
Health expenditure indicators	Lower-middle-income countries	Cambodia
GDP per capita (Current UD Dollars)	1006-3955	1387
CHE per capita (Current UD Dollars)	138	82
CHE per capita (% of GDP)	5.0	5.9
Domestic GGE for health (% of CHE)	44.0	23.8
Domestic GGE for health (% of GGE)	8.0	6.1
Domestic GGE for health (% of GDP)	2.0	1.4
Out-of-pocket (% of CHE)	39.0	60.4
Voluntary health insurance (% of CHE)	2.0	0.6

Note: GDP – Gross Domestic Product, CHE – Current health expenditure, GGE – General Government Expenditure
Source: World Health Organization [37]

1.5.1. The health protection policy frameworks

Figure 4 presents the current health protection system in Cambodia. The HEF is the co-financing social assistance mechanism established by the government and development partners.[39] The HEF covers around 20 % of poor households [40] and population in the specific occupations announced by the Prime Minister in January 2018, including commune council members, village representatives, the government-sponsored athletes, land-mine cleaners, and tricycle drivers.[41, 42] Whereas, about 10% of the formal-sector workers have been enrolled in the NSSF health insurance since 2016.[43] Thus, about 70% of the population is left uninsured in Cambodia. The government plans to extend the NSSF health insurance to the rest of the population by 2025.[44] The

Ministry of Economy and Finance (MEF) acknowledges a need of subsidy to cover the contributions of the poor and vulnerable, as well as additional targeted groups to the extent that the national budget allows. In the meantime, MEF also states that eligibility for full or partial subsidy should be assessed based on clear and appropriate criteria in order to avoid moral hazard.[44]



Source: NSSF [45], Ministry of Planning [40]

Figure 4: Cambodian health financial protection system

1.5.2. The NSSF health insurance scheme

The current NSSF health insurance scheme was established by the Sub-decree on

Establishment of Social Security Scheme on Health Care for Persons Defined by the Provisions of the Labour Law in January 2016.[43] The benefit package of the scheme comprises curative and preventive medical services (in-kind benefits) as specified in the Prakas on Health Care Benefits by the Ministry of Labour and Vocational Training that was adopted in March 2016. [46]

Provider payments are specified as either case-based or fee-for-service methods and defined to reimburse health facilities on the Inter-ministerial Prakas on Provider Payment Methods for Health Care by the Ministry of Labour and Vocational Training and the Ministry of Health.[47] The fee-for-service method is employed for selected high-cost items. It is payable only if decided to be necessary by the designated health facility as a follow-up case covered by the NSSF case-based payment method. The health insurance also covers costs required for patient referral services and corpse transportation.[47] The service provider network consists mainly of public hospitals, including health centers, district (referral) hospitals, provincial hospitals, and national hospitals (NHs). However, it also includes a few private facilities recognized by the Ministry of Health.[46] The case-based fees reflecting types of treatment and levels of care, are specified in the inter-ministerial Prakas on Provider Payment Methods for Health Care adopted in May 2016

[46] and revised in August 2017.[48]

1.5.3. Assessment of household capacity to pay

In Cambodia, the proxy means test, a widely used poverty assessment tool [49] has been practiced to identify poor households as beneficiaries of the social assistance programs, including the HEF.[40] The proxy means test has been carried out based on the questionnaires that consist of scoring and non-scoring proxy indicators that differentiate poor households from non-poor households.[40] The proxy means test assesses the level of poverty by scoring households. However, it does not provide information on how much a household earns or spends. In addition, performance of the proxy means test has not been regularly evaluated, while the results of the tool have been verified through discussions in communities.[40] Therefore, it is not most appropriate to use the proxy means test for insurance contribution estimates that require reliable information on the household CTP in monetary form.

On the other hand, the National Institute of Statistics (NIS), under the Ministry of Planning in Cambodia, annually conducts the Cambodia Socio-Economic Survey (CSES), which estimates social and economic status, including household income and

consumption which measure household CTP in the country.[50] The CSES provides information on the sample household CTP, and estimates a median of the Cambodian household CTP. However, it does not provide information on how much each household earns or spends, which is necessary for insurance contribution estimation. Besides, the CSES has the lengthy questionnaires, which are unlikely to be used by the local administrative staff typically with minimum capacity and maximum workload. Thus, the CSES data could be used to develop an efficient regression model with a limited number of indices, that allows an easy, quick, but accurate assessment of the household CTP.

Studies have attempted to develop efficient scales to measure household welfare or poverty status, mainly for social assistance programs [40, 49, 51-56], or a singular value decomposition, such as a principal component analysis, for research purposes.[57-59] Nonetheless, these tools merely identified poor households or ranked households by their welfare status. A couple of studies have attempted to predict household income or consumption. However, one dichotomized the households using 50,000 US Dollars as a cut-off point.[60] The other predicted national average household incomes for multiple countries.[61] These attempts implied the possibility of estimating household economic status using a limited number of indices. However, no studies have focused on predicting

income or consumption of each household on a monetary basis to be applied for health insurance contribution estimation.

1.6. Aim and objectives

While the health financing policy is to be decided by the government of Cambodia, this study aimed to help the government extend the contributory health insurance to the general population as planned in the policy framework, by collecting contributions according to household CTP. The objectives of this study were: (1) to develop and validate efficient regression models to predict annual household consumption, (2) to estimate hypothetical insurance contribution options, and (3) to evaluate equity and progressivity of the hypothetical insurance contribution options based on predicted household consumptions using the regression model.

1.7. Organization of the thesis

The thesis is organized as follows. Chapter 2 presents the development and validation of regression models that predict annual household consumption in Cambodia with the national cross-sectional, population-based survey data. Chapter 3 presents the estimates of the hypothetical health insurance contribution options using the household

consumption predictive model developed in Chapter 2. Chapter 4 evaluates equity and progressivity of the hypothetical health insurance contribution options. Chapter 5 summarizes the discussion points derived from the preceding chapters and extend them to policy implications. Chapter 6 concludes the thesis with recommendations for health financing policies and future studies in Cambodia.

2. CONSTRUCTION OF HOUSEHOLD CONSUMPTION PREDICTIVE MODEL

2.1. Objective

This chapter presents the development and validation of regression models that predict annual household consumption in Cambodia with the national cross-sectional, population-based survey data.

2.2. Methods

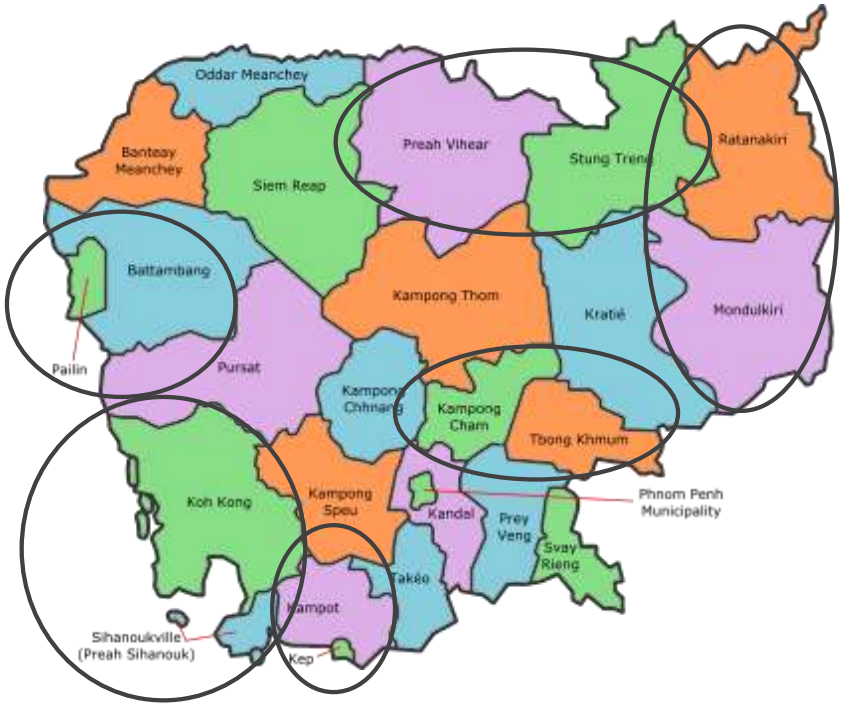
2.2.1. Data source

This study used the data of the Cambodia Socio-Economic Survey (CSES) conducted annually between 2010 and 2017 [50, 62-68], that are made available for research and analysis according to the procedure specified in the Statistics Law.[69] The CSES is a nationally representative cluster sample survey. The CSES was initiated in 1993/1994, and has been conducted annually since 2007 by the NIS, under the Ministry of Planning. A large-scale survey is conducted every five years. The CSES data are used in the calculation of national accounts, income, and agricultural statistics, as wells as statistics on issues of vulnerability and victimization. The main user of the CSES is the Royal Government of Cambodia, university researchers, analysts, international organizations, such as the World Bank and NGOs.[50]

The CSES uses systematic sampling with probabilities proportional to the size, based on the number of households per village retrieved from the public information sources, including the General Population Census, the Cambodian Inter-censal Population Survey, the National Census of Agriculture of Cambodia, the regularly updated administrative information in the Commune Data Base and official information from the Ministry of Interior.[50] Probability sampling weights were added over all sample households within each stratum. The sum of the weights is an estimate of the total number of households in the stratum. This estimate was compared to the number of households according to demographic projections based on the above-mentioned public information sources.[70] The CSES targets at all people living in normal households.[50] Non-response rate of the CSES is less than 1%.[50]

Figure 5 shows the country's 24 provinces and a municipality that were grouped into 19 in the process of constructing the CSES sampling framework. The provinces that had smaller population were paired with an adjoining province. Each group was further divided into urban and rural strata, and a total of 38 strata were formed. Urban-rural definition of PSUs in the CSES followed the Population Census that defined communes as urban if (1) total population exceeds 2,000, (2) population density exceeds 200 per

km^2 and (3) percentage of male employment in agriculture is below 50%, and the rest were considered as rural.[70] The villages primarily constitute PSUs in the CSES with a few exceptions of large villages that are represented by more than one PSU.

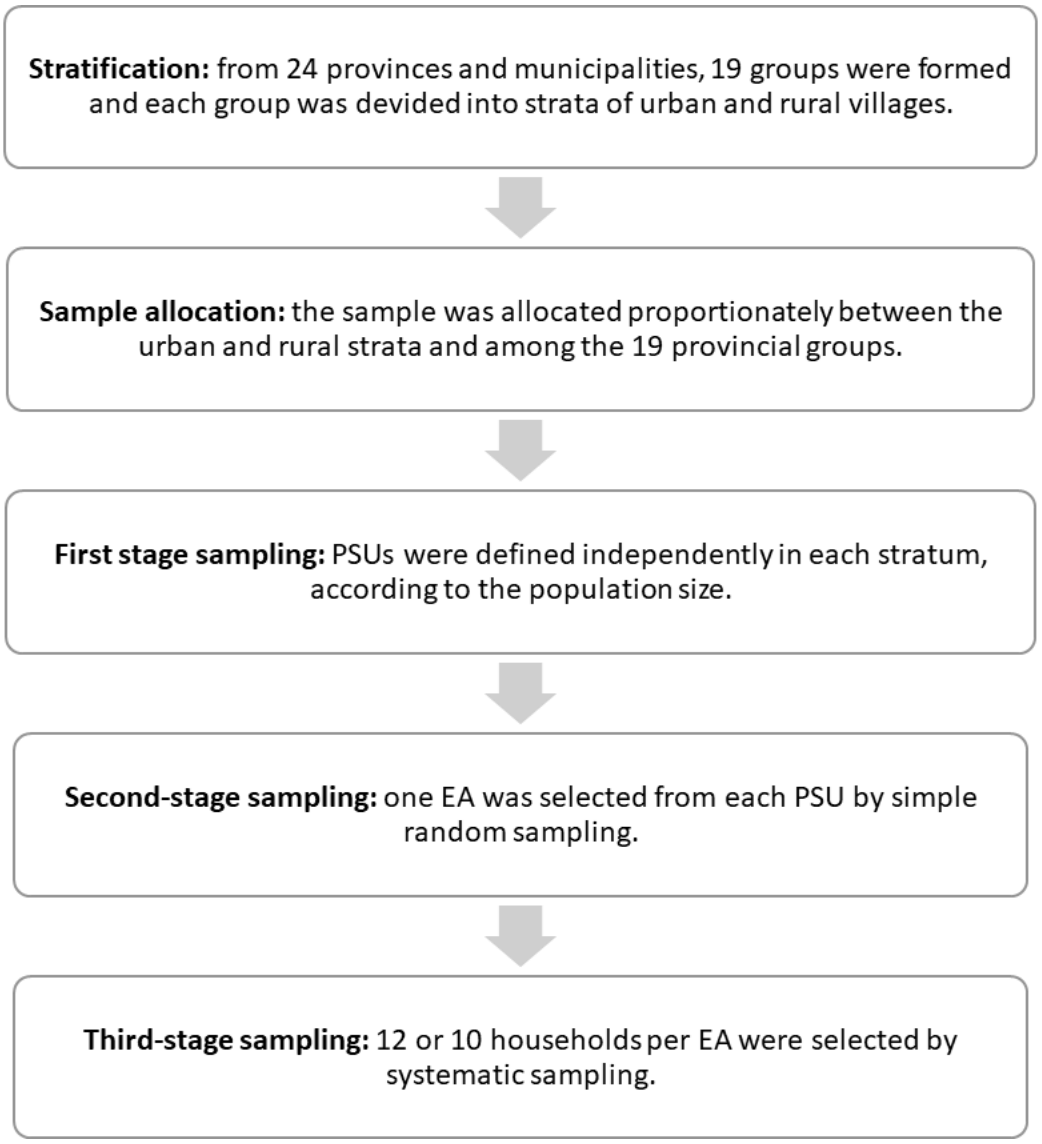


Note: PSU – primary sampling unit, CSES – Cambodia Socio-Economic Survey, a circle indicates a paired provincial group

Source: Cambodia Socio-Economic Survey 2014 [50]

Figure 5: 19 provincial groups constituting the Cambodia Socio-Economic Survey strata

The CSES is designed in the three-stage sampling. Figure 6 describes the procedure of the sampling. First, preliminary data from the latest public information is used to construct the CSES sampling frame. Then a systematic sample with probabilities proportional-to-size is conducted to select PSUs from each stratum. Second, a mapping of enumeration areas (EAs) is conducted in each selected PSU, and one EA is selected from each PSU by simple random sampling. An EA comprises 50 to 100 households.[71] Finally, a household mapping is conducted in each selected EA, and 12 households for a large-scale survey and 10 households for a small-scale survey are selected from each EA by systematic sampling.[50] The interviews were conducted with a household head, his/her spouse, or any other adult household member if the head and spouse were both absent.



Note: PSU – primary sampling unit, EA – enumeration area

Source: National Institute of Statistics [50]

Figure 6: Sampling procedure of Cambodia Socio-economic Survey

For this study, I used pooled data of 38,472 households covered in the CSES 2010–2017: 3,592 in 2010 and 2011, 3,840 in 2012 and 2013, 12,090 in 2014, 3,839 in 2015 and 2016, and 3,840 in 2017.[50, 62-68] The large data pool increased precision and power.

Table 3 presents descriptive statistics and socioeconomic characteristics of the survey respondents. From 2010 to 2017, the household size was constant at 4.4 to 4.6, household head age increased by three years, and the average annual household consumption also increased by 74% in the eight years. An overall increase in the price level in the study period was deflated [72], according to the consumer price index with 2010 as the base year.[73]

Table 3: Descriptive statistics and socioeconomic characteristics of the survey respondents

	No of HHs	No of Indi-viduals	HH size[1]	HH head age[1]	F-headed HHs[2]	Per-capita HH annual consumption[3]	CPI[4]
2010	3592	16510	4.6 (1.9)	46.2 (13.9)	22 (21-24)	678 (666)	100.000
2011	3592	16327	4.5 (1.9)	46.8 (14.0)	23 (21-25)	707 (628)	105.479
2012	3840	17644	4.6 (1.8)	47.3 (13.8)	22 (20-23)	814 (710)	108.572
2013	3840	17225	4.5 (1.8)	47.5 (13.6)	21 (20-23)	895 (695)	111.767
2014	12090	53968	4.5 (1.8)	47.8 (13.8)	22 (21-23)	894 (718)	116.076
2015	3839	17301	4.5 (1.7)	49.2 (13.7)	24 (22-25)	1036 (843)	117.493
2016	3839	16985	4.4 (1.8)	49.3 (13.9)	23 (21-24)	1178 (910)	121.071
2017	3840	16090	4.4 (1.7)	49.2 (13.7)	23 (21-25)	1179 (912)	124.572
Total	38472	172050					

Notes: No – number, HH – household, F-headed – female-headed, CPI – consumer price index, 1. Mean (standard deviation), 2. Percentage (95% confidence interval), 3. Median (interquartile range) in current US dollars (1 USD = 4061.15 riels as of 23 June, 2020) [74] , 4. The 2010-base consumer price index [73] was used to adjust all household consumption data for inflation. Per-capita household annual consumption in this table are crude data.
Source: Cambodia Socio-Economic Survey 2010-2017 [50, 62-68]

2.2.2. Composition of household consumption

Equitable health insurance contribution should be determined based on one's CTP, which is not simply defined as a current income function. It should be more precisely defined

as a non-subsistence effective income.[33] Effective income is further defined as the income that households would behave as if they have when making consumption decisions.[33] Households tend to smooth consumption over time by saving and borrowing [75], taking into account expected variations in income over the year, their assets, and future earning potentials.[33] Additionally, the policy paper suggested that consumption-based measure is more relevant in a lower-income setting where many households are borrowers, rather than savers.[76] The household survey of 2015 reported that 93% of the survey respondents in Cambodia had no saving, while 41% had experience of debt or asset sale for emergency.[77] Moreover, household consumption should be defined over a period of one year because it is a natural unit of time for most households to encompass many predictable fluctuations in income and expenditure.[33] Therefore, annual household consumption was used as the basis to estimate the household CTP in Cambodia. The household consumption in each year was transformed into the value of 2010 based on the consumer price index [73] to adjust for inflation in the eight years.

Table 4 shows the consumption items that are included in the household consumption aggregates.[70] The CSES household questionnaire is designed to collect consumption

data on purchase in cash, consumption of own production, and consumption of items received in kind, by asking approximately 130 questions. The data were aggregated following the Guideline for Constructing Consumption Aggregates for Welfare Analysis [78], the most widely referenced guideline of household consumption aggregates, albeit excluding consumer durables due to insufficient information in the CSES data.[70]

Table 4: Composition of household consumption aggregates

1.	**Food consumption**
	Cereals; fish; meats and poultry; eggs; dairy products; oil and fats; fresh vegetables; tuber; pulses and legumes; prepared and preserved vegetables; fruits; dried nuts and edible seeds; sugar, salt, and spices; tea, coffee, and cocoa; non-alcoholic beverages; alcoholic beverages; tobacco products; other food products; food taken away from home; prepared meals bought outside and eaten at home
2.	**Non-food consumption**
	Medical care; medical products; purchase of vehicles; operation of transport equipment; transportation services; communication and postal services; communication equipment; personal care; clothing footwear; furniture; domestic salaries; accommodation services within the country; recreation within the country; recreation abroad; education; personal effects; gambling; transfer to charity; regular inter-household transfers; miscellaneous expenditure
3.	**Housing consumption**
	Utility (water, sewage, garbage, electricity, gas, kerosene, firewood, charcoal, battery, and other); rent or rent value of the dwelling; maintenance and repair of the dwelling

Source: National Institute of Statistics [70], World Bank [78]

2.2.3. Household consumption predictors

Based on the previous discussions in similar studies [51-61, 79], 369 predictor variables were created with the CSES data, to explore every possibility to find the best predictors of the outcome. For example, four variables were created for possession of cars: (1) whether the household had a car at the time of interview, (2) the total number of cars that the household had at the time of interview, (3) the number of cars that the household had acquired within a year from the day of interview, and (4) the number of cars that the household had acquired more than a year earlier than the day of interview. To avoid recall bias and underreporting, most of the predictor variables were created to be proxy indicators, such as information on occupation and education, instead of direct measurement of consumption, such as food expenditure. Table 5 shows the summary of the predictors.

Table 5: Summary of the household consumption predictors

Residential Area
Province; urban/rural settings
Household members' characteristics
Sex, age, ethnicity, and educational level of household members; household size; dependent rate; total working hours
Real estate property
Number, area and use of own land; number, area, use and price value of own buildings; investment on buildings
Housing/living conditions
Size and construction materials of the dwelling; source of lightening; source of drinking water; type of toilet; utility charges; consumption of luxury food
Land use
Number, area, and use of land parcels operated
Farming activities
Harvested land area; production; type of livestock, fishery, and forestry activities
Durable goods
Possession, number, and newness of durable goods in both urban and rural settings
Work
Type of employer; employment status; occupation; type of industry
Income and liabilities
Type of income; number and amount of loans
Survey year

2.2.4. Constructing the household consumption predictive models

The data were divided into a training set and a test set using an 8:2 ratio randomly. Subsequently, the analyses were conducted in two steps.

In the first step, using the training set, linear regression models that related a set of

predictor variables (X) to observed household consumption (y), the value reported in the CSES, were constructed, as follows:

$$y_t = \sum_k \beta_k X_{k,t} + e_t$$

where β_k is a coefficient parameter to be estimated using ordinary least squares (OLS) method and e_t is the error term, which is assumed to follow the normal distribution. With the new information on the predictor variables in time $t + 1$, the corresponding household consumption ($\hat{y}_{t+1}$) can be predicted by plugging the estimated parameters into the above equation. The multiple-stage survey design was reflected on each analysis to designate variables that contain information about the survey design, such as the sampling units and weights. The sampling unit variables were re-created by grouping the variables for each survey year.

For the linear and mixed-effects models, all the variables were screened using a partial correlation coefficient with significance at 90% or higher as the cut-off point.[61] The predictor variables were manually selected using a backward-elimination technique to construct Model A (Manually-selected Linear Model). The backward-selection technique

within a stepwise regression analytical framework with a 0.1 level of significance was also used as the cut-off point for removing variables [61] to construct Model B (Stepwise Linear Model). Subsequently, Model C (Mixed-effects Linear Model) was constructed with the remainder of the stepwise selection, assuming a random effect across the same province. To avoid overfitting, Model D was constructed with elastic net regression, which was finally functioned with L1 penalty term of the regression coefficients, the least absolute shrinkage and selection operator (LASSO) regression. In addition, it was further made into adaptive LASSO by adding data-dependent weights to obtain more unbiased estimates. Ten-fold cross-validation was used to select the regularization parameter in the adaptive LASSO model.[80, 81] All the available predictor variables were used for Model D since adaptive LASSO can automatically perform the variable selection to improve the prediction performance and interpretability of the statistical model while ensuring the model parsimony. Because the models were developed only for predictive purposes, not for determination purposes, it would not be appropriate for making conclusions regarding any cause-effect relationship between household consumption and the predictors. Likewise, multicollinearity was not an issue in these analyses.[9, 61]

In the second step, the trained models were applied to the test data. With this subset,

the household consumption values were predicted, and the results were compared with the values reported by the CSES, which used the full-length questionnaires.

Finally, the prediction performance was evaluated using three measures, namely mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE). MAE was used because it evaluates the prediction performance of the model most simply by taking the absolute difference between the actual and predicted values and finds the average as follows [82]:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t|$$

RMSE squares the difference, finds the average of all the squares, and then finds the square root, as shown below. RMSE was additionally used because it is more sensitive to larger errors as it creates an exponential change in the base number by squaring the difference.[82]

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}t)^2}$$

While MAE and RMSE are useful methods to compare the prediction performance of different models for the same dataset, they do not provide information about the prediction model's relative performance. MAPE is the percentage of the error compared to the actual value according to the following equation [83], which provides more context to explain the model's average performance.

$$MAPE = \left\{ \frac{1}{n} \sum_{t=1}^{n} \frac{|y_t - \hat{y}t|}{|y_t|} \right\} \times 100$$

All analyses were conducted in Stata 16.0. A P-value <0.05 was considered statistically significant. The protocol of this study has been published in an academic journal.[84]

2.3. Results

Figure 7 shows the conceptual framework of predictor variable selection. Out of 369 predictor variables, 98 remained after removing variables with 0.1 or greater partial correlation coefficients. Subsequently, 51 predictor variables were selected for Model A, 86 for Models B and C, and 162 remained for Model D. Appendix 1 shows details of the remaining predictor variables and the coefficients in each model.

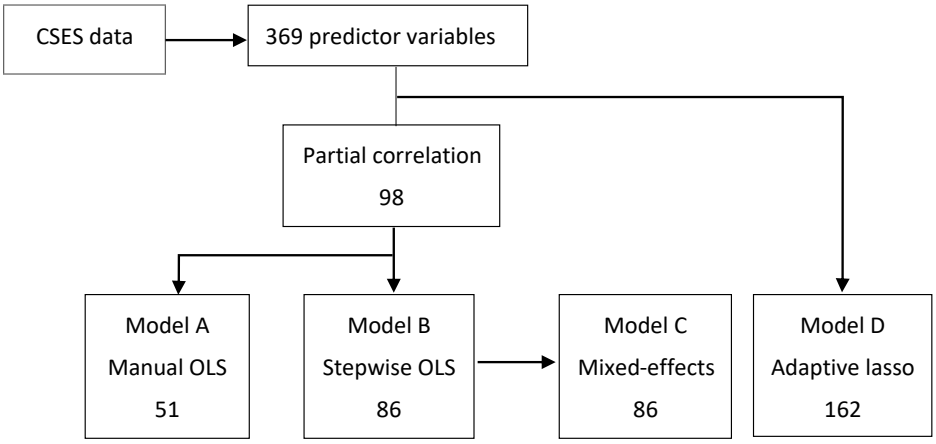


Figure 7: Conceptual framework of predictor variable selection

Figure 8 shows scatter plots of observed versus predicted household annual consumption values with the four alternative prediction models. Logarithmic transformation was performed for the outcome values. Overall, a positive linear relationship between observed and predicted household annual consumptions was found in all four models, with the data points concentrated along the regression fit lines. There was a subtle trend that the middle-class household consumption was likely to be underestimated. In contrast, the high-class household consumption was over-estimated in all four models, while the trend was less noticeable in Model D.

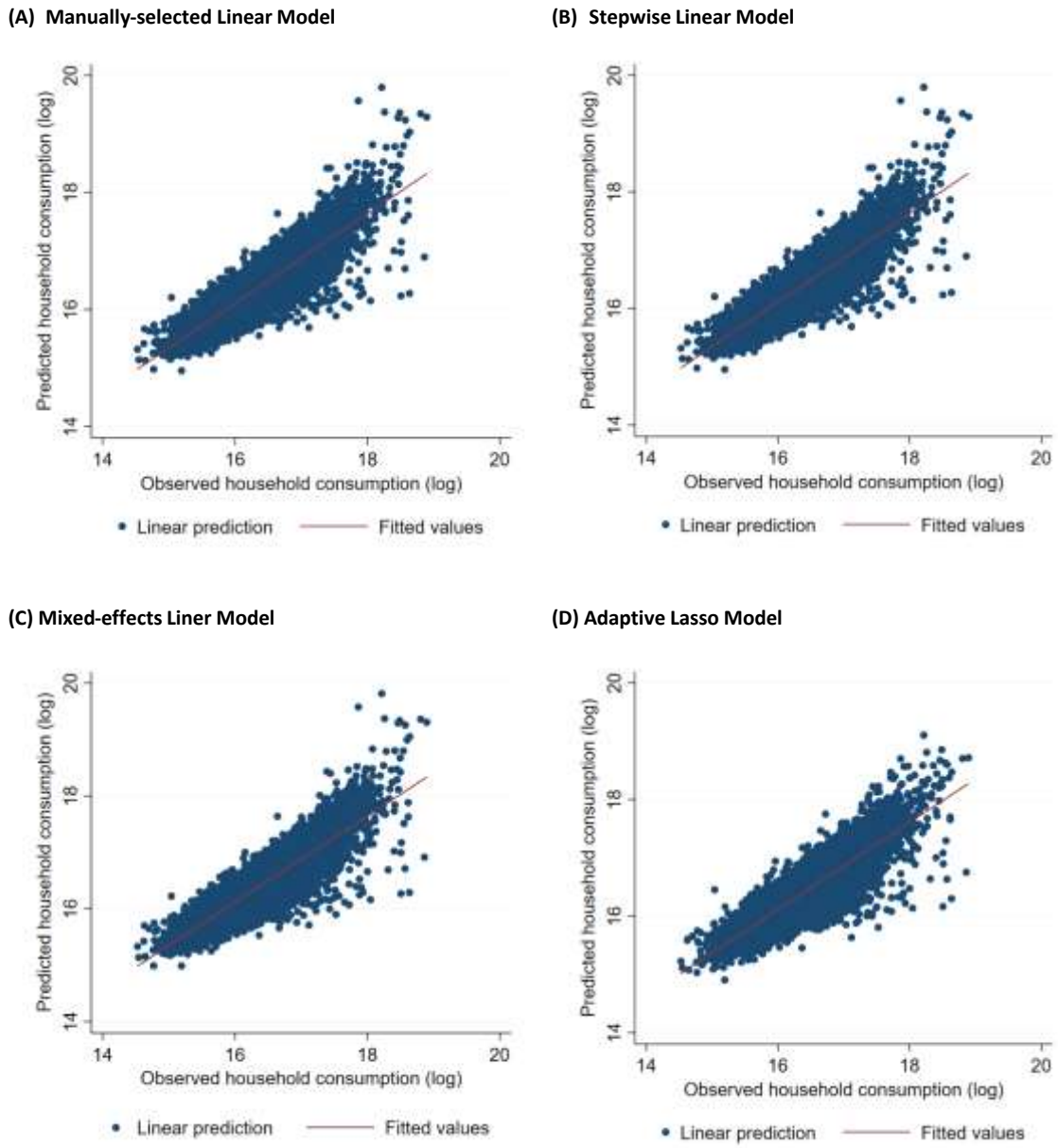


Figure 8: Observed vs. predicted household annual consumption in Cambodia in 2010-2017

Table 6 shows the MAE, RMSE, and MAPE values for the four alternative prediction models. MAE and RMSE are expressed in logarithmically converted Cambodian riels. Smaller values of all these statistical measurements are preferred. The smallest MAE of

0.227, was calculated for Model B, followed by Model C with 0.228, Model D with 0.230, and Model A with 0.242. The trend was not different for RMSE, which should have reacted more pronouncedly to larger errors, with the values of 0.301 for Model B, 0.302 for Model C, 0.305 for Model D, and 0.320 for Model A. The percentage of the predictive error, MAPE was 1.376% for Model B, 1.380% for Model C, 1.394% for Model C, and 1.469% for Model A. The rank was consistent for all three statistical measurements.

Table 6 Prediction performance of alternative predictive models (95% confidence intervals)

	Model A	Model B	Model C	Model D
	Manually-selected Linear Regression	Stepwise Linear Regression	Mixed-effects Linear Regression	Adaptive LASSO Regression
MAE (95% CI)	0.242 (0.238 – 0.247)	0.227 (0.223 – 0.231)	0.228 (0.223 – 0.232)	0.230 (0.225 – 0.234)
RMSE (95% CI)	0.320 (0.312 – 0.327)	0.301 (0.293 – 0.309)	0.302 (0.294 – 0.309)	0.305 (0.297 – 0.313)
MAPE (95% CI)	1.469 (1.441 – 1.497)	1.376 (1.349 – 1.402)	1.380 (1.354 – 1.406)	1.394 (1.367 – 1.421)

Notes: LASSO – least absolute shrinkage and selection operator, MAE – mean absolute error, RMSE – root mean squared error, MAPE – mean absolute percentage error, CI – confidence interval

3. ESTIMATION OF HYPOTHETICAL HEALTH INSURANCE CONTRIBUTIONS FOR THE CURRENTLY UNINSURED POPULATION IN CAMBODIA

3.1. Objective

This chapter presents the estimates of the hypothetical health insurance contribution options based on the predicted household consumption using the household consumption predictive model developed in the preceding chapter.

3.2. Methods

3.2.1. Data source

The following five data sources were used to estimate ten options of the hypothetical health insurance contribution for the currently uninsured population in Cambodia within the NSSF health insurance framework.

(1) The NSSF health insurance case-based and high-cost fee-for-service fee schedules in the Inter-ministerial Prakas between the Ministry of Labour and Vocational Training and the Ministry of Health (NO. 327 LV/PrK. NSSF) stipulated in August 2017 [48]

(2) Assumptions on 1) relative share of healthcare utilization by types of treatment and levels of healthcare, 2) relative share of healthcare utilization by benefit

categories (outpatient and inpatient care) and levels of healthcare, and 3) the medical transportation costs, i.e. referral on health care and corpse transportation, made by the International Labour Organization, mainly based on the data obtained from the Health Insurance Project, the pilot scheme conducted by a French non-governmental organization between 2009 and 2012 for selected garment factory workers [85]

(3) The number of outpatients and inpatients discharged from public health facilities in the Operational District Indicator Reports 2016, provided by the Cambodian MOH [86]

(4) The data of CSES 2010-2017 that were used for the regression model development in the preceding chapter

(5) The Cambodian household consumption data predicted using the regression model developed in the preceding chapter

3.2.2. Estimating the hypothetical health insurance contributions

The health insurance contribution is an amount of money that the insured must pay regularly to the health insurance fund to get a refund in case of illness or injury.[9] The health insurance contributions were estimated in three steps, following the consistent

methodology applied when the NSSF health insurance contribution rates were initially estimated for the Cambodian private-sector employees in 2015.[87] Figure 9 describes the conceptual framework of the household health insurance contribution estimation.

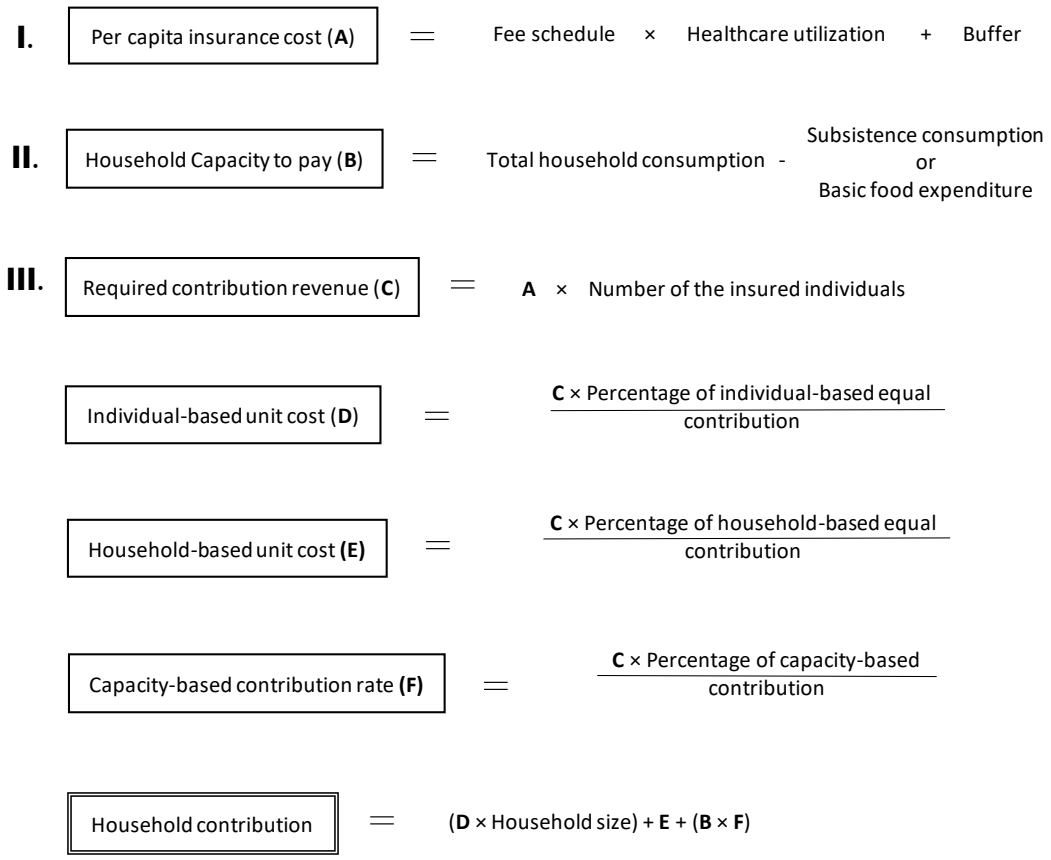


Figure 9: Conceptual framework of the household contribution estimation for the hypothetical health insurance

First, per-capita insurance cost was estimated. Second, household CTP was calculated based on the observed and predicted household consumptions using the regression

model developed in the preceding chapter. Finally, ten options of the hypothetical insurance contribution were estimated based on per-capita insurance benefit cost, administration cost, the household CTP, and additional considerations, including the need for a safety margin to account for statistical variations and to ensure the accumulation of contingency reserves.[87]

3.2.2.1. Estimating costs of per-capita insurance benefit and administration

The per-capita insurance benefit cost of general healthcare services (outpatient and inpatient care) was estimated based on the NSSF health insurance case-based fee schedule and the assumptions on healthcare utilization. Table 7 shows the current NSSF health insurance case-based fee schedule as stipulated in the revised Inter-Ministerial Prakas in 2017.[48]

Table 7: Case-based fee schedule of the National Social Security Fund health insurance (USD)

Case Description	MPA	CPA1	CPA2	CPA3	NH
Consultation	1.48	2.95	3.94	5.91	14.77
Birth Control (short-term)	2.46				
Birth Control (long-term)	7.39			98.49	147.74
Minor Surgical Activities	2.95	4.92	5.91	9.85	24.62
Surgery			36.94	49.25	147.74
Moderate Surgical Intervention				147.74	246.24
Major Surgical Intervention			98.49	246.24	369.35
Emergency/Referral	4.92	29.55	59.10	78.80	196.99
Adult General Medicine		24.62	29.55	39.40	98.49
Child General Medicine		22.65	26.59	31.52	86.18
Tuberculosis		39.40	44.32	49.25	73.87
Delivery	19.70	24.62	29.55	39.40	98.49
Gynecology		24.62	36.94	49.25	98.49
Abortion		24.62	29.55	36.94	98.49

Note: MPA – Minimum Package of Activities, CPA – Complementary Package of Activities, Attached numbers to CPA indicate the level of healthcare: greater numbers indicate more comprehensive care, NH – National Hospital that provides the most comprehensive care in Cambodia, USD – US dollars, fees are presented in the value of 2017 (1 USD = 4097.14 riels as of 16 August, 2017)[74]

Source: Cambodian Ministry of Labour and Vocational Training, Ministry of Health [48], World Bank [88]

The fee schedule is set for pre-defined cases reflecting types of treatment, such as consultation, surgery, and general medicine, and level of healthcare provided at the Cambodian public health facilities. These services are categorized into Minimum Package of Activities (MPA), Complementary Package of Activities (CPA) 1 to 3, and the National Hospitals. The numbers 1 to 3 attached to CPA indicate the level of healthcare, and the

greater numbers indicate more comprehensive healthcare. Table 8 presents the number of beds and types of healthcare services provided at MPA and CPA 1 to 3. The National Hospitals are the top-referral hospitals that provide the most comprehensive health services in Cambodia.

Table 8: Number of beds and types of healthcare services provided at MPA and CPA1-3 in Cambodia

Level of healthcare	Number of beds	Types of healthcare services provided
MPA	—	Antenatal check-up, delivery, basic pediatrics, immunization, tuberculosis, malaria, and health education
CPA1	Up to 60	Internal medicine, pediatrics, obstetrics/gynecology, emergency, radiology, sterilization, clinical testing, and pharmacy
CPA2	61-100	CPA 1 services and surgery
CPA3	More than 100	CPA 2 services, ophthalmology, otolaryngology, and blood bank

Note: MPA – Minimum Package of Activity, CPA – Complementary Package of Activity
Source: Ministry of Health of Cambodia[89, 90]

Table 9 shows the assumed relative share of healthcare utilization by types of treatment and healthcare levels, made referring to the previous assumption prepared when the NSSF health insurance contribution rates were estimated for private-sector employees.[87] The initial assumption did not grasp child as the target group as it was

made for employees. However, child was added in this study because children are part of the currently uninsured population and the fee of child general medicine is exclusively scheduled in the NSSF health insurance system. Child was defined as individuals under the age of 15 years, following the international standard recommended by the United Nations Children's Fund (UNICEF) [91] and the Cambodian Labour Law.[92]

Table 9: Assumed relative share of healthcare utilization by types of treatment and levels of healthcare for adult male, adult female and child (%)

Population group	Male Adult					Female Adult					Child				
Proportion in total population	0.32					0.34					0.34				
	MPA	CPA1	CPA2	CPA3	NH	MPA	CPA1	CPA2	CPA3	NH	MPA	CPA1	CPA2	CPA3	NH
Outpatient care															
Outpatient consultation	95.0	96.0	92.0	88.0	90.0	93.0	96.0	92.0	88.0	90.0	95.0	96.0	92.0	88.0	90.0
Minor surgical consultation	2.0	4.0	8.0	12.0	10.0	2.0	4.0	8.0	12.0	10.0	2.0	4.0	8.0	12.0	10.0
Emergency	3.0	-	-	-	-	3.0	-	-	-	-	3.0	-	-	-	-
Birth control (short-term)	-	-	-	-	-	2.0	-	-	-	-	-	-	-	-	-
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Inpatient care															
Emergency	-	2.0	2.0	2.0	2.0	-	1.0	1.0	1.0	1.0	-	2.0	2.0	2.0	2.0
Surgery	-	-	5.0	10.0	5.0	-	-	2.5	5.0	4.0	-	-	5.0	10.0	5.0
Major surgical intervention	-	-	10.0	20.0	35.0	-	-	5.0	10.0	20.0	-	-	10.0	20.0	35.0
Adult general medicine	-	97.0	82.0	52.0	37.0	-	48.5	41.0	26.0	18.5	-	-	-	-	-
Child general medicine	-	-	-	-	-	-	-	-	-	-	-	97.0	82.0	52.0	37.0
Tuberculosis	-	1.0	1.0	1.0	1.0	-	0.5	0.5	0.5	0.5	-	1.0	1.0	1.0	1.0
Moderate surgical intervention	-	-	-	15.0	20.0	-	-	-	6.5	5.0	-	-	-	15.0	20.0
Delivery	-	-	-	-	-	98.0	30.0	30.0	30.0	30.0	-	-	-	-	-
Gynecology	-	-	-	-	-	-	10.0	10.0	10.0	10.0	-	-	-	-	-
Abortion	-	-	-	-	-	-	10.0	10.0	10.0	10.0	-	-	-	-	-
Birth control (long-term)	-	-	-	-	-	2.0	-	-	1.0	1.0	-	-	-	-	-
Total	-	100	100	100	100	100	100	100	100	100	-	100	100	100	100

Note: MPA – Minimum Package of Activity, CPA – Complementary Package of Activity, NH – National Hospital

Source: International Labour Organization [87] and General Population Census of the Kingdom of Cambodia 2019 [35]

The unit cost of each treatment by levels of healthcare was obtained by multiplying the case-based fees by the relative share of healthcare utilization by types of treatment and levels of healthcare for male and female adults and children. The unit cost of each population group was further averaged by multiplying the proportion of the group in the total population and summed up.

The further assumption on the relative share of healthcare utilization by benefit categories (outpatient and inpatient care) and levels of healthcare was also used to estimate an average unit cost of outpatient and inpatient care across all levels of healthcare. Table 10 shows assumed relative share of healthcare utilization by levels of healthcare that was made based on the previous assumption prepared for private-sector employees.[87]

Table 10: Assumed relative share of healthcare utilization by levels of healthcare (%)

Service category	MPA	CPA1	CPA2	CPA3	NH	Total
Outpatient care	10.0	10.0	10.0	20.0	50.0	100.0
Inpatient care	12.5	5.0	7.5	10.0	65.0	100.0

Note: MPA – Minimum Package of Activities, CPA – Complementary Package of Activities, NH – National Hospital

Source: International Labour Organization [87]

The annual healthcare cost was obtained by multiplying the average unit costs of outpatient and inpatient care by the healthcare utilization rates obtained from the Cambodian MOH.[86] Table 11 shows the population, the total number of outpatients and inpatients discharged from the public health facilities in each province of Cambodia, and the healthcare utilization rates in 2016. The national average utilization rate of outpatient care was 0.58 times, and that of inpatient care was 0.02 times per person per year.

Table 11: Population, total number of outpatients and inpatients discharged from the public health facilities and the healthcare utilization rates by province in 2016

	Province	Population	Outpatient care (case)	Inpatient care (case)
1	Banteay Meanchey	730,596	682,892	18,689
2	Battambang	1,173,416	611,644	20,043
3	Kampong Cham	1,052,864	731,363	19,661
4	Kampong Chhnang	537,513	315,257	11,936
5	Kampong Speu	825,107	290,740	5,966
6	Kampong Thom	688,167	396,703	12,216
7	Kampot	629,383	217,407	8,897
8	Kandal	1,238,353	694,576	21,296
9	Koh Kong	130,325	71,932	2,601
10	Kratie	369,033	105,007	7,728
11	Mondul Kiri	73,702	72,263	2,994
12	Phnom Penh	1,464,856	268,804	13,640
13	Preah Vihear	208,953	218,568	11,505
14	Prey Veng	1,181,100	809,172	28,074
15	Pursat	434,003	228,633	7,170
16	Ratanak Kiri	190,479	170,720	6,302
17	Siemreap	1,018,979	945,785	16,496
18	Preah Sihanouk	203,844	54,725	N/A
19	Stung Treng	135,670	87,976	609
20	Svay Rieng	604,715	320,352	16,477
21	Takeo	975,986	688,682	23,173
22	Oddar Meanchey	242,052	152,010	4,849
23	Kep	39,296	20,767	N/A
24	Pailin	66,976	23,490	N/A
25	Tbong Khmum	789,287	587,968	22,637
	Total	15,004,655	8,767,436	282,959
	Utilization rate		0.58	0.02

Note: Utilization rate is per person per year

Source: Ministry of Health of Cambodia [86]

The annual costs of outpatient and inpatient care was combined to estimate the total cost of the general healthcare services per capita per year to be covered by the

hypothetical health insurance.

For high-cost items and medical transportation, i.e., referral on health care and corpse transportation, the per-capita annual costs, previously estimated for private-sector employees in 2015 [87] were converted to the value of 2017 to be used in this study. Table 12 presents the selected high-cost healthcare items and estimated unit costs, annual incidence rates, and annual per-capita costs [47], that were made referring to the previous estimation prepared for private-sector employees.[87] For metal osteosynthesis, that was added after the previous estimation, the average annual per capita cost of the other eight items was computed as an annual per capita cost of the service. The per-capita annual cost of high-cost items was estimated to be 2.07 US Dollars and converted to be 2.19 US Dollars in the 2017 value.

Table 12: High-cost healthcare items, estimated unit cost, annual incidence rates, and annual per-capita cost

Case description	Estimated per capita cost (USD)	Annual incidence rate	Annual per capita cost (USD)
Hemodialysis	78.80	0.00028	0.02
MRI	147.74	0.00200	0.30
CT scan	118.19	0.00400	0.47
Radiotherapy	14.77	0.00341	0.05
Cancer cell analysis	24.62	0.00100	0.02
Trepanation	689.46	0.00018	0.12
Cardiovascular surgery	3939.77	0.00020	0.79
Emergency treatment for heart disease	147.74	0.00050	0.07
Metal osteosynthesis	–	–	*0.23*
Total			2.07
Total (converted to 2017 value)			2.19

Note: MRI – magnetic resonance imaging, CT – computed tomography, USD – US dollars, costs are presented in the value of 2017 (1 USD = 4097.14 riels as of 16 August, 2017)[74]
Source: Cambodian Ministry of Labour and Vocational Training and Ministry of Health [47], International Labour Organization [87], and the World Bank [73]

The average per capita annual cost of medical transportation was estimated based on the assumption that ten percent of hospitalized patients would need to be transferred, and ten percent of the referred cases are transferred to other provinces.[87] The distance of medical transportation was assumed to be 20 km if it was within the province and 400 km if it was between provinces, as suggested in the previous assumption.[87] It was estimated to be 0.34 US Dollars in 2015 and converted to 0.36 US Dollars in the value of 2017.[73]

The total cost of healthcare per capita per year for all benefits was obtained by adding the average costs of the outpatient and inpatient care, high-cost items, and medical transportations. The per capita annual health insurance benefit cost was then obtained by adding an administration cost allocation assumed at ten percent of the total benefit cost. Ten percent is the maximum proportion allowed as administration cost in the Cambodian Social Security Law.[87]

Finally, capital buffer was added to the above subtotal, assuming that the contribution collection is imperfect. It was assumed that 95% of the contributions would be collected. Additionally, 90% was set as the density factor that accounts for the fact that on average the duration during which contributions are paid is shorter than that during which entitlement to benefits exists, due to employment turnover and other factors.[87]

3.2.2.2. Estimating household capacity to pay

The hypothetical health insurance contribution was estimated based on household CTP because a household is a unit that owns and manages the productive resources in a society [33, 93, 94]. Fairness in health financing is usually defined as equal shares of household CTP to the health system.[33]. Household-based enrollment is also more

efficient than individual-based enrollment in the extension of the insurance coverage and risk reduction of adverse selection.[16] Household CTP was estimated using the CSES 2010-2017 data that were previously used for the regression model development. Household CTP equals total household consumption minus subsistence expenditure.[95] The household subsistence spending is the minimum requirement to maintain basic life in a society.[95] In this study, the subsistence spending was defined as the average food expenditure of households whose food share in the total household expenditure was between the 45th and 55th percentile of the Cambodian households, following the WHO guideline.[95] The sampling weights were used to calculate the average food expenditure.

Household consumption was adjusted for household size and composition since a household with many adults typically consumes more than a household with fewer adults. Economies of scale were also taken into account as the household members share the consumption. Thus, household consumption was redefined in the adult equivalent scale as follows [75]:

$$AE = (A + aK)^{\theta}$$

where A is the number of adults, K is the number of children in the household, α is the cost of children, and θ reflects the degree of economies of scale.[75] In this study, 0.5 was set for α and 0.75 for θ, as suggested by previous studies given the fact that economies of scale are relatively limited in LMICs. [75, 96] In case a household basic food expenditure was lower than the subsistence expenditure, the household CTP was calculated as total household consumption minus the basic food expenditure.[95]

3.2.2.3. Estimating the hypothetical household health insurance contributions

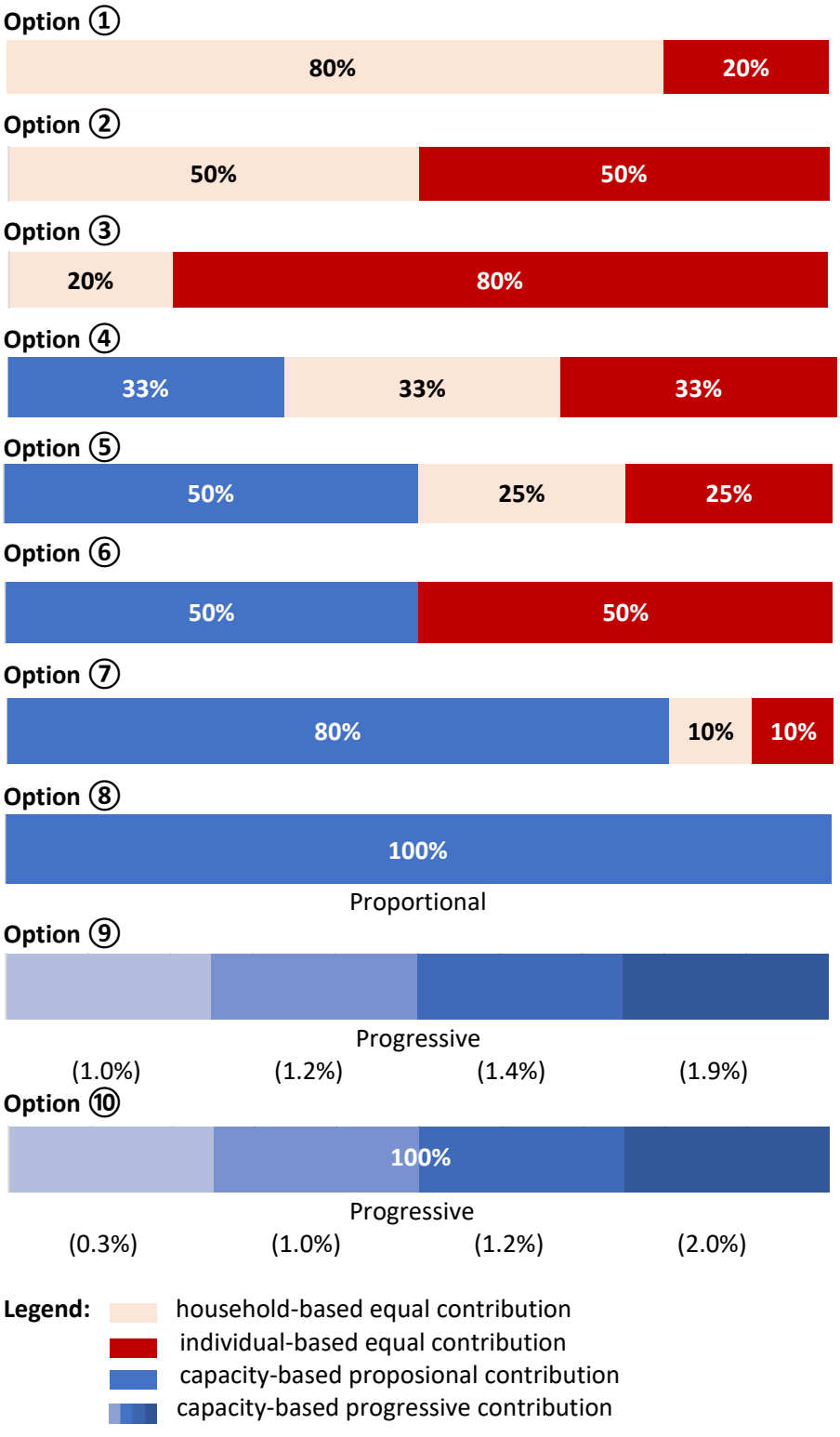
Based on the estimated health insurance cost and the household CTP, ten options of the hypothetical household health insurance contributions were calculated on both observation and prediction bases. Household health insurance contributions have been defined in various ways in the world. For example, it is defined based on household disposable income and assets, and participation as a household or an individual for the Japanese National Health Insurance (NHI) today.[97] The combination and ratio of insurance contribution fractions are decided by each insurer, considering local contexts.[97] The participatory contribution, i.e., household- and individual-based flat-rate contribution, is collected for the NHI in Japan because (1) each family member is considered to be the insured, rather than a dependent in the self-employed family, and

(2) it is considered to be unfair if a smaller household has to bear the medical expenses of a larger household.[19] Whereas, the insurance contribution is defined based on household disposable income and the number of dependents in Taiwan [98], and household disposable income and assets in Korea and Rwanda.[25, 27, 28] The contribution is collected from the informal sector in two distinct rates based on self-reported income in the Philippines [99] and at a flat rate for the Urban Resident Basic Medical Insurance and the New Rural Basic Medical Insurance in China.[26]

In this study, ten options of the hypothetical health insurance contribution were proposed, comprising the four basic insurance contribution collection methods, namely (1) household-based flat-rate contribution, (2) individual-based flat-rate contribution, (3) capacity-based proportional contribution, and (4) capacity-based progressive contribution. Asset-based contribution collection, practiced in Japan, Korea, and Rwanda, was not considered in this study because the CSES data do not contain the information that allows calculation of household asset values.

Initially, the total insurance contribution revenue, required to cover the hypothetical health insurance benefit costs for the currently uninsured population in Cambodia, was

calculated by multiplying the estimated per-capita insurance cost by the expected number of the insured. Subsequently, the unit costs of the household- and individual-based flat-rate contribution were calculated, by dividing the defined revenue fraction by the expected number of participating households or individuals. Then, the capacity-based insurance contribution rate was calculated, by dividing the revenue fraction by the total household CTPs of the participating households. For the capacity-based progressive contribution options, contribution rates were pre-defined and set higher for the richer and lower for the poorer. Figure 10 describes details of ten options of the hypothetical health insurance contribution.



Notes: Bar charts represent the required insurance contribution revenue, and the percentages in the bars indicate proportions of the revenue to be collected by the designated method. For Options ⑨ and ⑩, the percentages in parentheses indicate the CTP share of the insurance contribution paid by Q2 to Q5 households.

Figure 10: Proposed ten options of the hypothetical health insurance contribution

The bar charts represent the required health insurance revenue that needs to be filled with the insurance contributions, which does not include government subsidy or any other revenue source. The percentages in the charts indicate proportions of the revenue collected by the designated methods. For Option ⑨ and ⑩, the percentages in parentheses indicate the CTP share of the insurance contribution paid by Q2 to Q5 households.

For Options 1 to 3, household-based and individual-based flat-rate contributions were collected at 8:2, 5:5 and 2:8 ratio, respectively. The capacity-based contribution was added to the flat-rate contributions for Options 4 to 7. For Option 4, equal fractions of the revenue were collected by the three different collection methods. For Option 5, one half of the revenue was collected based on household CTP, 25% on household-based flat-rate contribution, and 25% on individual-based flat-rate contribution. For Option 6, one half of the revenue was collected based on household CTP, and the other half on individual-based flat-rate contribution. For Option 7, 80% of the revenue was collected based on household CTP, and ten percent each on household-based and individual-based flat-rate contributions. For Options 8 to 10, all revenue was collected based on household CTP. For Option 8, all contribution was collected on household CTP

proportionally. For Options 9, contribution was collected progressively at 1.0% from Q2, 1.2% from Q3, 1.4% from Q4 and 1.9% from Q5, and 0.3% from Q2, 1.0% from Q3, 1.2% from Q4 and 2.0% from Q5 for Option 10.

The beneficiaries of the HEF, 20% of the poorest households, were excluded for this estimation as they are exempted from healthcare payment. Formal-sector employees were also excluded for the estimation since their contribution is already paid by their employers. Given the fact that the NSSF requires only those enterprises that employ eight or more workers for the initial period [100], and the turnover rate is generally high in Cambodia, it was not realistic to estimate the number of insured household members based on the survey data. Instead, household size was indiscriminately reduced by ten percent, the proportion of the formerly employed population in Cambodia, to make it closer to the overall number of uninsured members in each household.

Moreover, the hypothetical insurance contributions were also estimated for the scenario that the government makes subsidization. In this scenario, the hypothetical health insurance contribution was subsidized when it exceeded three percent of household CTP. The cut-off point was set at three percent because 80% of the National Health Insurance

enrollees in Tokyo were contributing as much as three percent of their disposable income in the 1960s, the first decade after the compulsory insurance was introduced in Japan.[101] Model B, developed and evaluated as the most predictable model in the preceding chapter, was used for the household consumption prediction.

Appropriateness of the hypothetical health insurance contributions was further assessed regarding the amount and share in household CTP. The amount of household insurance contributions on a prediction basis was compared with those on an observation basis. The share of the insurance contribution in household CTP was calculated by dividing the household insurance contribution on a prediction basis by household CTP on an observation basis. Median, interquartile range (IQR), minimum, and maximum share of the insurance contribution in household CTP were presented by quintile for each option with and without subsidy to assess burden of a household.

3.3. Results

3.3.1. The hypothetical health insurance cost per capita

Table 13 provides the estimated unit cost of outpatient and inpatient care. The estimated unit costs of outpatient and inpatient care were calculated to be 10.03 US Dollars and 147.55 US Dollars, respectively.[88]

Table 13: Unit cost of outpatient and inpatient care (USD)

	MPA	CPA1	CPA2	CPA3	NH	Estimated unit cost
Outpatient care	0.16	0.30	0.41	1.28	7.88	10.03
Inpatient care	2.43	1.21	2.67	8.75	132.49	147.55

Note: MPA – Minimum Package of Activity, CPA – Complementary Package of Activity, NH – National Hospital, USD – US dollars, unit costs are presented in the value of 2017 (1 USD = 4097.14 riels as of 16 August, 2017) [88]

Table 14 shows the estimated hypothetical health insurance cost per capita per year. The per-capita annual cost of general healthcare services (outpatient and inpatient care) was estimated to be 8.77 US Dollars. By adding the cost of fee-for-service (high-cost) care and medical transportation, the total benefit cost per capita per year came to be 11.32 US Dollars. Finally, considering ten percent of administration cost [87], 95% as the completed contribution collection rate, and 90% as density factor, the total health insurance cost per capita per year was estimated to be 14.57 US Dollars.

Table 14: Estimated hypothetical health insurance cost per capita

	Unit cost (USD)	Utilization rates (times)	Annual cost per capita (USD)
Outpatient care	10.03	0.58	5.82
Inpatient care	147.55	0.02	2.95
Subtotal			8.77
Transport cost			0.36
High-cost care			2.19
Total benefit cost			11.32
Administration cost		10%	1.13
Expected collection rate		95%	0.66
Density factor		90%	1.46
Total insurance cost			14.57

Note: USD – US dollars, costs are presented in the value of 2017 (1 USD = 4097.14 riels as of 16 August, 2017) [88]

3.3.2. Hypothetical household health insurance contribution options

Table 15 shows the median and IQR of hypothetical health insurance contribution options per capita per year by quintiles estimated on observed and predicted household CTP with and without subsidy. The households were ranked by per capita household CTP and grouped into quintiles on an observation basis. Quintile 1 (Q1) is the poorest, and Quintile 5 (Q5) is the richest 20% of households in Cambodia.

Table 15: Median (IQR) of hypothetical household annual health insurance contribution options by quintiles on observation and prediction bases with and without subsidy (USD)

		①		②		③		④		⑤		⑥		⑦		⑧		⑨		⑩	
Q1	Observation basis	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
	Prediction basis	0.00	(56.09)	0.00	(54.72)	0.00	(53.36)	0.00	(41.44)	0.00	(35.65)	0.00	(34.06)	0.00	(25.27)	0.00	(17.90)	0.00	(11.27)	0.00	(3.50)
	with subsidy	0.00	(32.95)	0.00	(32.95)	0.00	(32.38)	0.00	(32.94)	0.00	(32.09)	0.00	(30.22)	0.00	(25.27)	0.00	(17.90)	0.00	(11.27)	0.00	(3.50)
Q2	Observation basis	56.96	(7.86)	55.27	(19.67)	53.58	(31.47)	44.94	(14.13)	39.46	(11.77)	38.67	(21.36)	27.44	(9.17)	19.91	(8.10)	13.55	(5.51)	4.06	(1.65)
	Prediction basis	56.09	(10.49)	54.72	(26.23)	53.36	(41.96)	42.29	(23.17)	37.17	(21.06)	35.30	(34.79)	26.96	(19.48)	20.03	(19.13)	13.14	(14.72)	4.25	(15.08)
	with subsidy	36.88	(34.73)	36.82	(32.92)	35.46	(32.92)	36.05	(29.70)	34.59	(26.73)	32.20	(32.50)	26.96	(19.48)	20.03	(19.13)	13.14	(14.72)	4.25	(15.08)
Q3	Observation basis	56.96	(7.86)	55.27	(19.67)	53.58	(31.47)	47.89	(13.30)	43.49	(12.34)	43.01	(20.06)	34.70	(11.64)	29.03	(11.95)	23.70	(9.76)	19.75	(8.13)
	Prediction basis	56.09	(5.25)	54.72	(13.11)	53.36	(20.98)	46.78	(16.45)	42.47	(16.45)	41.94	(24.62)	33.87	(19.32)	27.83	(21.59)	20.50	(20.42)	17.01	(23.87)
	with subsidy	50.84	(24.56)	48.17	(26.74)	44.68	(28.51)	45.03	(20.59)	41.74	(18.59)	40.11	(24.51)	33.87	(19.32)	27.83	(21.59)	20.50	(20.42)	17.01	(23.87)
Q4	Observation basis	56.96	(5.25)	55.27	(13.11)	53.58	(20.98)	52.66	(15.21)	50.53	(15.17)	50.02	(22.80)	46.85	(17.55)	44.36	(19.17)	42.25	(18.26)	36.22	(15.65)
	Prediction basis	56.09	(5.25)	54.72	(13.11)	53.36	(20.98)	51.27	(17.24)	48.50	(19.54)	48.55	(25.77)	44.08	(26.09)	40.83	(30.47)	35.10	(37.68)	30.15	(42.79)
	with subsidy	56.09	(8.43)	54.72	(15.70)	53.36	(20.98)	50.60	(18.33)	48.42	(19.86)	47.72	(25.80)	44.08	(26.09)	40.83	(30.47)	35.10	(37.68)	30.15	(42.79)
Q5	Observation basis	56.96	(5.25)	55.27	(13.11)	53.58	(20.98)	69.30	(28.26)	74.36	(36.18)	74.50	(41.80)	84.67	(51.92)	91.13	(62.29)	101.17	(69.15)	112.78	(77.09)
	Prediction basis	56.09	(5.25)	54.72	(13.11)	53.36	(20.98)	63.24	(30.68)	66.58	(39.55)	66.12	(45.42)	72.90	(58.61)	77.23	(71.88)	87.87	(88.97)	95.61	(99.95)
	with subsidy	56.09	(5.25)	54.72	(13.11)	53.36	(20.98)	63.00	(30.72)	66.58	(39.55)	65.84	(45.47)	72.90	(58.61)	77.23	(71.88)	87.87	(88.97)	95.61	(99.95)
Total subsidy required			42,347.91		39,806.91		40,272.75		14,318.96		5,677.46		6,724.59		3.36		0		0		0

Notes: IQR – interquartile range; USD – US dollars, amount of contribution and subsidy are presented in the value of 2017 (1 USD = 4097.14 riels as of 16 August, 2017); Q – quintile, Q1 is the poorest households and Q5 is the richest households, ① to ⑩ represent the option numbers, the following indicate fractions of contribution revenues collected by 1) household-based flat-rate contribution (HFC), individual-based flat-rate contribution (IFC), or contribution based on household capacity to pay (CTP). ① – 80% HFC and 20% IFC, ② – 50% HFC and 50% IFC, ③ – 20% HFC and 80% IFC, ④ – 34% CTP, 33% HFC and 33% IFC, ⑤ – 50% CTP, 25% HFC and 25% IFC, ⑥ – 50% CTP and 50% IFC, ⑦ - 80% CTP, 10% HFC and 10% IFC, ⑧ – 100% CTP proportionally, ⑨ – 100% CTP progressively at 1.0% for Q2, 1.2% for Q3, 1.4% for Q4 and 1.9% for Q5, ⑩ – 100% CTP progressively at 0.3% for Q2, 1.0% for Q3, 1.2% for Q4 and 2.0% for Q5.

Overall, the median household contributions were slightly lower on a prediction basis than that on an observation basis, except Q2 of Options ⑧ and ⑩, that were slightly higher on a prediction basis, and Q4 and Q5 of Options ① to ③ that were consistent. Meanwhile, the IQRs were slightly higher on a prediction basis than that on an observation basis, except Q3 of Options ① and ②, that were slightly lower on a prediction basis, and Q4 and Q5 of Options ① to ③ that were consistent.

Although households in Q1 were excluded from the hypothetical insurance contribution estimates on an observation basis, some households appeared in Q1 when the insurance contributions were estimated on a prediction basis, due to the prediction errors. Consequently, the hypothetical insurance contribution of the household at the 75th percentile in Q1 became IQRs of Q1 for all options.

The medians of the hypothetical household insurance contributions from Q2 to Q5 without subsidy were flat for Options ① to ③ on both observation and prediction bases: 56.96, 55.27, and 53.58 US Dollars on an observation basis, and 56.09, 54.72, and 53.36 US Dollars on a prediction basis, respectively. The government subsidy reduced the amounts of insurance contribution imposed on the poorer households, namely Q1

to Q3. For the three options, the subsidy reduced the median household annual contribution of Q2 from 56.09 to 36.88 US Dollars, 54.72 to 36.82 US Dollars, and 53.36 to 35.46 US Dollars, and that of Q3 from 56.09 to 50.84 US Dollars, 54.72 to 48.17 US Dollars, and 53.36 to 44.68 US Dollars, respectively. The insurance contributions at the 75^{th} percentile in Q1 were also reduced from 56.09 to 32.95 US Dollars, 54.72 to 32.95 US Dollars, and 53.36 to 32.38 US Dollars, respectively. The total subsidy required annually for Options ① to ③ was estimated to be 42,347.91, 39,806.91, and 40,272.75 US Dollars, respectively. The household contribution increased as fraction of the household-based flat-rate contribution revenue increased. The total subsidy required for the hypothetical health insurance was the lowest when the household- and individual-based flat-rate contribution ratio was at 5:5.

From Option ④ onward that included a revenue fraction of capacity-based contribution, the median household insurance contribution increased as household economic status progressed on both observation and prediction bases. For Options ④ to ⑦, which contained both flat-rate contribution and capacity-based contribution fractions, the difference in insurance contributions between the rich and the poor widened as the capacity-based revenue fraction increased. Simultaneously, the total government

subsidy declined as the capacity-based revenue fraction increased. While Option ⑤ and ⑥ contained 50% of flat-rate contributions, the difference in amount of insurance contributions between the rich and the poor was wider for Option ⑥, in which flat-rate contribution was totally individual-based, than Option ⑤, which contained 25% household-based contribution. The required government subsidy was higher for Option ⑥ than Option ⑤. The government subsidy somewhat reduced amounts of household contributions in all quintiles, except Q5 in Option ⑤. There was little effect of the government subsidy for Option ⑦, in which subsidy was estimated to be minimal: 3.36 US Dollars per year.

The difference in insurance contributions between the rich and the poor was wider for Options ⑧ to ⑩, compared to that for Options ① to ⑦. The difference was even wider for options that collected capacity-based contributions progressively, and it was the widest for Option ⑩. No subsidy was required for Options ⑧ to ⑩.

3.3.3. Health insurance contribution as a share in household capacity to pay

Table 16 shows the median, IQR, minimum and maximum shares of insurance contributions in household CTP by quintiles. The contributions were estimated on a prediction basis for the ten options with and without government subsidy.

Table 16: Share of hypothetical insurance contribution in household capacity to pay by quintiles estimated on a prediction basis for the ten options with and without government subsidy (%)

Options			①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩
Q1	Without subsidy	Median (IQR)	0.00 (6.21)	0.00 (6.15)	0.00 (5.94)	0.00 (4.82)	0.00 (4.21)	0.00 (4.02)	0.00 (3.04)	0.00 (2.12)	0.00 (1.30)	0.00 (0.39)
		Min/Max	0.00 / 22.26	0.00 / 20.05	0.00 / 17.85	0.00 / 15.14	0.00 / 12.83	0.00 / 11.00	0.00 / 8.50	0.00 / 6.29	0.00 / 5.41	0.00 / 4.63
	With subsidy	Median (IQR)	0.00 (3.90)	0.00 (3.90)	0.00 (3.90)	0.00 (3.90)	0.00 (3.78)	0.00 (3.59)	0.00 (3.03)	0.00 (2.12)	0.00 (1.30)	0.00 (0.39)
		Min/Max	0.00 / 11.58	0.00 / 11.54	0.00 / 11.25	0.00 / 10.68	0.00 / 10.34	0.00 / 10.34	0.00 / 8.50	0.00 / 6.29	0.00 / 5.41	0.00 / 4.63
Q2	Without subsidy	Median (IQR)	3.88 (2.11)	4.00 (1.53)	3.90 (1.29)	3.25 (1.21)	2.86 (1.12)	2.77 (1.11)	2.10 (1.04)	1.54 (1.09)	0.96 (0.93)	0.29 (1.10)
		Min/Max	0.00 / 11.29	0.00 / 8.21	0.00 / 6.45	0.00 / 6.69	0.00 / 5.98	0.00 / 5.13	0.00 / 5.92	0.00 / 6.40	0.00 / 7.33	0.00 / 7.98
	With subsidy	Median (IQR)	2.80 (1.72)	2.84 (1.77)	2.84 (1.75)	2.80 (1.53)	2.67 (1.32)	2.58 (1.27)	2.10 (1.04)	1.54 (1.09)	0.96 (0.93)	0.29 (1.10)
		Min/Max	0.00 / 7.00	0.00 / 6.92	0.00 / 5.61	0.00 / 6.69	0.00 / 5.98	0.00 / 5.13	0.00 / 5.92	0.00 / 6.40	0.00 / 7.33	0.00 / 7.98
Q3	Without subsidy	Median (IQR)	2.81 (1.07)	2.85 (0.76)	2.81 (0.69)	2.43 (0.63)	2.22 (0.64)	2.17 (0.64)	1.77 (0.79)	1.47 (0.94)	1.08 (1.05)	0.90 (1.26)
		Min/Max	0.00 / 7.86	0.00 / 5.72	0.00 / 4.21	0.00 / 5.18	0.00 / 4.92	0.00 / 3.87	0.00 / 4.45	0.00 / 4.16	0.00 / 4.77	0.00 / 5.19
	With subsidy	Median (IQR)	2.38 (1.02)	2.46 (0.96)	2.46 (0.91)	2.32 (0.78)	2.18 (0.74)	2.12 (0.72)	1.77 (0.79)	1.47 (0.94)	1.08 (1.05)	0.90 (1.26)
		Min/Max	0.00 / 7.60	0.00 / 5.72	0.00 / 4.06	0.00 / 5.18	0.00 / 4.92	0.00 / 3.87	0.00 / 4.45	0.00 / 4.16	0.00 / 4.77	0.00 / 5.19
Q4	Without subsidy	Median (IQR)	1.87 (0.70)	1.85 (0.55)	1.80 (0.51)	1.73 (0.46)	1.66 (0.50)	1.62 (0.50)	1.52 (0.67)	1.42 (0.81)	1.22 (1.14)	1.04 (1.44)
		Min/Max	0.00 / 5.30	0.00 / 3.86	0.00 / 3.05	0.00 / 3.68	0.00 / 4.35	0.00 / 4.53	0.00 / 5.60	0.00 / 6.44	0.00 / 7.38	0.00 / 8.03
	With subsidy	Median (IQR)	1.74 (0.65)	1.75 (0.54)	1.72 (0.49)	1.71 (0.49)	1.66 (0.51)	1.61 (0.52)	1.52 (0.67)	1.42 (0.81)	1.22 (1.14)	1.04 (1.44)
		Min/Max	0.00 / 4.91	0.00 / 3.66	0.00 / 3.05	0.00 / 3.68	0.00 / 4.35	0.00 / 4.53	0.00 / 5.60	0.00 / 6.44	0.00 / 7.38	0.00 / 8.03
Q5	Without subsidy	Median (IQR)	0.93 (0.54)	0.92 (0.52)	0.90 (0.51)	1.06 (0.44)	1.12 (0.46)	1.09 (0.47)	1.21 (0.60)	1.29 (0.70)	1.48 (0.95)	1.61 (1.13)
		Min/Max	0.00 / 3.13	0.00 / 2.28	0.00 / 1.70	0.00 / 4.09	0.00 / 5.74	0.00 / 5.83	0.00 / 8.84	0.00 / 10.90	0.00 / 12.49	0.00 / 13.59
	With subsidy	Median (IQR)	0.91 (0.53)	0.90 (0.51)	0.88 (0.50)	1.05 (0.45)	1.12 (0.46)	1.09 (0.47)	1.21 (0.60)	1.29 (0.70)	1.48 (0.95)	1.61 (1.13)
		Min/Max	0.00 / 3.01	0.00 / 2.28	0.00 / 1.70	0.00 / 4.09	0.00 / 5.74	0.00 / 5.83	0.00 / 8.84	0.00 / 10.90	0.00 / 12.49	0.00 / 13.59

Notes: Q – quintile, Q1 is the poorest households and Q5 is the richest households; IQR – interquartile range; Min – minimum, Max – maximum, ① to ⑩ represent the option numbers, the following indicate fractions of contribution revenues collected by 1) household-based flat-rate contribution (HFC), individual-based flat-rate contribution (IFC), or contribution based on household capacity to pay (CTP). ① – 80% HFC and 20% IFC, ② – 50% HFC and 50% IFC, ③ – 20% HFC and 80% IFC, ④ – 34% CTP, 33% HFC and 33% IFC, ⑤ – 50% CTP, 25% HFC and 25% IFC, ⑥ – 50% CTP and 50% IFC, ⑦ - 80% CTP, 10% HFC and 10% IFC, ⑧ – 100% CTP proportionally, ⑨ – 100% CTP progressively at 1.0% for Q2, 1.2% for Q3, 1.4% for Q4 and 1.9% for Q5, ⑩ – 100% CTP progressively at 0.3% for Q2, 1.0% for Q3, 1.2% for Q4 and 2.0% for Q5.

All the median shares of insurance contributions in household CTP were under three percent with the exception of Q2 in Options ① to ④, which also came to be under three percent with the government subsidy.

For Options ① to ③ that collected all the required contribution revenue through the flat-rate contribution collection methods, the CTP share of insurance contribution was the highest for Q2, and it gradually declined towards Q5. The CTP share increased as the size of the household-based fraction increased. For Options ④ to ⑦ that included a revenue fraction collected based on household CTP, the CTP share also declined from Q2 to Q5, albeit to a lesser degree in comparison with Options ① to ③. The difference in the CTP share between the rich and the poor was widened as the capacity-based contribution fraction size increased. Although the CTP share was set uniformly for Option ⑧, the variation was observed due to the prediction errors. The CTP share steadily increased from Q2 to Q5 for Options ⑨ and ⑩, despite the prediction errors.

For Options ① to ⑥, the maximum CTP shares of Q1 exceeded ten percent, even with the government subsidy. The same trend was observed in Q5 for Options ⑧ to ⑩, and it was the highest at 13.59% for Q5 of Option ⑩. All minimum shares of insurance

contribution in household CTP were zero. This indicates that some households in all quintiles were erroneously predicted to be in Q1 and exempted from the insurance contribution payment.

4. EVALUATION OF HYPOTHETICAL HEALTH INSURANCE CONTRIBUTION OPTIONS

4.1. Objectives

This chapter aims to evaluate equity and progressivity of the hypothetical health insurance contribution options proposed in the preceding chapter.

4.2. Methods

The equity and progressivity of the hypothetical insurance contribution options in healthcare financing on both observation and prediction bases were assessed by examining how the cumulative proportion of the insurance contribution payment changed with the cumulative proportion of the households ranked by per capita household CTP, using the Gini coefficient, concentration curves and the Kakwani indexes (KIs).[26] The KI is the difference between the concentration index for the insurance contribution payment and the Gini coefficient. The Gini coefficient is associated with per capita household CTP before the insurance contribution was paid.

The KI ranges from -2 to 1, with a positive value indicating that the insurance contribution payment yields progressive healthcare financing and a negative value indicating that the insurance contribution payment yields regressive healthcare

financing. A zero value indicates that the insurance contribution payment yields proportional healthcare financing.[75] Progressivity means that the rich households contribute a larger proportion of their CTP to the healthcare financing, and regressivity means that the poor households contribute a larger proportion of their CTP to the healthcare financing.[75]

The estimates of the Gini coefficient and concentration index were obtained, as shown below:

$$G = \frac{2}{N\mu} c_i r_i - 1 - \frac{1}{N}$$

where N is the number of observations, c_i is the insurance contribution payment or CTP of household i, μ is the mean insurance contribution payment or CTP, and r_i is the fractional rank of household i in the CTP distribution with i = 1 for the poorest and i = N for the richest.[75] The Gini coefficient ranges from 0 to 1, with 0 representing perfect equality and 1 representing perfect inequality. A higher Gini coefficient means greater inequality. If every household of a nation had the same CTP, the Gini coefficient would be zero. If only one household had all CTPs of a nation and the rest had zero CTP,

the Gini coefficient would be 1.[102]

The concentration index ranges from -1 to 1, with a negative value indicating that poor households are paying a disproportionally higher proportion of insurance contribution and a positive value indicating that the rich are making a higher proportion of insurance contribution. The value zero indicates that there is no inequality in the insurance contribution payment.[75]

The KI was computed using OLS regression method, and it was obtained as follows:

$$2\sigma^2\left(\frac{s_i}{\mu}-\frac{t_i}{\eta}\right)=a+\theta x_i+\varepsilon$$

where s_i is the insurance contribution payment of household i, μ is an estimate of its mean, t_i is CTP variable, η is an estimate of its mean, x_i is the household fractional rank according to CTP distribution and σ^2 is its variance. The OLS value of θ is an estimate of the KI.[26] In this study, the concentration index and the KI of OOP payment, the current practice of healthcare financing in Cambodia, were also calculated and compared with those of the proposed ten options of hypothetical health insurance

contribution. The data of household OOP healthcare payment were obtained from the CSES 2010-2017 household consumption data.

Additionally, dominance tests were conducted to determine whether the insurance contribution would help reduce economic inequality in Cambodia, in the sense that the poorer households would contribute a smaller proportion of their CTP to the healthcare financing system than the richer households. The tests examined whether a concentration curve of each healthcare financing scheme dominates, i.e., lying above or crossing, against the 45° line, the Lorenz curve, or another concentration curve.[103] Dominance indicates regressivity, and non-dominance indicates progressivity in the dominance test.

The Lorenz curve defines the Gini coefficient drawn according to the percentiles of the population plotted on the horizontal axis of the graph according to per capita household CTP and the cumulative household CTP of the population on the vertical axis. The concentration curves displayed the cumulative shares of healthcare financing plotted on the vertical axis of the graph that are accounted for cumulative proportions of households ranked from the poorest to the richest in the country on the horizontal

axis.[104-106]

For the dominance tests, the standard errors of the concentration curve ordinates were computed in addition to their point estimates. A multiple comparison approach to testing was adopted [107] to reject the null hypothesis of non-dominance in favor of dominance if there is at least one significant difference between a curve and 45° line, the Lorenz curve or another concentration curve in one direction, and no significant difference in the other direction across 19 evenly-spaced points from 0.05 to 0.95.[75]

4.3. Results

Table 17 presents the quintile-based distribution of per-capita household CTP, OOP payment, and the proposed ten options of hypothetical health insurance contribution with and without government subsidy. The Gini coefficient, the concentration indexes, and the KIs were also presented to describe equity and progressivity of each healthcare financing scheme.

Table 17: Distribution of household capacity to pay, and equity and progressivity of the current healthcare financing and the hypothetical health insurance options in Cambodia

	CTP	OOP	①		②		③		④	
			without subsidy	With subsidy	without subsidy	with subsidy	without subsidy	with subsidy	without subsidy	with subsidy
Q1	4.28	5.71	8.59	6.27	8.48	6.24	8.38	6.21	6.66	5.62
Q2	7.05	9.44	15.74	13.52	15.81	13.59	15.88	13.61	12.88	11.98
Q3	10.65	13.15	19.34	18.75	19.51	18.94	19.69	19.01	16.73	16.51
Q4	19.50	20.79	24.78	26.19	24.79	26.29	24.79	26.35	23.31	23.86
Q5	58.52	50.91	31.55	35.27	31.41	34.94	31.27	34.82	40.42	42.03
Gini coefficient	0.439									
Concentration Index - predicted			0.142	0.206	0.141	0.204	0.140	0.202	0.226	0.256
Concentration Index - observed		0.344	0.197	0.251	0.192	0.246	0.187	0.243	0.289	0.306
Kakwani index		-0.094	-0.296	-0.232	-0.298	-0.235	-0.299	-0.236	-0.212	-0.183
	⑤		⑥		⑦		⑧	⑨	⑩	
	Without Subsidy	with subsidy	without subsidy	with subsidy	without subsidy	with subsidy				
Q1	5.80	5.30	2.06	5.19	4.18	4.18	3.11	2.06	1.13	
Q2	11.50	11.13	5.23	11.09	8.92	8.92	7.20	5.23	3.67	
Q3	15.42	15.34	9.21	15.45	12.96	12.96	11.33	9.21	7.67	
Q4	22.62	22.86	19.44	22.88	21.32	21.32	20.46	19.44	18.76	
Q5	44.66	45.38	64.05	45.38	52.61	52.61	57.91	64.05	68.77	
Concentration Index - predicted	0.270	0.284	0.269	0.284	0.360	0.360	0.425	0.505	0.575	
Concentration Index - observed	0.337	0.341	0.333	0.339	0.434	0.434	0.504	0.568	0.653	
Kakwani index	-0.168	-0.155	-0.169	-0.154	-0.079	-0.079	-0.013	0.067	0.136	

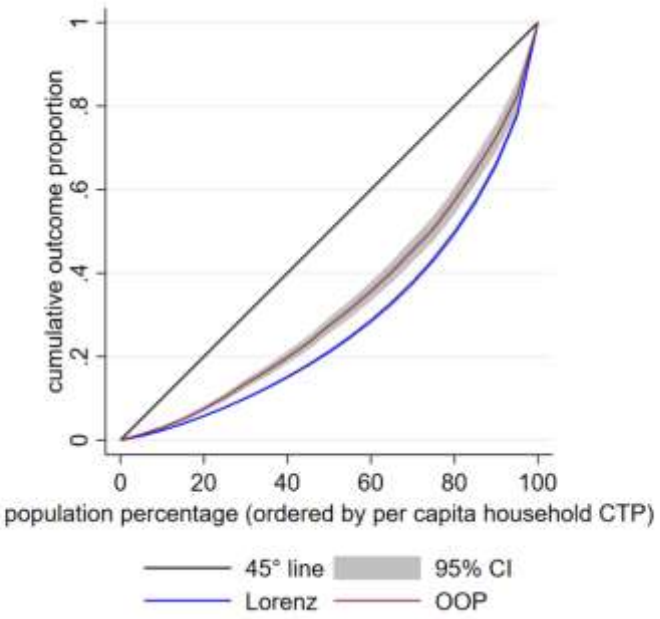
Notes: CTP – capacity to pay, OOP – out-of-pocket payment, Q – quintile: Q1 is the poorest and Q5 is the richest quintile, ① to ⑩ represent the option numbers that indicate fractions of contribution revenue collected by 1) household-based flat-rate contribution (HFC), individual-based flat-rate contribution (IFC), or contribution based on household capacity to pay (CTP). ① – 80% HFC and 20% IFC, ② – 50% HFC and 50% IFC, ③ – 20% HFC and 80% IFC, ④ – 34% CTP, 33% HFC, and 33% IFC, ⑤ – 50% CTP, 25% HFC, and 25% IFC, ⑥ – 50% CTP and 50% IFC, ⑦ - 80% CTP, 10% HFC and 10% IFC, ⑧ – 100% CTP proportionally, ⑨ – 100% CTP progressively at 1.0% for Q2, 1.2% for Q3, 1.4% for Q4 and 1.9% for Q5, ⑩ – 100% CTP progressively at 0.3% for Q2, 1.0% for Q3, 1.2% for Q4 and 2.0% for Q5.

The values of the concentration index for the OOP healthcare payment and all the insurance contribution options on both prediction and observation bases were positive. These results confirm that richer households are currently contributing a larger proportion of healthcare payments than poorer households and the trend would remain with any of the proposed options in Cambodia. Meanwhile, the concentration index values on a prediction basis were consistently lower than those on an observation basis for all options. This result indicates that the prediction errors would reduce equity in healthcare financing. For the contribution options that included government subsidy, namely Options ① to ⑦, the values of concentration indexes were higher for the options with government subsidy than those without subsidy, except Option ⑦ that had minimal government subsidy. This result indicates that the government subsidy would regain equity in healthcare financing. Furthermore, the concentration index of the OOP healthcare payment was higher than those of Options ① to ⑥, but lower than those of Options ⑦ to ⑩.

Meanwhile, the values of KIs were all negative, except Options ⑨ and ⑩. This result confirms that richer households are currently contributing a smaller proportion of healthcare payment in comparison to their CTP, and the same trend was observed

under Options ① to ⑧. In contrast, richer households were likely to contribute a larger proportion of healthcare payment in comparison to their CTP under Option ⑨ or ⑩.

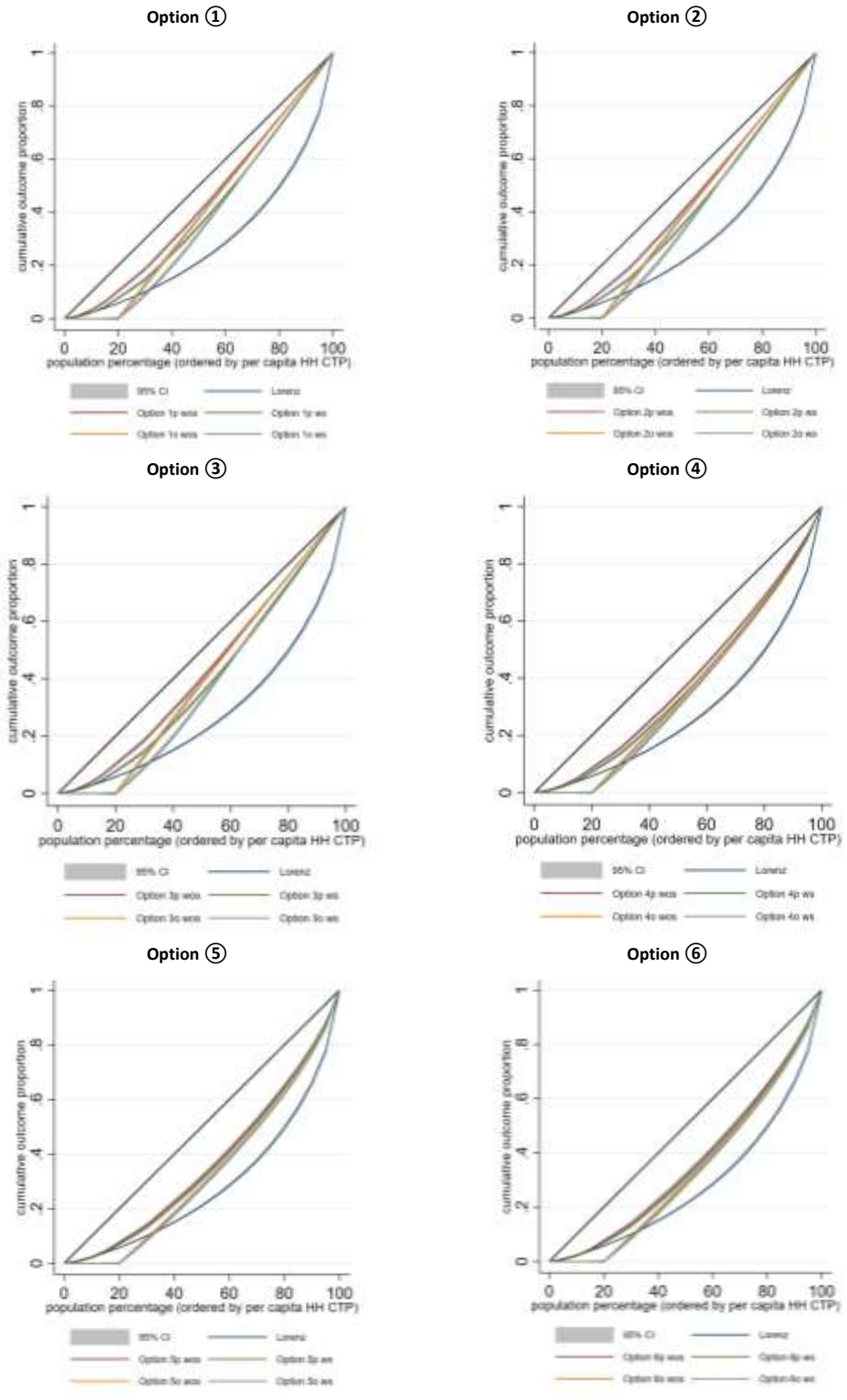
Figure 11 presents the graphical result of the dominance test for the OOP healthcare payment against the 45° line and the Lorenz curve. The Lorenz curve was drawn below the 45° line. This result represents that wealth is not equally distributed in Cambodia. On the other hand, the concentration curve, associated with the OOP healthcare payment, was dominated by the 45° line, but dominated the Lorenz curve. This result indicates that the richer are paying a larger proportion of healthcare payment, but not in comparison to their household CTP under the current healthcare system.



Notes: CTP – capacity to pay, CI – confidence interval, Lorentz – Lorenz curve, OOP – out-of-pocket healthcare payment

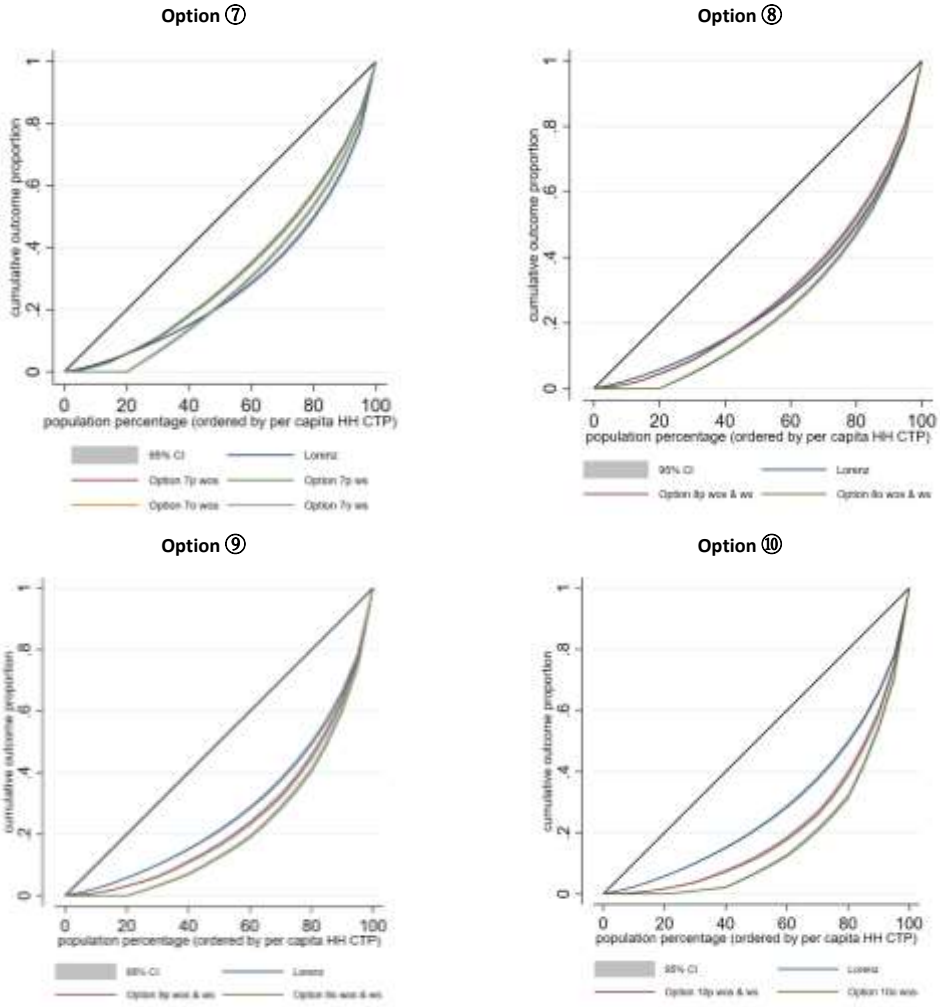
Figure 11: Dominance test of out-of-pocket health expenditure against 45° line and the Lorenz curve in Cambodia (2010-2017)

Figure 12 presents the graphical results of dominance tests conducted for each insurance contribution option on both observation and prediction bases, with and without government subsidy against the 45° line in black and the Lorenz curve in dark blue.



Notes: HH – household, CTP – capacity to pay, CI – confidence interval, Lorentz – Lorenz curve, p – prediction basis, o – observation basis, wos – without subsidy, ws – with subsidy

Figure 12: Dominance test of the hypothetical health insurance contribution options against 45° line and the Lorenz curve



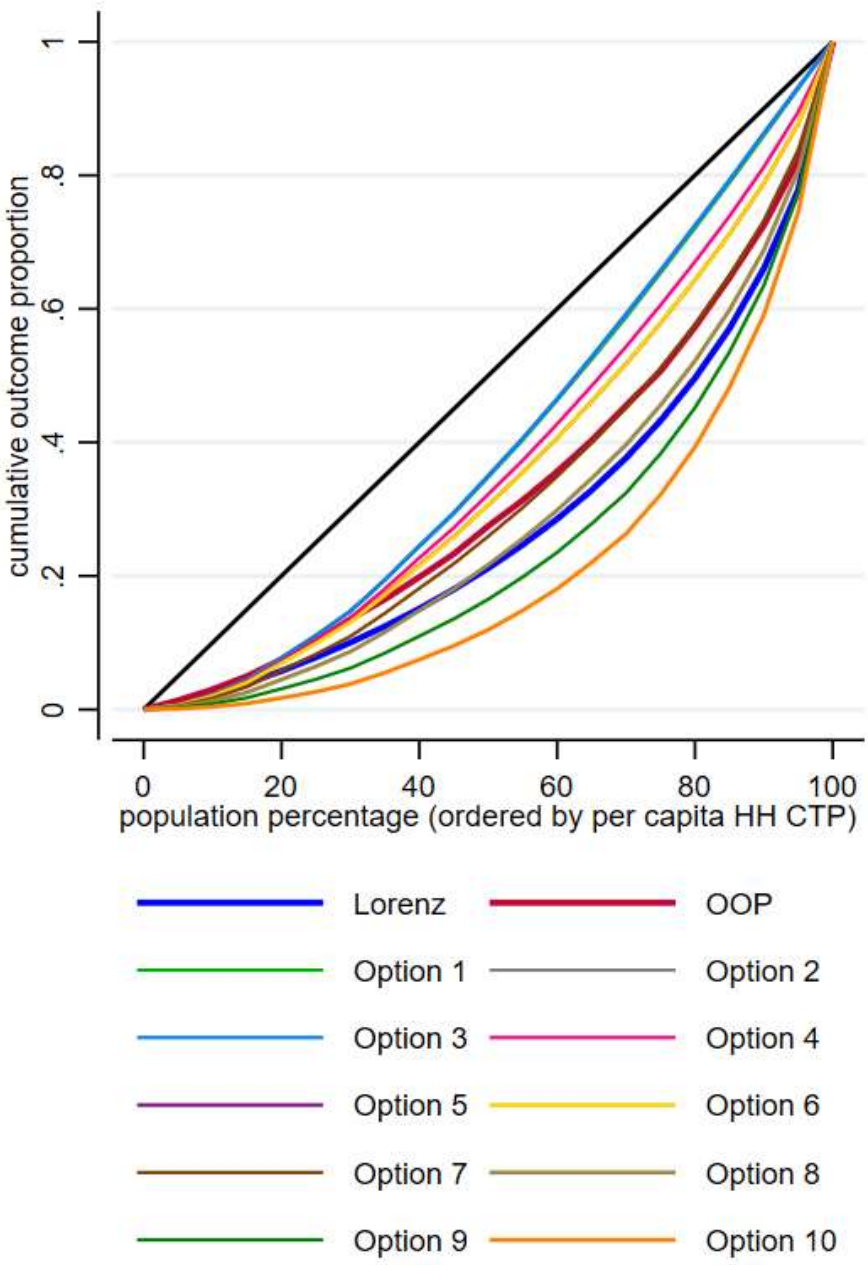
Notes: HH – household, CTP – capacity to pay, CI – confidence interval, Lorentz – Lorenz curve, p – prediction basis, o – observation basis, wos – without subsidy, ws – with subsidy

Figure 12: Dominance test of the hypothetical health insurance contribution options against 45° line and the Lorenz curve (continued)

The concentration curves associated with all options on an observation basis were flat on the bottom for the poorest 20%, while those on a prediction basis were not. This result confirms that the attempted exemption for Q1 was not applied on a prediction basis. On the other hand, all the concentration curves with government subsidy were

dominated by those without subsidy. This result indicates that the government subsidy would help reduce inequity. The concentration curves associated with Options ① to ⑦ dominated the Lorentz curve on both observation and prediction bases. This result indicates that the insurance contribution payments would not help to reduce the inequity under these options. While the concentration curve associated with Option ⑧ was dominated by the Lorentz curve on an observation basis, it crossed the Lorenz curve on a prediction basis. This result indicates that Option ⑧ was supposed to be a progressive healthcare financing scheme, but it was not on a prediction basis due to the prediction errors. The concentration curve associated with Option ⑨ was dominated by the Lorenz curve even on a prediction basis. The same trend was observed for Option ⑩ to a greater degree. These indicate that Options ⑨ and ⑩ would help to redistribute income and reduce inequity in Cambodia.

Figure 13 presents relative progressivity of the current practice of the OOP healthcare payment and the predicted insurance contribution options with government subsidy.



Notes: HH – household, CTP – capacity to pay, Lorenz – Lorenz curve, OOP – out-of-pocket healthcare payment

Figure 13: Relative progressivity of the hypothetical health insurance contribution options on a prediction basis with government subsidy and the OOP healthcare payment

The concentration curves associated with Options ① to ③, that have overlapped in the figure, dominated all other concentration curves. This result indicates that Options ① to ③ are the most regressive healthcare financing schemes among the ten. The concentration curves associated with Options ④ to ⑥ were dominated by those of Option ① to ③, but dominated the rest, including the OOP healthcare payment. This result means that Options ④ to ⑥ are less regressive than Options ① to ③, but more regressive than others, including the OOP payment. The concentration curve associated with Option ⑦ crossed that of OOP payment. This result indicates that the regressivity of Option ⑦ is comparable to that of the OOP payment. Likewise, the concentration curve associated with Option ⑧ crossed the Lorenz curve. This result indicates that Option ⑧ is nearly, but not exactly a progressive healthcare financing scheme. The concentration curves associated with Option ⑨ and ⑩ were dominated by the Lorenz curve, and Option ⑩ was dominated by Option ⑨. This result concludes that Options ⑨ and ⑩ are progressive healthcare financing schemes, and Option ⑩ is the most progressive healthcare financing option among the ten in Cambodia.

5. DISCUSSION

5.1. Discussion of the findings

This study suggested that it is possible to make a reasonable prediction of household consumption by developing a parsimonious regression model with the existing population survey data in Cambodia. While the four alternative prediction models had different functions, a simple linear model performed comparably to more sophisticated regression models. It was inferred that capacity-based contribution collection could increase relative amount of insurance revenue, and reduce burden on the poor and the government. Progressive contribution collection may reinforce the effects, and mitigate regressivity of the current healthcare financing in the country.

The four alternative prediction models developed in this study did not show significant difference in performance, particularly among Models B, C, and D, despite the difference in functions. In other words, inclusion of random effects in Model C and the regularization technique with consideration of data-dependent weights in Model D were not particularly effective in this setting. This result was consistent with the previous income prediction study, comparing 16 techniques by ten different performance measurements, which concluded that traditional linear regression performed

comparably to more sophisticated non-linear and two-stage models in income prediction.[108] This was also in line with the study that chose the stepwise linear regression model to predict the annual household consumptions in 28 Sub-Saharan African countries.[61]

Model B was suggested to best suit Cambodia, as it was the most predictive and the second most parsimonious model among the four alternatives. Parsimoniousness is an important criterion in the model selection for insurance contribution estimation, as a parsimonious model yields a simple questionnaire. Although Model A was the most parsimonious model, it was not to be the best model because its prediction performance was the worst among all. While Model B finally contained 86 predictors, the number of questions could be curtailed, as multiple predictors are attributed to one information source. Appendix 2 presents a sample questionnaire that provides the necessary information for Model B. It was estimated to take 15 to 20 minutes to answer the questions. Although the number of questions required for Model B was still larger than that for a simple scale to classify households into a few socio-economic groups [51-53, 57, 60], it was certainly more parsimonious than the CSES survey [50], and Model B allows consumption prediction on a monetary basis.

Furthermore, the questionnaire for Model B should be preferrable to that of the CSES with regards to preventing recall bias and underreporting. A respondent may not remember exact amount of consumed items, or may underreport if he or she knew that the primary purpose of the interview was to estimate the amount of insurance contribution or tax to be paid by the respondent.[111] While the CSES directly asks the respondents how much they spent for what purposes [50], the questionnaire for Model B is mostly comprised of proxy indicators, such as educational attainment and occupation of the household members. Moreover, the direct questions in the questionnaire for Model B were mostly evidence-based. That is, the respondents are asked to show their utility bills, for example, to confirm how much he or she had paid for electricity or water in the previous month.

The study presented a single model for the entire country. Although the study attempted to develop different models for urban and rural areas of the country, no significant difference was found between the two models. This was somewhat inconsistent with the previous studies that suggested urban-rural difference particularly in asset ownership. [56, 58, 110] The inconsistency might be attributed to a large proportion of the non-asset predictor variables that constitute the models developed in this study.

On the other hand, a previous study on wealth indices suggested that some indices tended to have stronger agreement with consumption expenditure in middle-income settings compared with lower-income settings.[109] This further suggested that performance of the household consumption predictive model could be better if it was developed to be income-level specific, such as by applying quantile regression. However, quantile regression was not applicable in this setting as household income level is unknown when the household consumption predictive model is applied in real settings.

The highest median annual household contribution was estimated to be 95.61 (IQR: 99.95) US Dollars for Q5 households under Option 10. It was within the range of willingness to pay suggested by the household survey conducted in Battambang and Kampong Speu provinces in Cambodia in 2017/2018.[112] The survey suggested that a great number of households in the provinces were willing to pay up to 2.50 US Dollars per person per month for health insurance, and 7.50 US Dollars per person per month might be the upper limit.[112] These were further estimated to be 121.50 US Dollars and 364.50 US Dollars per household per year, respectively. Although it was a rather small-scale survey in two provinces, it inferred that the estimated amount of contributions in this study were not unreasonable with regards to Cambodian population's willingness to

pay.

This study suggested that capacity-based contribution collection would reduce burden on poorer households, and thus increase their enrollment. This is in line with the rationale behind capacity-based contribution of the Ghanaian National Health Insurance that aimed to ensure the membership.[8] Capacity-based contribution itself may also promote the enrollment because capacity-based insurance contribution is generally believed to be fairer than flat-rate contribution.[33] Ultimately, the ensured membership may bring health financing equity, as experienced in Thailand by persistently reducing catastrophic expenditure under the Universal Coverage scheme.[113]

Although the study suggested that capacity-based contribution is a key to equitable healthcare financing, capacity-based contribution collection might be interpreted as unfair for some households as a smaller household may have to bear medical expenses of a larger household. Consequently, a fraction of flat-rate contribution might have to be sought, as experienced in Japan.[114] However, the fraction of flat-rate contribution should be kept minimum, as flat-rate contribution is likely to make the insurance system regressive.[26, 115] Asset-based contribution could be considered as alternative

capacity-based contribution, particularly in rural areas where land and agricultural assets account for a major part of the household economy.[97]

This study suggested that burden on the poor would be further diminished if contribution is collected progressively, and the progressive collection may help mitigate the increased regressivity derived from the prediction errors. On the other hand, it was found that the progressive collection, in return, could impose excessive contribution on the rich. This might violate the principle of benefit-sharing on which health insurance is based.[19] According to the principle, the insurance contribution should not exceed the value of benefits one can expect from the healthcare system.[19] When this argument arises, upper limit might have to be placed on the contribution collection, which is likely to require increased amount of government subsidy in return.

While this study suggested that capacity-based insurance contribution collection could reduce the amount of government subsidy, and subsidy might not be required when the insurance contribution was entirely collected based on household CTP. However, these estimates were based on the CSES data and the premise of compulsory participation. The amount of required government subsidy in reality depends on the number of the

insured, range of the service coverage, and contribution collection rate. Careful estimation is necessary when the limited number of citizens is enrolled, particularly in the transitional period. Meanwhile, cross-subsidization of the NSSF health insurance and the HEF might alternate the government subsidy, as suggested by a bilateral aid agency.[116]

This study suggested that healthcare financing is currently regressive in Cambodia, and the regressivity would remain with a contributory health insurance, unless contribution is largely collected on a capacity basis. This finding was in line with the previous studies in China, Tanzania and Ghana, where the contributory health insurance schemes were evaluated to be regressive [26, 115] and the payroll-based insurance scheme to be progressive.[26] Although the informal sector contribution of the Ghanaian National Health Insurance is supposed to be capacity-based, it was evaluated to be regressive because the collection system is incomplete, as discussed earlier.[8] Nonetheless, progressivity analysis should be critically examined within the context of the health financing arrangements in the country.[117]

5.2. Strengths and limitations

This was the first comprehensive study to investigate the possibility of developing a tool to assess household CTP to enroll the general population in a contributory health insurance in Cambodia to make the country move towards achieving UHC goals. A number of studies have analyzed situations of health financial protection in LMICs and assessed progress of their efforts towards achieving UHC. However, few studies have suggested a concrete solution for these countries. It is expected that the findings of this study would help other LMICs, particularly middle-income countries in Asia, where the largest concentration of the world population with catastrophic health spending resides, [5] to accelerate their efforts to move towards UHC achievement.

This study initially maximized the number of predictors by looking at the available data from different angles to explore every possibility to select the best predictors for the household consumption predictive model. Subsequently, the number of predictors was reduced to make the model parsimonious for the practical purposes. This attempt was supported by the evidence that the increased number and types of indicators in a wealth index could result in modest gains in the strength of its agreement with consumption expenditure.[109]

Nonetheless, there were some limitations in the study. First, all analyses in this study were subject to the assumption that the CSES provides nationally representative data with accuracy. For example, out-of-sample validation of the household consumption predictive model was performed by comparing the predicted values with the observed values in the CSES data. Validity of the models was dependent upon reliability as well as representativeness of the survey data. Although the CSES is the best nationally representative socio-economic survey that currently exists in Cambodia, the survey does not include the homeless, the boat population and the transient population in the sampling framework.[50] In addition, it captures less than 4,000 households annually, except for a large-scale survey that is conducted once in five years. Exclusion of the households in the survey design and the limited sample size per year might have affected the study results.[117] Caution is required for interpretation of the study results.

The household consumption aggregate, the outcome of the regression models, did not include consumer durables due to insufficient information in the CSES data.[70] Although consumption attributed to consumer durables does not comprise the major part of household consumption in Cambodia [50], absence of the item in the aggregate might have diminished validity of the estimated values of the outcome, as the World

Bank recommends to construct household consumption aggregate with food items, non-food items, housing consumption and consumer durables.[78]

This study found that there were some errors when the household consumptions were predicted using the regression model. Although, the impact of the household consumption prediction errors on the insurance contribution estimation was not significant, there is a chance that the prediction errors impose contribution on the poor, or exempt the non-poor from contribution payment. Additionally, it was also suggested that the errors could reduce equity in the capacity-based contribution. Therefore, the predicted results should be re-assessed on demand, as practiced in the process of identifying the HEF beneficiaries.[40]

The per-capita health insurance cost was estimated to be 14.57 US Dollars in this study. This was lower than the estimate made for the Cambodian private sector employees in 2015 [87], which is equivalent to 19.95 US Dollars in the value of 2017. The difference in the estimated costs was presumably derived from the difference in healthcare utilization rates. That is, the health insurance cost might increase once the health insurance is introduced as the healthcare utilization is also expected to increase.[17, 118] Moreover,

medical costs may also increase due to the use of advanced technology, and the changes in disease structure, from communicable to non-communicable diseases that usually require more expensive medication for longer period.[119] The age structure of the currently uninsured population, including the elderly and children, may reinforce the rise in medical costs, since they are more prone to ill health.[120] It should be noted that the health insurance costs are subject to the change in the fee schedules in the future.

This study has developed and validated the household consumption predictive models. However, the models have not yet been applied in the real settings. They should be tested gradually with regards to predictability of household consumption, capability of the local administration to operate the model, acceptability of the predicted results by the community members, and functionality of the model in the health insurance system. In the meantime, the study results should be carefully interpreted that these are based on the hypothetical health insurance with various assumptions.

5.3. Policy implications

This study has implied the possibility of extending the existing contributory health insurance to the general population, including the informal sector, by collecting capacity-

based contributions from them, using the household consumption predictive model in Cambodia. This finding may help Cambodia to solve the dilemma that neither a tax-based nor a contributory health insurance would work to improve the country's health protection system. A contributory scheme allows Cambodia to maintain financial disciple, and possibly improve the quality of healthcare, as well.[18, 19] Furthermore, a single financing scheme is preferable, in terms of risk pooling, and financial efficiency and sustainability.[16]

This study suggested that capacity-based contribution would increase the total contribution revenue. However, it should be noted that the suggestion was based on the premise of compulsory enrollment. Although a policy maker may intend to exclude the rich from the health insurance system to reduce the amount of subsidization for their contributions, the intrinsic support of low-risk, high-income earners should not be lost to enhance the financial sustainability of the insurance fund.[18] A study has also reported that a contributory insurance can be progressive in countries where coverage is universal and regressive in those that exclude the high-income group.[33]

Since the household consumption predictive model has not been tested in real settings,

a feasibility study is required before it is used in the contributory health insurance in Cambodia. Before the National Health Insurance was enacted in Japan, a field survey on 36 villages in 13 prefectures was conducted, and 12 health insurance unions were subsequently established nationwide on a trial basis to realize the system.[32] Since the Ministry of Finance in Japan did not approve the budget for the feasibility study, the study was conducted with the subsidy provided by the charitable organization.[32] In Cambodia, ideally, the first test should focus on confirming predictability of the model in one or two villages. If the predictability is confirmed, the subsequent test could be conducted in one or two communes to assess capability of the local administration in terms of human capacity and infrastructure. Simultaneously, it could be tested whether the community members accept the household-consumption prediction as a fair system. If the capability and acceptability are confirmed, the third test could be conducted in one or two operational districts to examine functionality of the household consumption predictive model in the health insurance system. If all these are confirmed, a feasibility study could be conducted in a few provinces.

It must be stressed that the actual expenditure of the health insurance for the general population will only be known after its introduction since behavioral aspects cannot be

easily predicted by the estimation.[87] This is particularly relevant for the phasing period, during which the insured gradually adjust their behavior and increasingly use their entitlements under the system. Therefore, it is relevant to allow for future adjustments of the contributions on the future financing requirements and the development of financial parameters, particularly insurable earnings, the amounts of contributions, benefit expenditure, and the accumulation of reserve funds.[87] It is recommended that actuarial evaluation be conducted periodically to review the experience and financial status of the insured, and reassess the adequacy of financing provisions and contribution rates.[87]

The household consumption predictive model requires households to provide personal information, such as types and amount of household property and their utility payment. Such personal information should be properly managed and protected against misuse. For example, in Japan, such personal data are protected by the Act on the Protection of Personal Information, and the National Civil Service Act, that prohibits the civil servants to use the data for purposes other than official duties.[121] The General Data Protection Regulation of the European Union (EU) has made most of the EU nations establish the Data Protection Authorities or Regulators, independent from the government, to be

guardians of data protection.[122, 123] In Thailand, the Personal Data Protection Act was also enacted in 2019.[124] Cambodia might also have to pursue such institutionalization, once the health insurance is in place.

Although the household consumption predictive model was primarily made for health insurance contribution collection, the model could be used for other purposes, such as income taxation and identification of the poor households. While the proxy means test might not be applicable for insurance contribution estimates, the reverse may be possible. When feasibility of this model is proved, it is worth testing the model for the dual purposes to make the Cambodian social security system more efficient.

6. CONCLUSION AND RECOMMENDATIONS

This study suggested that it is possible to develop a regression model using population survey data to make a reasonable prediction of household consumption in Cambodia. The predicted household consumption could further estimate capacity-based health insurance contribution for all households, including the informal sector. Moreover, if the health insurance becomes compulsory, the capacity-based contribution could increase the insurance revenue, reduce burden on the poor, and eventually reduce regressivity of the healthcare financing.

The final product of the household consumption predictive model will be an automated tool with selected predictor variables and respective regression coefficients. Although various issues remain for UHC achievement, incorporating the developed tool into the existing health financing system in Cambodia may enhance its current efforts to reduce catastrophic health expenditure and make progress towards UHC achievement.

The household consumption predictive model should be gradually tested in real settings. Subsequently, a well-designed feasibility study should be conducted to evaluate and improve the model. The feasibility study could also provide empirical data of how much

administration cost would be required in what settings. Simultaneously, rigorous discussions should be held among the stakeholders to decide what social values should be considered to make the best health financing policy in the country, and how the personal data should be protected.

Future study could explore developing non-linear and region-specific models in Cambodia, particularly when more data become available. Similar studies could be conducted in other LMICs to apply the attempts of this study with some modification as the model development and application in health financing are highly context-specific. Although this study has merely provided a tool to promote a possible health financing measure, it is expected that these efforts would help Cambodia and other LMICs to strengthen their healthcare financing, and ultimately move towards UHC goals.

References

1. World Health Organization. World Health Assembly Resolution 58.33, Sustainable health financing, universal coverage and social health insurance. Geneva: World Heath Organization, 2005.

2. United Nations. Transforming our world: the 2030 agenda for sustainable development, 2015. https://sdgs.un.org/2030agenda (accessed Mar 15, 2020).

3. World Health Organization, World Bank. Tracking universal health coverage: 2017 global monitoring report, 2017. https://apps.who.int/iris/bitstream/handle/10665/259817/9789241513555-eng.pdf?sequence=1 (accessed Nov 10, 2018).

4. World Health Organization. Primary health care on the road to universal health coverage 2019 monitoring report, 2019. https://www.who.int/healthinfo/universal_health_coverage/report/uhc_report_2019.pdf?ua=1 (accessed Mar 12, 2020).

5. World Health Organization, World Bank. Global monitoring report on financial protection in health 2019, advance copy, 2019. https://www.who.int/healthinfo/universal_health_coverage/report/fp_gmr_2019.pdf?ua=1 (accessed Mar 12, 2020).

6. Wagstaff A, Neelsen S. A comprehensive assessment of universal health coverage in 111 countries: a retrospective observational study. *Lancet Glob Health* 2020; 8: e39–49.

7. World Bank. High-performance health financing universal health coverage: driving sustainable, inclusive growth in the 21st century, 2019. https://www.mof.go.jp/english/international_policy/convention/g20/annex8_2.pdf (accessed Mar 12, 2020).

8. Amporfu E. Equity of the premium of the Ghanaian national health insurance scheme and the implications for achieving universal coverage. *Int J Equity Health* 2013; 12: 1–9.

9. Tsugawa Y. The world's most easy-to-understand textbook of health policy [Japanese]. Tokyo: Igakushoin, 2020.

10. Mano T. Introduction health policy [Japanese]. Tokyo: Chuokoron-Shinsha, Inc, 2012.

11. Ushikoshi H. Introduction to health economics [Japanese]. Tokyo: Iwanami Shoten, 2009.

12. Brook RH, Ware JE, Rogers WH, et al. The effect of coinsurance on the health of adults: results from the Rand health insurance experiment, 1984. https://www.rand.org/content/dam/rand/pubs/reports/2006/R3055.pdf (accessed May 10, 2020).

13. Light DW. Universal health care: lessons from the British experience. *Am J Public Health* 2003; 93: 25–30.

14. Bauernschuster S, Driva A, Hornung E. Bismarck's health insurance and the mortality decline. Munich: International Platform of Ludwigs Maximilians University's Center for Economic Studies and Information Institute, 2018. https://www.ifo.de/DocDL/cesifo1_wp6601.pdf (accessed July 22, 2020).

15. Physicians for a National Health Program. Health care systems - four basic models. https://www.pnhp.org/single_payer_resources/health_care_systems_four_basic_models.php. (accessed July 28, 2020).

16. Kwon S. Health care financing in Asia: key issues and challenges. *Asia Pac J Public Health* 2011; 23: 651–61.

17. Escobar M-L, Griffin CC, Shaw RP, editors. The impact of health insurance in low- and middle-income countries. Washington DC: Brookings Institution, 2010. https://www.brookings.edu/wp-content/uploads/2016/07/theimpactofhealthinsurance_fulltext.pdf (accessed Mar 3, 2019)

18. Shimazaki K. The path to universal health coverage - experiences and lessons from Japan for policy actions -. Tokyo: Japan International Cooperation Agency, 2013.

19. Shimazaki K. Health care in Japan: institutions and policies [Japanese]. Tokyo: University of Tokyo Press, 2020.

20. Mathauer I, Koch K, Zita S, et al. Revenue-raising potential for universal health coverage in Benin, Mali, Mozambique and Togo. *Bull World Health Organ* 2019; 97:620–30.

21. Reich MR, Harris J, Ikegami N, et al. Moving towards universal health coverage: lessons from 11 country studies. *Lancet* 2015; 387: 811–6.

22. International Labour Office. Women and men in the informal economy: a statistical picture. Geneva: International Labour Organization, 2018. https://www.ilo.org/wcmsp5/groups/public/---dgreports/---dcomm/documents/publication/wcms_626831.pdf (accessed July 22, 2020).

23. International Labour Office. The ILO recommendation n°204 recommendation concerning the transition from the informal to the formal economy: conference at the 140th session. Geneva: International Labour Organization, 2015. https://www.ilo.org/wcmsp5/groups/public/---ed_norm/---relconf/documents/meetingdocument/wcms_377774.pdf (accessed July 22, 2020).

24. Dayrit MM, Lagrada LP, Picazo OF, Pons MC, Villaverde MC. The Philippines health system review. New Delhi: World Health Organization Regional Office for Southeast Asia, 2018. https://apps.who.int/iris/bitstream/handle/10665/274579/9789290226734-eng.pdf (accessed July 22, 2020).

25. International Labour Office. Rwanda: progress towards universal health coverage. Geneva: International Labour Office, 2016. https://www.social-protection.org/gimi/gess/RessourcePDF.action?ressource.ressourceId=53613 (accessed July 23, 2020).

26. Chen M, Palmer AJ, Si L. Improving equity in health care financing in China during the progression towards universal health coverage. *BMC Health Serv Res* 2017; 17: 1–8.

27. Li L. Comparative analysis on the public health insurance policy in South Korea and Taiwan: industrialization, democratization, and social policy in late-comer societies—PhD thesis. Tokyo: Waseda University, 2007.

28. Ii M. Health care systems in Asia [Japanese]. Tokyo: Health Care Science Institute, 2009.

29. Kawamura H, Ishihara T, Kizuki M. The detailed description of the National Health Insurance Act [Japanese]. Tokyo: Ganshodo Shoten, 1939.

30. National Health Insurance Association. History of the National Health Insurance [Japanese]. Tokyo: Toyo Printing, 1948.

31. Kose Y, Nagase K. Remarks of the National Health Insurance Act, exposition of the Health Insurance Act [Japanese]. Tokyo: Nihontosho Center, 2006.

32. Nitta H. Insurers of the National Health Insurance: theoretic discussions from the system foundation to its municipality-based management [Japanese]. Tokyo: Shinzansha, 2009.

33. Murry CJL, Knaul F, Musgrove P, Xu K, Kawabata K. Defining and measuring fairness in financial contribution to the health system. Geneva: World Health Organization, 2000. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2053994 (accessed Dec 10, 2019).

34. World Bank. World Bank country and lending groups. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups (accessed Aug 13, 2020).

35. National Institute of Statistics. General population census of the Kingdom of Cambodia 2019 provisional population totals. Phnom Penh: National Institute of Statistic, Ministry of Planning, 2019. http://www.nis.gov.kh/nis/Census2019/Provisional%20Population%20Census%202019_English_FINAL.pdf (accessed Aug 13, 2020).

36. World Bank. GDP per capita (current US$) – Cambodia. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=KH (accessed Aug 15, 2020).

37. World Health Organization. Global health expenditure database. https://apps.who.int/nha/database/Regional_Averages/Index/en (accessed Aug 20, 2020).

38. Xu K, Soucat A, Kutzin J, et al. Global spending on health: a world in transition. Geneva: World Health Organization, 2019. https://apps.who.int/nha/database/Regional_Averages/Index/en (accessed Aug 20, 2020).

39. Ministry of Health of Cambodia. Health Equity Fund operational manual. Phnom Penh: Ministry of Health of Cambodia, 2016.

40. Identification of Poor Households Programme, Ministry of Planning, Kingdom of Cambodia. Leave no one behind: Cambodia national poverty identification system. https://www.idpoor.gov.kh/ (accessed Aug 15, 2020).

41. Kingdom of Cambodia. Notification on registration of the informal workers in the National Social Security Fund in order to access health care scheme through the Health Equity Fund system and the additional allowance for the female workers when delivery. Phnom Penh: Kingdom of Cambodia, 2017.

42. Cambodian Ministry of Labour and Vocational Training, Ministry of Economy and Finance, Ministry of Health. Inter-ministerial prakas on implementation of health care scheme through Health Equity Fund system for informal workers and provision of additional allowance for female workers when delivery. Phnom Penh: Cambodian Ministry of Labour and Vocational Training, Ministry of Economy and Finance, Ministry of Health, 2017.

43. Royal Government of Cambodia. Subdecree on establishment of social security scheme on health care for persons defined by the provisions of Labour Law. Phnom Penh: Royal Government of Cambodia, 2016.

44. Royal Government of Cambodia. National social protection policy framework 2016-2025. Phnom Penh: Royal Government of Cambodia, 2017. http://inndec.com/library/docs/SPPF%20English%20-%20Final%20Ver.pdf (accessed Mar 10, 2020).

45. National Social Security Fund. Achievement over the period of 10 years (2008–2017) and direction for 2018. Phnom Penh: National Social SecurityFund, 2018.

46. Ministry of Labour and Vocational Training. Prakas on health care benefit. Phnome Penh: Ministry of Labour and Vocational Training, 2016.

47. Cambodian Ministry of Labour and Vocational Training, Ministry of Health. Inter-ministerial prakas on provider payment methods for health care. Phnom Penh: Cambodian Ministry of Labour and Vocational Training, Ministry of Health, 2016.

48. Cambodian Ministry of Labour and Vocational Training, Ministry of Health. Inter-ministerial prakas on revision of annex 1 of inter-ministerial prakas no. 173 LV/PRK. Phnom Penh: Cambodian Ministry of Labour and Vocational Training, Ministry of Health, 2016.

49. Social Protection & Labor Team. Measuring income and poverty using proxy means test. Dhaka: World Bank, 2020. https://olc.worldbank.org/sites/default/files/1.pdf (accessed Aug 4, 2020).

50. National Institute of Statistics. Cambodia Socio-Economic Survey 2014. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2015.

51. El-Gilany A-H, El-Wehady A, El-Wasify MA. Updating and validation of the socio-economic status scale for health research in Egypt. *E Mediterr Health J* 2012; 18: 962–8.

52. Gaur KL. Socio-economic status measurement scale: thirst area with changing concept for socio-economic status. *Int J Innov Res Dev* 2013; 2: 139–45.

53. Singh T, Sharma S, Nagesh S. Socio-economic status scales updated for 2017. *Int J Res Med Sci* 2017; 5: 3264–7.

54. Naga RA, Burgess R. Prediction and determination of household permanent income. October, 2001 (unpublished).

55. Vollmer F, Alkire S. Towards a global assets indicator: re-assessing the assets indicator in the global multidimentional poverty index. Oxford: Oxford Poverty and Human Development Initiative, University of Oxford, 2018. https://www.ophi.org.uk/wp-content/uploads/OPHIRP53a_vs3.pdf (accessed Jan 20, 2019).

56. Ngo DKL. A theory-based living standards index for measuring poverty in developing countries. *J Dev Econ* 2018; 130:190–202.

57. Tajik P, Majdzadeh R. Constructing pragmatic socioeconomic status assessment tools to address health equality challenges. *Int J Prev Med* 2014; 5: 46–51.

58. Chasekwa B, Maluccio JA, Ntozini R, et al. Measuring wealth in rural communities: lessons from the Sanitation, Hygene, Infant Nutrition Efficiency (SHINE) trial. *PLoS ONE* 2018; 13: e0199393.

59. Ferguson BD, Tandon A, Gakidou E, Murry CJL. Estimating permanent income using indicator variables. Geneva: World Health Organization, Evidence and Information for Policy Cluster, 2003. https://www.who.int/healthinfo/paper44.pdf (accessed Aug 5, 2020).

60. Chakrabarty N, Biswas S. A statistical approach to adult census income level prediction. Presentation at 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN); Greater Noida, India; Oct 12–13, 2018.

61. Benin S, Randriamamonjy J. Estimating household income to monitor and evaluate public investment programs in sub-Saharan Africa. Washington DC: International Food Policy Research Institute, 2008. https://core.ac.uk/download/pdf/6388489.pdf (accessed Mar 22, 2019).

62. National Institute of Statistics. Cambodia Socio-Economic Survey 2010. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2011. http://nis.gov.kh/nis/CSES/Final%20Report%20CSES%202010.pdf (accessed July 12, 2018).

63. International Labour Organization Joint Projects Office. Cambodia Socio-Economic Survey 2011, ILO microdata repository. Phnom Penh: International Labour Organization Joint Projects Office, 2017. file:///C:/Users/GHP/Downloads/ddi-documentation-english-340%20(2).pdf (accessed July 12, 2018).

64. National Institute of Statistics. Cambodia Socio-Economic Survey 2012. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2013. http://nis.gov.kh/nis/CSES/Final%20Report%20CSES%202012.pdf (accessed July 12, 2018).

65. National Institute of Statistics. Cambodia Socio-Economic Survey 2013. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2014. http://nis.gov.kh/nis/CSES/Final%20Report%20CSES%202013.pdf (accessed July 12, 2018).

66. National Institute of Statistics. Cambodia Socio-Economic Survey 2015. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2016. http://nis.gov.kh/nis/CSES/Final%20Report%20CSES%202015.pdf (accessed July 12, 2018).

67. National Institute of Statistics. Cambodia Socio-Economic Survey 2016. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2017. http://nis.gov.kh/nis/CSES/Final%20Report%20CSES%202016.pdf (accessed July 12, 2018).

68. National Institute of Statistics. Cambodia Socio-Economic Survey 2017. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2018. http://nis.gov.kh/nis/CSES/Final%20Report%20CSES%202017.pdf (accessed July 12, 2018).

69. Kingdom of Cambodia. Statistics Law. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2017. https://www.nis.gov.kh/index.php/en/national-statistical-systems/legislation/statistics-law (accessed Feb 3, 2021).

70. Institute of Statistical Mathematics (ISM), Statistical Information Institute for Consulting and Analysis (SINFONICA). User's manual for handling resampled micro data of Cambodia Socio-Economic Survey (CSES), CSES 2009 Version 2.0. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2016. https://www.sinfonica.or.jp/eng/information/research/cambodia/Manual%20Cambodia%20CSES%202009%20Ver2.0.pdf (accessed July 12, 2018).

71. National Institute of Statistics. General population census of Cambodia 2008 instructions for cartographic field workers on drawing enumeration area maps. Phnom Penh: National Institute of Statistics, Ministry of Planning, 2006.

72. Deaton A. The analysis of household surveys – a microeconometric approach to development policy. Washington DC: World Bank, 2018.

73. World Bank, International Monetary Fund. Consumer price index (2010=100). https://data.worldbank.org/indicator/FP.CPI.TOTL?locations=KH (accessed Aug 13, 2020).

74. XE com Inc. XE currency converter. Newmarket, Canada: XE.com Inc. https://www.xe.com/currencycharts/?from=USD&to=KHR&view=5Y (accessed Oct 11, 2020).

75. O'Donnell O, Van Doorslaer E, Wagstaff A, Lindelow M. Analyzing health equity using household survey data: a guide to techniques and their implementation. Washington DC: World Bank, 2008. https://openknowledge.worldbank.org/bitstream/handle/10986/6896/424800ISBN978011OFFICIAL0USE0ONLY10.pdf?sequence=1&isAllowed=y (accessed Dec 5, 2019).

76. Wagstaff A. Catastrophic medical expenditures reflections on three issues. Washington DC: World Bank, 2018. https://openknowledge.worldbank.org/bitstream/handle/10986/30878/WPS8651.pdf?sequence=1&isAllowed=y (accessed Feb 12, 2019).

77. Global Link Management Inc. Data collection survey on the social health protection system in the Kingdom of Cambodia. Tokyo: Japan International Cooperation Agency, 2016.

78. Deaton A, Zaidi S. Guideline for constructing consumption aggregates for welfare analysis. Washington DC: World Bank, 2002. https://openknowledge.worldbank.org/bitstream/handle/10986/14101/multi0page.pdf?sequence=1&isAllowed=y (accessed May 9, 2020).

79. Saleem SM. Modified Kuppuswamy socioeconomic scale updated for the year 2019. *Indian J Forensic Community Med* 2019; 6: 1–3.

80. Santosa F, Symes WW. Linear inversion of band-limited reflection seismograms. *SIAM J Sci Stat Comp* 1986; 7: 1307–30.

81. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc* 1996; 58: 267–88.

82. Greene WH. Economic analysis. Harlow, England: Pearson Education Limited, 2020.

83. De Myttenaerea A, Golden B, Le Grand B, Rossic F. Mean absolute percentage error for regression models. July 11, 2017 (arXiv:1605.02541).

84. Nakamura H, Amimo F, Yi S, et al. Implementing a sustainable health insurance system in Cambodia: a study protocol for developing and validating an efficient household income-level assessment model for equitable premium collection. *Int J Equity Health* 2020; 19: 1–6.

85. Le Roy P, Phily C, Merry A. Learning journey: Health Insurance Project (HIP) for the garment sector in Cambodia – to refine the piloted health insurance programme in preparation for transfer to the National Social Security Fund (NSSF). PhnomPenh: GRET Cambodia Office, 2014.

86. Ministry of Health of Cambodia. Inpatient and outpatient indicator reports for ODs, provinces (January to December 2016). Phnom Penh: Ministry of Health of Cambodia, 2017.

87. Hennicot J-C. Cambodia technical note estimation of contribution rates for the health insurance, maternity, and sickness benefits' branch of the National Social Security Fund (NSSF). Phnom Penh: International Labour Organization Joint Project Office, 2015.

88. World Bank. Official exchange rate (LCU per US$, period average) - Cambodia. https://data.worldbank.org/indicator/PA.NUS.FCRF (accessed Nov 3, 2019).

89. Ministry of Health of Cambodia. Guidelines on minimum package of activities for health center development 2008–2015. Phnom Penh: Ministry of Health of Cambodia, 2007.

90. Ministry of Health of Cambodia. National guidelines on complementary package of activities for referral hospital development from 2006 to 2010. Phnom Penh: Ministry of Health of Cambodia, 2006.

91. United Nations Children's Fund. Children in Cambodia facts and figures. New York: United Nations Children's Fund, 2020. https://www.unicef.org/cambodia/children-cambodia (accessed Aug 10, 2020).

92. National Assembly of Kingdom of Cambodia. The Labour Law of Cambodia. Phnom Penh: National Assembly of Kingdom of Cambodia, 1997. http://www.cambodiainvestment.gov.kh/the-labor-law-of-cambodia_970313.html (accessed Aug 10, 2020).

93. Fuchs VR. Who shall live? Health, economics and social choice. New Jersey: World Scientific Publishing, 2011.

94. Iglehart JK. The American health care system – expenditures. *N Engl J Med* 1999; 340: 70–6.

95. World Health Organization, Department of Health System Financing, Cluster of Evidence and Information for Policy. Distribution of health payments and catastrophic expenditures methodology. Geneva: World Health Organization, 2005. https://www.who.int/health_financing/documents/dp_e_05_2-distribution_of_health_payments.pdf (accessed Mar 10, 2020).

96. Xu K, Evans DB, Kawabata K, Zeramdini R, Klavus J, Murray CJL. Household catastrophic health expenditure: a multicountry analysis. *Lancet* 2003; 362: 111–7.

97. Ministry of Health, Labour and Welfare of Japan. Guideline on calculation method of payment and standard insurance rate in the National Health Insurance system [Japanese]. Tokyo: Ministry of Health, Labour and Welfare of Japan, 2015.

98. Tikkanen R, Osborn R, Mossialos E, Djordjevic A, Wharton GA. International health care system profiles Taiwan. https://www.commonwealthfund.org/international-health-policy-center/countries/taiwan#:~:text=Role%20of%20public%20health%20insurance,approximately%203.4%20percent%20of%20GDP.&text=Taiwan's%20NHI%20is%20a%20predominantly%20premium%2Dbased%20social%20health%20insurance%20system (accessed Aug 21, 2020).

99. Medilo LR. PhilHealth contribution table and payment schedule for 2019. https://techpilipinas.com/philhealth-contribution-table-payment-schedule/ (accessed Mar 4, 2019).

100. National Social Security Fund. Report on the annual achievement in 2014 and the action plan for 2015. Phnom Penh: National Social Security Fund, 2014.

101. Tokyo metropolitan government. The basic study report on the National Health Insurance in Tokyo. Tokyo: Tokyo metropolitan government, 1970.

102. World Population Review. Gini coefficient by country 2020. https://worldpopulationreview.com/country-rankings/gini-coefficient-by-country (accessed Aug 24, 2020).

103. O'Donnell O, Van Doorslaer E, Rannan-Eliya RP, et al. The incidence of public spending on healthcare: comparative evidence from Asia. *World Bank Econ Rev* 2007; 21: 93–123.

104. Kakwani NC. Measurement of tax progressivity: an international comparison. *Econ J* 1977; 87: 71–80.

105. Kakwani NC, Wagstaff A, Van Doorslaer E. Socioeconomic inequalities in health: measurement, computation, and statistical inference. *J Econometrics* 1997; 77: 87–103.

106. Wagstaff A, Paci P, Van Doorslaer E. On the measurement of inequalities in health. *Soc Sci Med* 1991; 33: 545–57.

107. Ashcroft R. Health inequities: evaluation of two paradigms. *Health Soc Work* 2010; 35: 249–56.

108. Kibekbaev A, Duman E. Benchmarking regression algorithms for income prediction modeling. *Inform Syst* 2016; 61: 40–52.

109. Howe LD, Hargreaves JR, Gabrysch S, Huttly SRA. Is the wealth index a proxy for consumption expenditure? A systematic review. *J Epidemiol Commun H* 2009; 63: 871-80.

110. Singh A. Predicting income level, an analytic case study in R. New Castle: CloudxLab, 2017. https://cloudxlab.com/blog/predicting-income-level-case-study-r/ (accessed Jan 4, 2019).

111. Lavrakas PJ. Underreporting. In: Encyclopedia of survey research methods. Thousand Oaks: Sage Publications, Inc, 2008. https://methods.sagepub.com/reference/encyclopedia-of-survey-research-methods/n604.xml?fromsearch=true (accessed Jan 3, 2021).

112. Global Link Management Inc. The project for development of social health insurance for the informal sector in the Kingdom of Cambodia: final report. Tokyo: Japan International Cooperation Agency, 2018.

113. Prakongsai P, Limwattananon S, Tangcharoensathien V. The equity impact of the universal coverage policy: lessons from Thailand. *Adv Health* 2007; 21: 57–81.

114. Shimazaki K. Concept and treatment of dependents under the Health Insurance Act [Japanese]. *Shakai Hosyo Kenkyu* 2016; 1: 612–6.

115. Mills A, Ataguba JE, Akazili J, et al. Equity in financing and use of health care in Ghana, South Africa, and Tanzania: implications for paths to universal coverage. *Lancet* 2012; 380: 126–33.

116. Picillo B, Sureresh M, Makinen M. HFG Cambodia cross subsidization report. Washington, DC: United States Agency for International Development, 2016.

117. Ataguba JE, Asante AD, Limwattananon S, Wiseman V. How to do (or not to do) ... a health financing incidence analysis. *Health Policy Plan* 2018; 33: 436–44.

118. Chen L, Yip W, Chang M-C, et al. The effects of Taiwan's National Health Insurance on access and health status of the elderly. *Health Econ* 2006; 16: 223–42.

119. Kankeu HT, Saksena P, Xu K, Evans DB. The financial burden from non-communicable diseases in low- and middle-income countries: a literature review. *Health Res Policy Sy* 2013; 11: 1–12.

120. Arora S. Aging and non-communicable disease. Cham: Springer International Publishing, 2015.

121. Sato H. Theory of confidentiality among the civil servants [Japanese]. *Waseda Law Journal* 1988; 63: 1–62.

122. European Data Protection Board. Members. https://edpb.europa.eu/about-edpb/board/members_en (accessed Dec 31, 2020).

123. European Commission. Data protection in the EU 2020. https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en (accessed Dec 31, 2020).

124. Government Gazette, Thailand. Personal Data Protection Act. https://thainetizen.org/wp-content/uploads/2019/11/thailand-personal-data-protection-act-2019-en.pdf (accessed Dec 31, 2020).

APPENDIX 1: Selected variables and coefficients of prediction models

Variable	Coefficients			
	Model A	Model B	Model C	Model D
Residential Area				
Urban/Rural		-0.0203	-0.0203	-0.0196
Province				
2. Battambang		0.0218		0.0048
3. Kampong Cham		0.0064		
4. Kampong Chhnang		0.0179		
5. Kampong Speu		0.0047		
6. Kampong Thom		0.0410		-0.0070
7. Kampot		0.0111		
8. Kandal		0.0447		0.0220
9. Koh Kong		0.0192		0.0255
10. Kratie		0.0454		0.0361
11. Mondul Kiri	0.1456	0.1172		0.0880
12. Phnom Penh		0.0009		
13. Preah Vihear		-0.0284		-0.0198
14. Prey Veng		-0.0148		-0.0453
15. Pursat		0.0305		
16. Ratanak Kiri		-0.0547		-0.0517
17. Siemreap		0.0247		
18. Preah Sihanouk		0.0367		0.0368
19. Stung Treng		0.0328		0.0002
20. Svay Rieng		-0.0083		
21. Takeo		0.0090		
22. Oddar Meanchey		0.0926		
23. Kep		0.0020		-0.0255
24. Pailin		-0.0221		
Household members				
Household size (sqrt)				0.3209
Adult equalized size (sqrt)		1.0710	1.0710	0.6169
Household head age	-0.0008	-0.0012	-0.0012	-0.0012
Household head sex	-0.0580	-0.0144	-0.0144	-0.0092
Number of working-age members		-0.0951	-0.0951	-0.1031
Number of children	0.0365	0.0768	0.0768	0.0304
Number of dependents		-0.1015	-0.1015	-0.0769
Dependency rate		-0.0642	-0.0642	-0.1563
All-female family (Y/N)				-0.0023
Khmer-ethnic (Y/N)		0.0028	0.0028	0.0129
Number of English-speaking members	0.0241	0.0157	0.0157	0.0096
Number of French-speaking members	0.0298	0.0264	0.0264	
Education				
Number of literate members	0.0532	0.0113	0.0113	0.0084
Number of members with secondary+ education				0.0112
Number of members currently in private school	0.0765	0.0503	0.0503	0.0350
Number of members receiving private lessons	0.0349	0.0339	0.0339	0.0338
Land use				
Number of land parcels used for rice				-0.0036
Number of land parcels used for planting rubber trees				0.0973
Number of land parcels used for bamboo trees	-0.0930	-0.0760	-0.0760	
Housing/living				
Number of own buildings		-0.0707	-0.0707	-0.0047

Total area of buildings used (m2)		-0.0003	-0.0003	-0.0001
Total price value of own buildings (riel)				1.4500
Total monthly rent value of own buildings (riel)	2.6700	3.0500	3.0500	1.4300
Number of buildings used for non-residential purposes		0.0523	0.0523	0.0206
Number of buildings rented out	-0.2144	-0.1795	-0.1795	-0.0547
Number of buildings renovated				0.0740
Total area of dwelling (m2) (log)	0.0932	0.0843	0.0843	0.0659
Number of households sharing the dwelling	-0.0616	-0.0613	-0.0613	-0.0214
Number of rooms in the dwelling	0.0453	0.0412	0.0412	0.0336
The dwelling is own property (Y/N)				-0.0333
Wall of the dwelling is made of bamboo (Y/N)				-0.0387
Wall of the dwelling is made of concrete (Y/N)				0.0277
Wall of the dwelling is made of iron (Y/N)				-0.0244
Roof of the dwelling is made of tiles (Y/N)				0.0293
Roof of the dwelling is made of cement (Y/N)				0.0432
Roof of the dwelling is made of grass (Y/N)	-0.1058	-0.0963	-0.0963	-0.0216
Floor of the dwelling is made of bamboo (Y/N)				-0.0188
Floor of the dwelling is made of tiles (Y/N)				0.0214
Lightening source is electricity (Y/N)				0.0832
Lightening source is car battery (Y/N)				0.0563
Lightening source is kerosene (Y/N)	-0.1190	-0.1001	-0.1001	
Use piped water for drinking (Y/N)				0.0146
Use protected water for drinking (Y/N)				-0.0075
Use latrine toilet (Y/N)				0.0160
Use no toilet (Y/N)				-0.0249
Treat drinking water by any means (Y/N)				0.0100
by boiling water (Y/N)		0.0171	0.0171	
by using chemical (Y/N)		-0.0562	-0.0562	-0.0275
by using alum (Y/N)		0.0396	0.0396	0.0245
by using filter (Y/N)		0.0245	0.0245	0.0025
Number of luxury food items consumed in the past week: 0-5 (fruits, tea/coffee, non-alcoholic drinks, alcoholic drinks and tobacco)	0.0802	0.0720	0.0720	0.0677
Water charge of the previous month (riel)	2.1000	1.7200	1.7200	1.7200
Sewage charge of the previous month (riel)	1.9700	1.7200	1.7200	1.0500
Garbage charge of the previous month (riel)	4.3300	4.1400	4.1400	1.2100
Electricity charge of the previous month (riel)	6.5100	5.2800	5.2800	5.3800
Gas charge of the previous month (riel)	1.8400	1.5800	1.5800	1.4300
Farming activities				
Total land area harvested in the past year (m2)				4.4700
Crop stored on 31 Dec last year (Y/N)				0.0003
Had irrigation system (Y/N)		0.0215	0.0215	0.0063
Number of cattle possessed		0.0063	0.0063	
Number of female cattle possessed				0.0029
Number of buffalo possessed				0.0147
Number of female buffalo possessed				-0.0206
Number of female pigs possessed				0.0015
Number of chickens possessed		0.0003	0.0003	
Number of ducks possessed				3.0700
Had ducks in the past year (Y/N)				0.0157
Operated pond for fish culture (Y/N)	-0.4552	-0.4933	-0.4933	0.0336
Total number of ponds owned	0.4660	0.5056	0.5056	

Caught fish in the past year (Y/N)		-0.0157	-0.0157	-0.0337
Collected firewood in the past year (Y/N)	-0.0822	-0.0582	-0.0582	-0.0319
Collected palm juice, etc. in the forest (Y/N)				-0.0215
Durable goods				
Possession (Y/N)				
air conditioners (AC)				0.0396
bicycles	-0.0477	-0.0374	-0.0374	
cabinets	0.0669	0.0523	0.0523	0.0460
cars	0.1271	0.2465	0.2465	0.1081
car batteries		-0.0418	-0.0418	
carts	0.0417			
computers				0.0347
fans				0.0266
generators		-0.0582	-0.0582	
hand tractors	0.0853	0.0712	0.0712	0.0255
harrows				-0.0638
irons	0.0751	0.0488	0.0488	0.0883
mobile phones		-0.1235	-0.1234	
motorcycles		0.0790	0.0790	0.0582
printers		0.5199	0.5199	
refrigerators				0.0542
row boats				-0.0619
satellite dishes		0.0420	0.0420	
stereos		-0.0203	-0.0203	-0.0451
suitcases				-0.0075
threshing machines				0.0324
TVs	0.0510	0.0312	0.0312	0.0056
Number of durable goods acquired prior to a year ago				
bulldozers				-0.0061
cabinets				-0.0070
cars		-0.1026	-0.1026	
car batteries		0.0175	0.0175	
carts		0.0316	0.0316	
cooking stoves	0.0373	0.0317	0.0317	
dish washers				0.0199
generators	0.0783	0.1175	0.1175	0.0555
harrows	0.0100	0.0082	0.0082	0.0044
mobile phones	0.0348	-0.0399	-0.0399	
motorcycles		0.0326	0.0326	
printers		-0.4983	-0.4983	
pumps				0.0127
satellite dishes		-0.0483	-0.0483	
stereos		0.0278	0.0278	0.0410
suitcases				0.0101
telephones				0.0136
Number of durable goods acquired within a year				
bicycles		0.0352	0.0352	0.0107
cabinets				0.0108
cars		0.1081	0.1081	0.4108
fans				0.0057
freezers	-0.0975			-0.1788
hand tractors				0.0482

mobile phones				0.0100
motorcycles	0.1638	0.1537	0.1537	0.1127
radios				-0.0117
satellite dishes				0.1002
threshing machines	-0.6619	-0.5547	-0.5548	
tractors	1.346	1.2809	1.2809	0.2591
TVs				0.0112
videos				0.0205
Total number of durable goods possessed				
air conditioners (AC)				-0.0248
car batteries				0.0247
cooking stoves				0.0293
dining sets				0.0027
harrows				0.0027
Irons				-0.0376
mobile phones (sqrt)		0.2002	0.2002	0.0678
motorcycles				0.0329
pumps				0.0030
radios				-0.0014
row boats				0.0502
sport goods				0.0018
tractors				0.0171
washing machines				0.0615
Work				
Run a family business (Y/N)	0.0441	0.0198	0.0198	0.0146
Total working hours of household members in the past week	0.0002	-0.0002	-0.0002	-0.0000
Employers: any member employed by … (Y/N)				
a foreign enterprise	0.0418	0.0221	0.0221	0.0006
an aid organization				0.0594
Employment status: any member is … (Y/N)				
an employee				-0.0118
Employment status: number of …				
self-employees				0.0047
unpaid workers				0.0059
Type of occupation: any member is … (Y/N)				
an armed force				0.0148
a high-skilled worker				0.0153
a plant worker		-0.0252	-0.0252	
a service worker				0.0181
an elementary worker	-0.0430	-0.0470	-0.0470	-0.0188
Type of occupation: number of …				
farmers	0.0116	0.0088	0.0088	
service workers				-0.0014
Type of industry: any member engaged in … (Y/N)				
agriculture		-0.0602	-0.0602	-0.0250
fishery	-0.0454	-0.0915	-0.0915	-0.0149
manufacturing	-0.0290	-0.0433	-0.0433	
communication business				-0.0340
construction				-0.0287
energy business				-0.0496
transportation business				-0.0164

sales				-0.0313
accommodation business				-0.0284
research & technologies				-0.0034
education				-0.0210
Type of industry: number of members engaged in …				
fishery		0.0342	0.0342	
manufacturing				-0.0136
energy business				0.0168
service industry				-0.0186
research & technologies				-0.0229
real estate business				0.0053
financial business	0.0723	0.0544	0.0544	
Income and liabilities				
Received scholarship (Y/N)				0.0339
Received gifts (Y/N)				-0.0018
Received remittance (Y/N)	0.0404	0.0437	0.0437	
Received income from lottery in the past year (Y/N)				0.0090
Total number of outstanding loans				0.0280
Total amount of outstanding loans (rial)	5.0000	4.9400	4.9400	1.7000
Sold farm land within a year (Y/N)				0.0205

APPENDIX 2: Sample questionnaire for the household health insurance contribution estimation with Model B

Sample questionnaire for the household health insurance contribution estimation with Model B

Household questionnaire for the health insurance

1	Date of interview	
2	Interviewer's name	

I. Basic Household Information

3	Name of household head	
4	National ID # of household head	
5	Province to reside	
6	Ethnicity	

II. Information of household members

	Name	Relationship to household head	Sex	Date of birth	Literacy	English-specking	French-speaking	Private school	Private lessons	Main occupation	Total working hours in last week	Employers	Industry
1)													
2)													
3)													
4)													
5)													
6)													
7)													
8)													
9)													
10)													
11)													
12)													
13)													
14)													

Notes: Sex (M for male, F for female), date of birth (day/month/year), Main occupation (1. Plant worker, 2. Elementary worker, 3. Farmer, 4. Other), English- and French-Speaking (Y/N), Literacy (able to read and write: Y/N), Private school (currently in private school: Y/N), Private lessons (currently receiving private lessons: Y/N), Employers (1. Foreign enterprise, 0. Not foreign enterprise), Industry (1. Agriculture, 2. Fishery, 3. Manufacturing, 4. Finance, 5. Other)

III. Housing

8	What is the area of your dwelling? (m^2)	
9	How many rooms are there in your dwelling?	
10	How many households share your dwelling?	
11	How much would you have to pay per month to rent a similar dwelling in the area? (riels)	
12	Is the roof of your dwelling made of grass? (Y/N)	
13	Is the lightening source of your dwelling kerosene? (Y/N)	
14	How do you treat your drinking water? (*)	
15	How much did you pay for the following utilities last month? (ask for the receipts)	
	Water (riels)	
	Sewage (riels)	
	Garbage (riels)	
	Electricity (riels)	
	Gas (riels)	

* 1. boil, 2. chemical, 3. alum, 4. filter, 5. None

IV. Property

16	How many buildings so you own?	
17	What is the total area of the buildings used? (m2)	
18	How many buildings do you use for non-residential purposes?	
19	How many buildings do you rent out?	

V. Agricultural activities

20	Do you have irrigation system for farming? (Y/N)	
21	How many cattle do you possess?	
22	How many chickens do you possess?	
23	How many ponds do you own?	
24	Does anyone in your household own or operate a pond for fish culture? (Y/N)	
25	Did anyone in your household catch fish, shrimp, crabs, oysters, etc., in the past 12 months? (Y/N)	
26	Did anyone in your household collect firewood, charcoal, timber or other forest products in the past 12 months? (Y/N)	
27	Did anyone in your household collect palm juice, root crops, herbs, honey or hunt wild animals or birds in the past 12 months? (Y/N)	
28	Does your household use land for growing bamboo trees? (Y/N)	

Note: Pond is a small body of standing water formed naturally or often artificially made. It is smaller than a lake.

VI. Durable goods

29	How many TVs does your household possess?	
30	How many stereos does your household possess?	
	How many of them did your household acquire within the year?	
31	How many mobile phones does your household possess?	
	How many of them did your household acquire within the year?	
32	How many bicycles does your household possess?	
	How many of them did your household acquire within the year?	
33	How many motorcycles does your household possess?	
	How many of them did your household acquire within the year?	
34	How many cars does your household possess?	
	How many of them did your household acquire within the year?	
35	How many car batteries does your household possess?	
	How many of them did your household acquire within the year?	
36	How many generators does your household possess?	
	How many of them did your household acquire within the year?	
37	How many satellite dishes does your household possess?	
	How many of them did your household acquire within the year?	
38	How many cooking stoves did your household acquire within the year?	
39	How many irons does your household possess?	
40	How many cabinets does your household possess?	
41	How many printers does your household possess?	
	How many of them did your household acquire within the year?	
42	How many carts did your household acquire within the year?	
43	How many tractors did your household acquire within the year?	
44	How many hand tractors does your household possess?	
45	How many threshing machines did your household acquire within the year?	
46	How many harrows did your household acquire within the year?	

VII. Others

47	Did anyone in your household consume the following food items in the past week?	
	Fruits (Y/N)	
	Tea/coffee (Y/N)	
	Non-alcohol drink (Y/N)	
	Alcohol drink (Y/N)	
	Tobacco (Y/N)	
48	Does your household run a family business? (Y/N)	
49	Did your household receive remittance from relatives or other households in the past year? (Y/N)	
50	What is the total amount of outstanding loan? (riels)	