Towson University
Department of Economics
Working Paper Series

Working Paper No. 2011-04

Does Health Insurance Decrease Health Expenditure Risk in Developing Countries? The Case of China*

by Juergen Jung and Jialu Liu

April, 2014

Does Health Insurance Decrease Health Expenditure Risk in Developing Countries? The Case of China

Juergen Jung* Towson University

Jialu Liu Streeter† Allegheny College

April 2, 2014

Abstract

This paper studies the impact of health insurance on individual out-of-pocket health expenditures in China. Using China Health and Nutrition Survey data between 1991 and 2006, we apply two-part and sample selection models to address issues caused by censored data and selection on unobservables. We find that, although the probability of accessing health care increases with the availability of health insurance, the level of out-of-pocket health expenditure decreases. Our results from a selection model with instrumental variables suggest that having health insurance reduces the expected out-of-pocket health expenditure of an individual by 29.42 percent unconditionally. Meanwhile, conditional on being subjected to positive health expenditure, health insurance helps reduce out-of-pocket spending by 44.38 percent. This beneficial effect of health insurance weakens over time, which may be attributable to increases in the coinsurance rates of health insurances in China.

JEL: I11, C33, C34

Keywords: Health insurance, medical spending, two-part model, bivariate sample selection model, China Health and Nutrition Survey (CHNS)

*Corresponding author. Towson University, Department of Economics, Stephens Hall, Room 101K, 8000 York Road, Towson, MD 21252, *e-mail:* jjung@towson.edu

†Allegheny College, Department of Economics, 520 Main street, Meadville, PA 16335. *e-mail:* jliu@allegheny.edu

1 Introduction

Current debates about health care reforms in developing countries often focus on improving access to health care at affordable prices in order to reduce the financial burden of households. These reforms invariably deal with the introduction or expansion of either private or public health insurance markets. A widely accepted finding for developed countries is that individuals with health insurance have much lower out-of-pocket health expenditures than their uninsured counterparts (Rubin and Koelln (1993); Waters, Anderson and Mays (2004)). However, the effects of health insurance are not as clear in the context of a developing country where health insurance contracts tend to be less generous (Asgary et al. (2004); Trujillo, Portillo and Vernon (2005); Wagstaff and Lindelow (2008)). In this paper we focus on the role of health insurance in China in the past two decades and examine the impact of health insurance on individual out-of-pocket health expenditure. We address two main questions: First, are insured individuals more likely to use health care and thus report positive medical expenditures more often; Second, among health care users, do the insured pay less out-of-pocket than the uninsured?

Since its economic reform in 1978, China has transformed from a centrally planned economy to a market economy. This transformation was aided by a multitude of large scale economic reforms. One such reform triggered a series of changes in the health care sector. Prior to the reform, hospitals were mostly state-run and non-profit. State Owned Enterprises (SOEs) were fully responsible for health care payments of their employees (Wu (2005)). Since the late 1980s, health care reforms have gradually shifted part of the health care costs from employers to employees (Liu, Tang and Liu (2009)). Government control of drug and health care prices was also loosened, so that prices now reflect health care costs more accurately. The reforms also granted hospitals more autonomy in

choosing more advanced, but often more expensive, technologies and treatments (Hu (1991)). Consequently, health care spending increased more than three-fold between 1990 and 2001, accounting for 5.4 percent of GDP in 2001 (Meng (2004)). Meanwhile, traditional health insurance coverage has been decreasing (Akin and Lance (2004), Du (2009)). The enrollment numbers of traditional insurance, such as public insurance and worker compensation, have declined over time (see Table 1). The share of public health insurance enrollment dropped from 52 percent in 1991 to 7 percent in 2006. Worker compensation insurance enrollment dropped from 32 percent in 1991 to 17 percent in 2004. At the same time, we observe a stark increase in the average coinsurance rate across all health insurance types (Figure 1).

The increase in coinsurance rates during this phase of economic transition can be attributed to several factors. First, urban residents who worked in the state sector lost part or all of their insurance coverage during the period of downsizing state-owned enterprises (Du (2009)). Second, senior professionals, unskilled workers and service workers in urban sectors suffered the largest decrease in insurance coverage during the 1990s (Akin and Lance (2004)). Lastly, individuals living in rural areas worked mostly for agricultural production teams or communes prior to 1978, when the socialist *Cooperative Medical System* (CMS, Hezuo Yiliao Baoxian) provided coverage for them. Economic reform replaced the cooperative production teams with household production units. The old CMS, which had relied on risk pooling among a large number of rural workers, was therefore on the verge of collapsing.

The combination of rising health care expenditures and increases in the coinsurance rates could expose individuals to greater financial risk. Given its implications for China's socioeconomic structure and sustainable economic growth, the effect of health insurance on the financial risk exposure of an individual needs to be carefully examined. Wagstaff

and Lindelow (2008) suggest that health insurance increases the probability of catastrophic out-of-pocket medical spending, based on data from the China Health and Nutrition Survey. Expenses are defined as "catastrophic" if they exceed a given percentage threshold of the household's per capita income. Five thresholds were used: 5, 10, 15, 20 and 25 percent. Wagstaff and Lindelow then use a Probit model to estimate whether having health insurance affects the probability of crossing one of these thresholds. They find that having insurance significantly increases the probability of catastrophic out-of-pocket expenditures at the 5 percent and 10 percent thresholds. They refer to this result as a "curious case" because one would intuitively expect that insurance decreases the probability of "catastrophic out-of-pocket expenses" caused by adverse health events. They explain that health insurance encourages patients to seek expensive care, while health care providers choose more costly treatments for the insured and engage in a type of price gouging. Similarly, Wagstaff et al. (2009) show that the extension of health insurance to rural areas has not reduced out-of-pocket expenses, even though it has increased outpatient and inpatient utilization. The focus of both studies is on whether the insured exhibit a greater likelihood of high health care spending.

Our study adds to this literature in that we analyze what causes these "curious" results triggered by health insurance. In order to accomplish this, we estimate two-part and sample selection models using data from the China Health and Nutrition Survey between 1991 and 2006. Health insurance affects out-of-pocket health expenditures via two channels. The first channel relates to the likelihood that someone uses health care, estimated by a selection equation. The second channel focuses on the amount an individual spends out-of-pocket, conditional on having used health care, estimated by an outcome equation. An instrumental variable (IV) methodology is adopted to address the endogeneity problem of health insurance.

Several findings emerge. Health insurance significantly increases the probability of health care utilization in China – the insured are more likely to seek medical care than the uninsured. This effect is significant across all model specifications. Second, health insurance has a significant effect on reducing the level of out-of-pocket health spending. The results from a selection model with instrumental variables suggest that, having health insurance reduces the expected out-of-pocket payments of an individual by 29.42 percent unconditionally. Meanwhile, conditional on incurring positive expenditures, health insurance helps reduce out-of-pocket spending by 44.38 percent. This effect is even larger in the earlier years of the survey when health insurance plans were more generous and has become weaker in the most recent survey waves.

The paper is organized as follows. Section 2 describes the data and provides descriptive statistics for key variables. Section 3 outlines the main models and estimation techniques. Section 4 analyzes the results and section 5 presents our conclusions. The Appendix contains all tables and figures.

2 Data

We use longitudinal data from the China Health and Nutrition Survey(CHNS from here onwards), collected collaboratively by the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention.[1]

The survey follows a large sample of communities, households, and individuals in 1989, 1991, 1993, 1997, 2000, 2004, 2006 and is still ongoing. Originally, the survey covered eight provinces, Liaoning, Shandong, Henan, Jiangsu, Hubei, Hunan, Guizhou,

[1]See the web site for the CHNS for more information:
http : *//www.cpc.unc.edu/projects/china/data*

and Guangxi, with Heilongjiang added after 1997. The nine provinces exhibit considerable variation in location and level of economic development. A multistage random cluster process is adopted to draw samples in each province. Counties are stratified by income, and a weighted sampling method is used to randomly choose four counties (one low-, two middle-, and one high-income). Additionally, two cities (the capital city and a random lower-income city) are selected. In the next stage, villages and townships within the four counties and urban/suburban areas within the two cities are randomly selected as primary sampling units. Lastly, twenty households are surveyed within each primary sampling unit and all individuals in a household are interviewed (Monda et al. (2008); Guo et al. (1999)). Survey weights at the household and individual level are unavailable, which is one of the reported weaknesses of the CHNS (Popkin et al. (2010)). The survey tracks socioeconomic and demographic information at the household and individual level, such as income, occupation, gender, health, location of residence, education, and marital status.

This study uses six waves of the CHNS between 1991 and 2006,[2] which consists of $27,193$ distinct individuals from $6,333$ households. After focusing on working age adults between 16 and 69, the sample reduces to $50,591$ observations from $20,405$ individuals. The number of observations is greater than the number of individuals because some individuals participated in more than one wave of the survey. In the sample, 25.16 percent of individuals participated in only one wave, 24.99 percent in two waves, 13.95 percent in three, 13.77 percent in four, 7.09 percent in five, and 15.04 percent in all six waves.

Table 2 provides variable definitions and summary statistics. Table 3 shows summary statistics by year for the entire sample and for a subsample of individuals who accessed health care. The base year in the paper is 2006. We use the Chinese consumer

[2]The 1989 wave was excluded due to its different survey design.

price index (CPI)[3] to convert nominal variables, such as income and expenditures, into real variables measured in 2006 Chinese currency Renminbi, which are then converted into units of 2006 US Dollars using the 2006 exchange rate.

2.1 Dependent variables

To evaluate how health spending is affected by insurance, we use two expenditure-related variables as the dependent variables. The first, *MedExp Is Positive*, is a dummy variable indicating whether any medical expenditure was incurred, and the second, *OOP Expenditure*, measures the level of medical spending paid by the patient out-of-pocket.

The variable *Medical Expenditure* denotes total medical expenditures in the four-week period prior to the interview. *Medical Expenditure* is the sum of an individual's costs at all treatment facilities, the cost of preventative care, the cost of immunizations, and any other additional spending on treatments for illnesses, diseases, or injuries. If an individual reports positive medical expenditures, i.e., *Medical Expenditure*> 0, then the indicator variable *MedExp Is Positive* equals one, otherwise it equals zero. About 10 percent of all respondents have positive medical expenditures, hence the value of 0.10 for variable *MedExp Is Positive* in Table 2. Variable *OOP Expenditure* equals *Medical Expenditure* net of any reimbursements from health insurance. For each of the above mentioned medical spending categories, the survey provides a question about the percentage of the spending amount that was reimbursed by insurance. The average out-of-pocket health expenditures for the entire sample (including individuals who did not use health care) is about \$12.9.[4]

[3]The CPI is from the National Bureau of Statistics of China.

[4]Unless otherwise specified, all income and spending figures are converted to constant 2006 US Dollars.

2.2 Explanatory variables

We next turn our attention to the explanatory variables. In this paper we focus on the effect of health insurance, while controlling for income, employment, demographics, health status, and geography.

Health insurance. There have been many types of health insurance in China. However, it is difficult to examine the effect of *each* insurance plan on medical spending, because the availability of insurance plans varies across location and time (Table 1), while the exact timing of the availability at the community level is unknown.

Some conventional plans became less important over time due to certain reforms. For instance, *Public Health Insurance* (Gongfei Yiliao) was mainly offered to government officials as well as retirees and urban employees of large state-owned enterprises (SOEs), while *Worker Compensation Health Insurance* (Laodong Baoxian) was available only for urban workers facing risks of work-related injuries. Due to the downsizing of SOEs in the 1990s and 2000s, the share of the public and worker compensation insurance plans has declined dramatically.

New insurance plans were established to replace the old ones and to accommodate new demand post reform. For example, the *Cooperative Medical System* (CMS, Hezuo Yiliao Baoxian) was established after the foundation of Communist China in 1949 as the main health care network in rural areas. The CMS consisted of individual income contributions, a village collective welfare fund, and subsidies from provincial and central governments (Liu, Hsiao and Eggleston (1999)). Once a crucial tool for improving public hygienic conditions, preventing communicable diseases, and widening rural citizens' access to health care, the CMS became increasingly unsustainable in the late 1980s so that it became difficult for the rural poor to afford health care. Consequently, the government rolled out a *New Rural Cooperative Medical Care System* (NRCMCS,

Xinxing Nongcun Hezuo Yiliao Baoxian) to overhaul the old system (*The Decisions on Improving Rural Public Health by the Chinese Communist Party and Congress* (2002)). As of 2006, this new rural health insurance scheme provided coverage for 61 percent of individuals in the survey.

Another example of new insurance plans is the *Urban Worker Insurance* (or *Uni-Plan*), which was launched in the late 1990s to cover urban employees, many of whom had lost their previous public insurance. Within this framework, employers and employees both pay into the *Basic Health Insurance Fund*, which is used to set up an "individual account" and a "social account" (*TongChou Jijin*). The individual account pays for smaller costs from out-patient services, and the social account pays for larger costs from in-patient medical treatments (Ministry of Labor and Social Security, PRC (2013)). As of 2006, 30 percent of surveyed individuals were covered by the *Urban Worker Insurance.*

Because the availability of health insurance plans has changed frequently, we cannot monitor the effect of each plan. Instead, we focus on the general impact of insurance on one's financial burden without attempting to control for the differences between the various insurance schemes. Our approach is consistent with Wagstaff and Lindelow (2008). We use a dummy variable, *Insurance*, to indicate whether an individual is enrolled in any health insurance plan. On average, 31 percent of individuals are covered by health insurance. The coverage decreases from 32 percent in 1991 to the lowest point of 23 percent in 2000, and then rises to 51 percent in 2006.

Table 4 presents summary statistics of insured and uninsured health care users. Both groups show comparable medical spending levels. However, the uninsured are exposed to a heavier financial burden. For example, in 1991, the average out-of-pocket expenditure to income ratio is 2.22 for the uninsured, compared to 0.07 for the insured.

The differences in income and wealth levels between the two groups are distinct. Insured individuals have about twice the income and considerable larger holdings of durables. In addition, a higher fraction of the insured works for the state sector. The fraction of government officials and executives is much smaller among the uninsured.

Income and employment. The variable *Total Income* represents annual individual income including all cash and non-cash income components from various income sources. The average income increases from $490 in 1991 to $988 in 2006 (Table 3). The average value of durable goods increases from $682 in 1991 to $826 in 2006. Close to three quarters of the individuals in our sample are employed. People with positive medical expenditures work less on average. In addition, the fraction of working individuals decreases from 83 percent in 1991 to 64 percent in 2006. The fraction of workers working at SOEs is about 36 percent in 1991 and decreases over time.

Demographic variables. Educational attainment improves over time with the average length of schooling rising from about six years in 1991 to over eight years in 2006. The average age in the sample is 40 and increases from 37.8 in 1991 to 46.2 in 2006. The fraction of the rural population is 68 percent and stays roughly constant over time, as does the fraction of females which is about 50 percent. The average household size is 4 individuals. About 36 percent of the households have children. Household size decreases over time, from an average of 4.6 individuals in 1991 to 3.8 in 2006. The fraction of married individuals is 79 percent on average and increases over time.

Health status. One of the survey questions asks about the general health status compared to peers in the same age cohort. Individuals report their health status as excellent (14 percent), good (56 percent), fair (26 percent), or poor (4 percent). Individuals with positive medical expenditures report lower health status on average, with only 5 percent in excellent health, 31 percent in good health, 42 percent in fair

health, and 22 percent in poor health.

Geography. The nine provinces in the data vary significantly in terms of geographic location, economy, culture, and ethnic groups. Therefore, province indicator variables are used to control for location effects. Liaoning and Heilongjiang provinces are in northern China. Shandong is located on the eastern coast by the Yellow Sea facing the Korean peninsula. Jiangsu is located south of Shandong and north of Shanghai and is specializing in commerce and export. Henan, Hubei and Hunan are three adjacent inland provinces, serving as transportation and production hubs in central China. Guizhou, located in the southeast, is economically underdeveloped but rich in cultural and environmental resources, with diverse ethnic communities. Guangxi is an autonomous region in the south, neighboring Guangdong province and bordering Vietnam.

3 Estimation strategy

Having discussed the key variables, we now explain the main estimation strategy, data issues, and respective techniques to resolve them.

Data issues. Three data issues arise: First, many individuals do not use health care services during the four week window and therefore report zero health expenditures. To address this data censoring issue, we use a two-part model and a bivariate sample selection model. Second, the distribution of out-of-pocket medical spending is highly right skewed. A logarithmic transformation is used to reduce the skewness of the expenditure data. Without the logarithmic transformation, out-of-pocket expenditures have a mean more than ten times the value of the median. The logarithmic transformation reduces this skewness significantly, with a mean of 3.03 close to the median of 2.75, and the skewness statistic falls from 9.01 to 0.67. The kurtosis is 3.29,

close to the normal value of 3.[5] Third, health insurance may be endogenous if unobserved factors affect an individual's enrollment in health insurance as well as her level of health spending. For example, individuals in greater need for health care have a stronger incentive to purchase an insurance plan, and these individuals are also more likely to use health care facilities and incur more costs. If the endogeneity issue is not addressed, the coefficient estimates will be inconsistent and could be biased.[6] A cluster robust test for exogeneity of the variable *Insurance*, based on the score test by Wooldridge (1995), results in a test statistic of 10.23 with a p-value of 0.004. The test therefore rejects exogeneity of the *Insurance* variable.[7]

Instrumental variables. A general approach to control for endogeneity is the instrumental variables (IV) method. Choosing a proper instrumental variable for the endogenous variable, *Insurance*, is a challenging task. The goal is to find a variable that affects the enrollment of health insurance but not the individual's health condition and medical expenditures directly. Two instruments are used in this study. The first is government official status (IV-1), which is used in Wagstaff and Lindelow (2008). The second is the fraction of individuals having health insurance in an individual's community (IV-2). In CHNS, a community refers to a Neighborhood Committee (Juweihui) in urban areas, and a Township Neighborhood Committee (Xianchen Juweihui) or a village (Cun) in rural areas.

The average health insurance take-up rate (i.e. the fraction of individuals with

[5]Although the logarithmic transformation may reduce skewness, it cannot completely eliminate heteroskedasticity (see Deb, Manning and Norton (2012) for a thorough discussion of these issues).

[6]A thorough discussion about the likely direction of this bias would require a theoretical model that simultaneously explains insurance choice, health care utilization, and other household consumption. Absent such a model we abstain from drawing conclusions as to the direction of the bias and leave this question for future research.

[7]Additional endogeneity issues could arise with respect to the health state variables. Wagstaff and Lindelow (2008) as well as other studies addressing health spending do ignore this potential source of bias and include these variables in an attempt to control for initial health conditions. We follow this procedure for comparability reasons but also because of a lack of suitable instruments in our data.

health insurance) in an individual's community is a good candidate for an instrumental variable for two reasons. First, the government has rolled out various insurance schemes over the years. Such reforms (e.g., adding new plans, phasing out old ones, changing names, and merging plans together) were usually experimented with in a few locations before being introduced gradually in a wider area. For instance, the New Rural Cooperative was first tried out in some townships in 2003 and wasn't available nationwide until 2010. Some insurance plans are specific to rural or urban residents. For example, *Urban Worker Health Insurance* is available only to urban workers, and *New Rural Cooperative Health Insurance* can be purchased only by rural residents. For this reason, the average fraction of health insurance in an individual's community provides a good proxy for the general availability of health insurance. Second, although the average fraction of health insurance in a community is correlated with whether one has insurance, it does not affect one's health conditions and health expenditure directly. This can be confirmed by checking the correlation coefficients: the correlation coefficient between an individual's holding of insurance and the average take-up rate of insurance in her community is 0.57, whereas the correlation between an individual's out-of-pocket spending level and the average take-up rate of insurance in her community is only 0.01.

We next evaluate the relevance and validity of our instrumental variables. As discussed above, the fraction of individuals with insurance in one's community (IV-2) is highly correlated with the endogenous variable, *Insurance.* Moreover, the adjusted R^2 in the first stage regression is 0.41, and the partial R^2 is 0.17. The test statistic of the Cragg-Donald Wald F-test for weak instruments (Cragg and Donald (1993)) exceeds the usual benchmark of 10. Similarly, IV-1 passed the test as a suitable instrument. However, its partial R^2 is much smaller. Therefore, the null hypothesis that the instruments are weak is rejected. [8]

[8]Stock and Yogo (2005) report critical values for the Cragg-Donald statistic for the presence of weak

We use IV-1 only to replicate the results in Wagstaff and Lindelow (2008) as we like to control for work type in the two-part and sample selection models. As one anonymous referee pointed out, the validity of IV-1 may even be compromised. Due to the wide spread corruption issues in China, government officials may have other sources of funds or income to pay for their medical care which is not reflected in the relatively low income that is reported in the survey. This means that medical expenditure as fraction of total income may appear to be high for government officials. If government officials are more likely to have insurance and are also more likely to incur catastrophic expenditure, then this instrument is invalid. Our choice of instrument (IV-2) is less exposed to this problem.

Model estimation. Having discussed the main data issues and respective solutions, we now focus on the empirical models. The goal is to examine the effect of health insurance on the probability of using health care, and on the level of health spending. Two models are suitable for this task: the two-part model and the sample selection model.

Both models contain two equations. First, a *Selection Equation* is a binary equation modeling the probability of positive medical expenditures (*MedExp Is Positive=1* if health spending is positive, and 0 otherwise). Second, an *Outcome Equation* focuses on the (log) level of out-of-pocket health expenditures for individuals reporting positive levels of medical spending. By separating the estimation into these two parts, the model provides a detailed explanation of the impact of health insurance on out-of-pocket health spending, accounting for the fact that some individuals do not use health care and report zero expenditures.

The two-part and the sample selection models differ with respect to the

instruments based on a two-stage least squares bias. Critical values are $16.38, 8.96, 6.66$ and 5.53 for the $5\%, 10\%, 20\%$ and 25% bias respectively. If the Cragg-Donald statistics is less than the critical value then the instruments are weak. This is clearly not the case here.

specification of the error term correlation between the selection and outcome equations. The two-part model assumes that the errors of the selection and outcome equations are independent; hence, there is no correlation between errors in the two equations. The selection model does not impose this independence. If the two error terms are uncorrelated after controlling for observed individual characteristics, then the two equations can be estimated separately and both models are appropriate. However, if the errors are indeed correlated, then the two-part model estimates will be biased, and the selection model is more appropriate (Cameron and Trivedi (2005)).

While estimation of the two-part model is straightforward using a Probit model and a linear regression model separately, estimating the selection model is more complicated. The estimation can be executed by (*i*) Heckman's two-step estimator which augments the OLS regression with an estimate of the omitted regressor using a Probit estimator on the selection equation[9] , (*ii*) Heckman's two-step estimator with exclusion restrictions[10], or (*iii*) a maximum-likelihood estimator which assumes errors are homoskedastic and follow a joint normal distribution. The first method may be only weakly identified. The second method adopts exclusion restrictions to solve the identification problem. Finally, the third method imposes stronger assumptions on the distribution of the error terms.

The choice of a proper exclusion restriction in the second method needs to be considered carefully. The exclusion restriction is included in the selection equation but not the outcome equation. Ideally, the excluded variable affects the probability of using health care but not the level of out-of-pocket spending directly. We consider the variable *Ill Or Injured* to be a good candidate.[11] Variable *Ill Or Injured* is an indicator variable

[9]This omitted regressor, the inverse Mills ratio, is then introduced into the outcome equation as an additional explanatory variable. The correlation between the two errors can then be estimated. The Heckman two-step estimator only requires a linear relationship between the two error terms, but not the joint normality as does the ML estimator. It is therefore less restrictive and more robust to potential misspecification of the error characteristics.

[10]This estimator does not exclusively rely on the non-linearity of the model for parameter identification, which could lead to weak identification and hence biases.

[11]Among the alternative variables that we used for exclusion restrictions were the log of travel costs to

for whether a person is injured, ill, or has some disease in the four-week window prior to the survey. We find that the correlation between *Ill Or Injured* and *MedExp Is Positive* is 0.71, whereas that between *Ill Or Injured* and *OOP Expenditure* is only 0.04. In other words, if a person is injured or ill, then she is more likely to seek health care and therefore incur positive medical expenditures; however, the level of out-of-pocket spending remains undetermined. Thus the variable *Ill Or Injured* is a good exclusion restriction.

4 Estimation results

Having discussed our main estimation strategy and key regression inputs, we now turn our attention to the interpretation of coefficient estimates. In this section, we will first replicate the results from Wagstaff and Lindelow (2008) estimating an IV-Probit model. We then show that their results are driven by aggregating two opposing effects. Next, we present our main contribution by comparing the results from the two-part model, and the selection model with and without exclusion restrictions. Our results are robust across model specifications.

4.1 The "curious" result from Wagstaff and Lindelow (2008)

We first reproduce the "curious" results from Wagstaff and Lindelow (2008). "Catastrophic" out-of-pocket health expenditures are defined as out-of-pocket health expenditures exceeding 5, 10, 15, 20, and 25 percent of income. We then generate dummy variables for when an individual incurs catastrophic out-of-pocket expenses and regress this variable on $Insurance$ and other covariates. Consistent with Wagstaff and Lindelow

the nearest hospital and the time it takes to travel to the nearest hospital by bike. In addition, we estimated the models using an alternative set of control variables which included the type of transportation one would use to access a health care facility, the type of the health care facility, the type of doctors available at the health care facilities, the type of contract one has with the health care facility as well as the availability of necessary drugs at the health care facility. All alternative specifications resulted in very similar estimates for the effect of health insurance.

(2008), we find that having health insurance increases the probability of incurring catastrophic out-of-pocket health expenses, as shown in Table 5.

To control for the endogeneity of health insurance, we first use government official as an instrumental variable (IV-1) as in Wagstaff and Lindelow (2008). As a robustness check, we also use the average insurance coverage in one's community as an instrumental variable (IV-2). The results are very similar. The coefficient of insurance in these Probit models is positive and significant. This result seems to indicate that health insurance exposes the insured to a higher financial risk than the uninsured. As this seems counter intuitive, the term "curious case" is used by Wagstaff and Lindelow (2008). We next present additional insight into these "curious" results.

4.2 Coefficient estimates from two-part and selection models

Table 6 compares the results from the two-part models (columns 1-4) and the selection models estimated using the Heckman two-step procedure (columns 5-10).[12]

Selection equation. The coefficients of health insurance are positive and significant in the selection equation in all model specifications (compare the odd numbered columns in the first row in Table 6). In the two-part model, the coefficient is 0.15 without IV and 0.46 with IV. In the selection model, the coefficient is 0.15 when neither IV nor exclusion restrictions are used, 0.13 when we use an exclusion restriction for model identification but no IV, and 0.17 when both IV and an exclusion restriction are used.

This positive relationship between insurance and the probability of health spending can be caused by several factors. First, individuals in frequent need of health care treatments are more likely to enroll in health insurance (Cutler and Zeckhauser (1998)).

[12]The results from the selection model by maximum-likelihood estimation (MLE) are very similar to our results from the two-step estimation procedure, so that we omit presenting the MLE results.

Second, risk aversion may play a role. Risk-averse individuals are more likely to enroll in health insurance plans and may be more diligent in visiting doctors and having regular check-ups. The third possible reason is moral hazard. In theory, individuals may have a tendency to take risks (taking less care of their health) because part or all of the costs will be covered by insurance. In the context of China where the coinsurance rate is relatively high, moral hazard may play a lesser role than in developed countries, where health insurances are more generous. Finally, some insured individuals may simply wish to use health care to justify the premium they had to pay for insurance. For example, people who paid for dental insurance feel more compelled to get their teeth cleaned regularly. On the other hand, the uninsured are more likely to forgo health care to avoid expensive health care bills.

Outcome equation. A negative and significant relationship is detected between insurance and the (log) level of out-of-pocket health expenditures (compare the even numbered columns in the first row in Table 6). In the two-part model, the coefficient of insurance is -0.89 with no IV and -1.29 with IV. In the selection two-step model, the coefficient is -0.98 with no IV or exclusion restrictions, -0.90 when we use an exclusion restriction but no IV, and -1.32 when both an IV and an exclusion restriction are used.

The negative estimates of insurance on the out-of-pocket health spending level highlight a key contribution of this study. Previous researchers found that the insured are more likely to incur catastrophic out-of-pocket expenses, and thus speculated that health care providers target the insured with more expensive treatments. Our results show that, despite the possibility of price gouging, the insured still pay less out-of-pocket than the uninsured when using health care. As a robustness check we have also tried alternative dependent variables for the outcome equation (i.e., the level of *catastrophic* out-of-pocket spending, and the ratio of out-of-pocket spending to income) which all

result in negative and highly significant coefficient estimates. [13]

Since coinsurance rates in China are very high, the demand for health care is more sensitive to the price of health care than in developed countries where health insurances are more generous. When hospitals and doctors attempt to introduce expensive treatments, patients may choose to opt out of the expensive treatments, and seek more affordable options (Hougaard, Osterdal and Yu (2008)). According to our results, health insurance helps reduce the financial burden of a patient despite the high coinsurance rates. The "surprising" observation in the literature – that is health insurance in China increases the probability of catastrophic health expenditures – could have more to do with increased access to health care and not so much with price gouging of insured patients by providers. In addition, the interaction term with the year 2000 dummy variable is positive, which indicates that the mitigating effect of health insurance is weaker in the three most recent waves (2000, 2003, and 2006). This is obviously a function of the increase in coinsurance rates over time.

Coefficients of other covariates. As shown in the third and fourth rows in Table 6, income and the value of durable possessions only show a weak positive effect on the probability of accessing health care. More affluent individuals show a slightly higher likelihood of using health care, which is in line with the literature that finds health care to be a necessity good in lower income countries (Farag et al. (2012)).

Working individuals are less likely to use health care than their unemployed counterparts and when they do access health care, their out-of-pocket payments are also much lower. This could be because working individuals, compared to the unemployed, are not only healthier, but have access to better health insurance plans through their employers. In addition, government officials and executives show a higher probability of

[13]Estimation results of these robustness checks are available upon request.

using health care.

The coefficients of the dummy variable for rural residence are negative in both equations. Rural residents are less likely to access health care. According to a report by the Ministry of Health in 2011, less than 20 percent of total health care resources are allocated to the rural population (Ministry of Health (2011)). The lack of health care facilities makes it more difficult for the rural residents to seek medical treatment.

Other significant demographic variables include age, gender, family size, education, and general health conditions. Females are more likely to access health care but do not differ in the levels of out-of-pocket expenses from men. Older individuals are less likely to access care, but once they do, they spend more out-of-pocket than younger individuals. Family size has a negative effect on the probability of accessing care. The coefficients of education are small but positive in both equations. Lastly, the estimates of the health status coefficients are consistent with what we would expect. Compared to those in excellent health, people in worse health conditions are more likely to access care and spend more out-of-pocket.

4.3 Marginal effects

So far we have concentrated on the qualitative aspects of health insurance and other covariates. We next turn to the quantitative effects and calculate marginal effects of the selection model following the procedure suggested in Hoffman and Kassouf (2005).[14]

We first need to distinguish between unconditional and conditional marginal effects. If we are interested in the marginal effect of insurance on out-of-pocket health expenditure for individuals who may or may not get sick and thus may or may not use

[14]Other methods that also deal with the log-transformation include the smearing estimator in Duan (1983) and the generalized linear model (GLM) with a log-link function suggested by Deb, Manning and Norton (2012). Both methods have the advantage that the coefficients can be interpreted directly in levels, however, only conditional marginal effects can be estimated. In addition, according to Deb, Manning and Norton (2012), both methods are susceptible to create major precision losses.

medical services, then we calculate the unconditional marginal effect using the entire sample. In comparison, the conditional marginal effect is calculated based on the subsample of individuals with positive medical expenditures. The marginal effects are presented in percentages, in Table 7. The mathematical derivation of these marginal effects is available upon request.

Unconditional marginal effects. Unconditionally, health insurance reduces the level of out-of-pocket expenses by 16.91 percent (compare column 1 in Table 7). If an instrumental variable is used, the reduction is estimated to be 29.42 percent (column 3 in Table 7).

Focusing on results from the selection model with IV, working individuals spend about 19.89 percent less out-of-pocket than the unemployed. Individuals working for state-owned enterprises can expect to spend 15.68 percent less than the non-state sector workers. Rural agents spend 23.68 percent less than their urban counterparts. An additional year of age adds about 5.24 percent to out-of-pocket medical spending. Females have 30.38 percent higher expected out-of-pocket health expenditures. Married individuals outspend their unmarried counterparts by 28.13 percent. Finally, compared to the healthiest group, individuals in fair health spend 166.78 percent more, and those in bad health spend over 600 percent more.

Conditional marginal effects. Conditional on having positive medical spending, health insurance decreases out-of-pocket expenditures by 28.76 percent (column 2 of Table 7). When the IV technique is used, health insurance helps reduce spending by about 44.38 percent among people with positive medical expenditures (column 4 in Table 7). The conditional marginal effects are stronger than the unconditional ones, suggesting that health insurance provides more benefits to individuals who actually do end up using health care.

Earlier we have discussed that some of the more generous insurance types in the 1990s, such as Public Insurance and Worker Compensation Insurance, have either been discontinued, merged, or simply replaced with less generous insurance schemes. In Table 6 we already pointed out that the interaction term with the year 2000 dummy variable is positive which indicates that the higher coinsurance rates in the more recent waves have weakened the benefits of insurance. If we break the sample into pre-2000 and post-2000 waves, we find a very strong negative effect of health insurance for the pre-2000 sample. Conditional on having accessed health care, health insurance decreases the expected out-of-pocket health spending by about 74 percent (column 6). In the post-2000 sample the conditional effect is much smaller at 17 percent. This indicates, that the overall sample effect discussed earlier is mainly driven by the earlier waves, when insurances were more generous in their reimbursements to patients.

4.4 Discussion

The differences between our results and the results in Wagstaff and Lindelow (2008) are driven by two main factors. First, we use a different instrumental variable to control for the endogeneity of the health insurance variable. Wagstaff and Lindelow (2008) adopt government official (IV-1) while we use the average insurance rate in one's community (IV-2). As discussed in Section 3, we argue that IV-2 is a stronger instrument than IV-1. Overall, the choice of instrument will affect our results only quantitatively (to some extent) but does not explain the qualitative difference of our results, i.e., the finding that health insurance reduces the level of out-of-pocket spending of individuals who incur positive health expenditure. This result is driven by our model choice of two-part and selection models which, both, separate the overall effect of health insurance on out-of-pocket expenses into two opposing channels. The first channel relates to the

likelihood that someone uses health care, estimated by a selection equation. The second channel focuses on the amount an individual spends out-of-pocket, conditional on having used health care, estimated by an outcome equation. A simple probit model relating health insurance with the probability of catastrophic spending is not able to distinguish these two channels. Therefore, our model provides a more complete description and estimation of China's health insurance markets.

5 Conclusion

This paper demonstrates that having health insurance in China increases the probability of accessing medical care services; however, among individuals who access health care and therefore report positive medical spending, having health insurance decreases the out-of-pocket health spending levels. In other words, health insurance in China helps reduce patients' financial burden even though health insurance is less generous than in the developed world. We have therefore provided a refinement of earlier results in the literature which demonstrated that health insurance increases the probability of catastrophic health spending. These studies do not distinguish between the "opposing" effects of insurance on the probability of accessing care and on the level of out-of-pocket expenses. In order to identify these two channels, we estimate two-part and sample selection models while also controlling for potential endogeneity issues.

The coefficient estimates of other control variables are similar to findings for the U.S. and other developed countries. Individuals who work, live in rural areas, or have a larger family tend to spend less out-of-pocket on health care. Individuals who are better educated, older, or female tend to spend more. Lastly, health status is shown to be a good predictor of out-of-pocket health care spending. Individuals with the worst health status spend over 600 percent more out-of-pocket than the healthiest individuals in the

sample. Finally, we also document a trend of growing coinsurance rates that weaken the negative effect of health insurance on out-of-pocket spending.

References

Akin, John S., Willam H. Dow and Peter M. Lance. 2004. "Did the distribution of health insurance in China continue to grow less equitable in the nineties? Results from a longitudinal survey." *Social Science and Medicine* 58:293–304.

Asgary, Ali, Ken Willis, Ali Akbar Taghvaei and Mojtaba Rafeian. 2004. "Estimating rural households' willingness to pay for health insurance." *European Journal of Health Economics* 5(3):209–215.

Cameron, Colin A. and Pravin K. Trivedi. 2005. *Microeconometrics, Methods and Applications*. New York: Cambridge University Press.

Cragg, John G. and Stephen G. Donald. 1993. "Testing Identifiability and Specification in Instrumental Variable Models." *Econometric Theory* 9(2):222–240.

Cutler, David M. and Richard J. Zeckhauser. 1998. "Adverse Selection in Health Insurance." *Forum for Health Economics and Policy* v1(2).

Deb, Partha, Willard Manning and Edward Norton. 2012. "Preconference Course: Modeling Health Care Costs and Counts." ASHEcon - Minneapolis Conference 2012.

Du, Juan. 2009. "Economic reforms and health insurance in China." *Social Science and Medicine* 69:387–395.

Duan, Naihua. 1983. "Smearing estimate: a nonparametric retransformation method." *Journal of the American Statistical Association* 78:605–610.

Farag, Marwa, A. K. NandaKumar, Stanley Wallack, Dominic Hodgkin, Gary Gaumer and Can Erbil. 2012. "The income elasticity of health care spending in developing and developed countries." *International Journal of Health Care Finance and Economics* 12(2):145–162.

Guo, X., B. M. Popkin, T. A. Mroz and F. Zhai. 1999. "Food Price Policy Can Significantly Reduce the Negative Effects of the Nutrition Transition in China." *Journal of Nutrition* 129:994–1001.

Hoffman, Rudolfo and Ana Lucia Kassouf. 2005. "Deriving conditional and unconditional marginal effects in log earnings equations estimated by Heckman's procedure." *Applied Economics* 37:1303–1311.

Hougaard, Jens L., Lars Peter Osterdal and Yi Yu. 2008. "The Chinese Health Care System: Structure, Problems and Challenges." *Applied Health Economics and Health Policy* 9(1):1–13.

Hu, T. 1991. "Medical technology transfer in major Chinese hospitals." *International Journal of Technlogogy Assessment in Health Care* 7(1):3–11.

Liu, Guoen, Yan Tang and Lizang Liu. 2009. "A Study on the Medical Insurance Policy of Urban Employees: Individual Account and Medical Expenditure." *Finance and Economics* 1.

Liu, Yuanli, William C. Hsiao and Karen Eggleston. 1999. "Equity in health and health care: the Chinese experience." *Social Science and Medicine* 10:1349–1356.

Meng, Qingyue, Clas Rehnberg Ning Zhuang Ying Bian Goran Tomson Shenglan Tang. 2004. "The impact of urban health insurance reform on hospital charges: a case study from two cities in China." *Health Policy* 68:197–209.

Ministry of Health. 2011. "Counterpart Technical Support between Urban Tertiary Hospitals and Rural Hospitals in China." Ministry of Health of China, 2011.
URL: *http://www.who.int/workforcealliance/forum/2011/09_China.pdf*

Ministry of Labor and Social Security, PRC. 2013. Online, accessed December 2013.
URL: *http://www.molss.gov.cn/gb/ywzn/2005-12/02/content_95815.htm*

Monda, K.L., L.S. Adair, F. Zhai and B.M.Popkin. 2008. "Longitudinal relationships between occupational and domestic physical activity patterns and body weight in China." *European Journal of Clinical Nutrition* 62:1318–1325.

Popkin, Barry M, Shufa Du, Fengying Zhai and Bing Zhang. 2010. "Cohort Profile: The China Health and Nutrition Survey: monitoring and understanding socio-economic and health change in China, 1989-2011." *International Journal of Epidemiology* 39:1435–1440.

Rubin, R.M. and K. Koelln. 1993. "Determinants of Household Out-of-Pocket Health Expenditures." *Social Science Quarterly* 74(4):721–735.

Stock, James H. and Motohiro Yogo. 2005. Testing for Weak Instruments in Linear IV Regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D. W. K. Andrews and J. H. Stock. Cambridge University Press pp. 80–108.

The Decisions on Improving Rural Public Health by the Chinese Communist Party and Congress. 2002. Technical report Chinese Congress.
URL: *http://www.gov.cn/gongbao/content/2002/content_61818.htm*

Trujillo, Antonio J., Jorge E. Portillo and John A. Vernon. 2005. "The impact of subsidized health insurance for the poor: evaluating the Colombian experience using

propensity score matching." *International Journal of Health Care Finance and Economics* 5:211–239.

Wagstaff, Adam and Magnus Lindelow. 2008. "Can insurance increase financial risk? The curious case of health insurance in China." *Journal of Health Economics* 27:990–1005.

Wagstaff, Adam, Magnus Lindelow, Gao Jun, Xu Ling and Qian Juncheng. 2009. "Extending health insurance to the rural population: An impact evaluation of China's new cooperative medical scheme." *Journal of Health Economics* 28:1–19.

Waters, Hugh R., Gerard F. Anderson and Jim Mays. 2004. "Measuring financial protection in health in the United States." *Health Policy* 69:339–349.

Wooldridge, Jeffrey M. 1995. Score diagnostics for linear models estimated by two stage least squares. In *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C. R. Rao*, ed. Maddala G. S., P. C. B. Phillips and T. N. Srinivasan. Oxford Blackwell pp. 66–87.

Wu, Jinglian. 2005. *Understanding and Interpreting Chinese Economic Reform.* ThomsonSouth-Western.

6 Appendix

6.1 Tables

Table 1: Major types of insurance plans

	Commercial (Shangye)	Public (Gongfei)	Worker Compensation (Laobao)	Cooperative (Hezuo)	Uni-Plan (Tongchou)	Urban Worker (Chenzhen zhigong)	No Plans
1991	NE	52%	32%	13% (old coop)	NE	NE	2.3%
1993	NE	58%	25%	1.5% (old coop)	11%	NE	4.1%
1997	NE	42%	19%	32% (old coop)	2.4%	NE	3.98%
2000	55%	44%	22%	25% (old coop)	2.5%	NE	1.65%
2004	9%	27%	17%	30% (new coop)	17%	NE	1.60%
2006	5%	7%	NE	61% (new coop)	NE	30%	1.51%

(a) NE: not exist or phased out.
(b) For individuals aged 20-65 year old.
(c) The "old coop" refers to *Cooperative Medical System* (Hezuo Yiliao Baoxian), which was replaced by the "new coop", *New Rural Cooperative Medical Care System* (Xinxing Nongcun Hezuo Yiliao Baoxian).
(d) In 2006, the *Urban Worker Plan* absorbed all individuals previously under *Uni-plan*.
(e) Only major insurance plans are reported, while small plans such as Immunization plan are not reported here.
(f) Individuals can enroll in more than one plan; therefore, the rows don't necessarily sum up to 100%.
(g) Source: CHNS 1991-2006.

Table 2: Definition of variables and summary statistics

	Definition	Mean	Std.err.
Medical expenditure	Total medical expenditure	17.00	(1.25)
MedExp is positive	Dummy variable, "= 1" if Medical expenditure>0	0.10	(0.00)
OOP expenditure	Out-of-pocket expenditure	12.90	(0.79)
OOPExp is positive	Dummy variable, "= 1" if OOP expenditure>0	0.09	(0.00)
Catastrophic Exp 5%	Dummy variable, "= 1" if OOP expenditure > 5% of income	0.09	(0.00)
Catastrophic Exp 15%	Dummy variable, "= 1" if OOP expenditure > 15% of income	0.07	(0.00)
Insurance	Dummy variable, "= 1" if has health insurance	0.31	(0.00)
Annual premium	Insurance premium	1.69	(0.06)
Directly calculated coinsurance	= OOP expenditure/Medical expenditure	0.75	(0.01)
Indirectly calculated coinsurance	Rates reported by insurance plans	0.39	(0.01)
Total Income	Annual household income per capita	718.49	(5.13)
Durables	Durables per capita	674.28	(4.05)
Education	Education years	6.88	(0.02)
Working	Dummy variable, "= 1" if one is working	0.76	(0.00)
Work for state	Dummy variable, "= 1" if work unit is state owned	0.24	(0.00)
Rural	Dummy variable, "= 1" if one resides in rural areas	0.68	(0.00)
Age	Age in years	41.45	(0.06)
Female	Dummy variable, "= 1" if one is female	0.51	(0.00)
Family size	Nr. of people in the household	4.11	(0.01)
Family with children	Dummy variable, "= 1" if one has children	0.36	(0.00)
Married	Dummy variable, "= 1" if one is married	0.79	(0.00)
Health is excellent	"= 1" if one reports excellent health (base category	0.14	(0.00)
Health is good	"= 1" if one reports good health	0.56	(0.00)
Health is fair	"= 1" if one reports fair health	0.26	(0.00)
Health is poor	"= 1" if one reports poor health	0.04	(0.00)
—Instrumental variable—			
Official (IV-1)	Dummy variable, "= 1" if one is an official	0.04	(0.00)
Mean Frac of Insured (IV-2)	Mean fraction of insured in one's community	0.29	(0.26)
—Exclusive restriction—			
Ill or injured	Dummy variable, "= 1" if one was ill or injured	0.10	(0.00)
Observations		50,591	

Source: CHNS 1991,1993,1997,2000,2004, and 2006
Year and province dummies are not reported.

Table 3: Summary statistics: Full sample vs. health care users

	1991	1993	1997	2000	2004	2006
FULL SAMPLE						
Medical expenditure	7.52	7.21	9.00	16.00	41.45	22.17
MedExp is positive	0.09	0.05	0.06	0.07	0.17	0.14
OOP expenditure	5.31	4.99	5.77	11.13	33.49	17.77
OOPExp is positive	0.07	0.04	0.06	0.06	0.16	0.14
Catastrophic exp 5%	0.04	0.02	0.03	0.04	0.07	0.07
Catastrophic exp 15%	0.02	0.02	0.02	0.03	0.05	0.05
OOP-income ratio	0.12	0.11	0.19	0.23	0.50	0.59
Insurance	0.32	0.27	0.27	0.23	0.27	0.51
Annual premium	0.35	0.32	0.43	1.44	4.20	2.78
Total income	490.63	557.39	510.24	724.81	1085.29	988.74
Durables	704.05	653.91	532.25	672.83	675.43	818.95
Education	6.23	6.30	6.41	6.27	8.02	8.12
Working	0.83	0.82	0.79	0.76	0.66	0.68
Work for state	0.36	0.35	0.26	0.22	0.12	0.12
Rural areas	0.68	0.70	0.67	0.67	0.66	0.68
Age	37.76	38.62	39.62	42.19	45.06	46.15
Female	0.51	0.50	0.50	0.52	0.52	0.52
Family size	4.55	4.50	4.17	3.90	3.65	3.81
Family with children	0.43	0.48	0.41	0.37	0.22	0.21
Married	0.75	0.75	0.75	0.81	0.85	0.87
Health is good	0.62	0.64	0.61	0.51	0.47	0.49
Health is fair	0.21	0.20	0.22	0.28	0.32	0.32
Health is poor	0.04	0.03	0.04	0.05	0.06	0.06
Observations	9,091	8,485	8,846	7,790	8,355	8,019
SUBSAMPLE: Anyone whose medical expenditure is positive						
Medical expenditure	102.02	158.10	143.24	248.26	251.35	162.19
MedExp is positive	1.00	1.00	1.00	1.00	1.00	1.00
OOP expenditure	72.04	109.46	91.80	172.78	203.05	129.99
OOPExp is positive	0.92	0.98	0.98	0.99	0.99	0.99
Catastrophic exp 5%	0.48	0.53	0.55	0.56	0.42	0.50
Catastrophic exp 15%	0.33	0.39	0.37	0.41	0.27	0.33
OOP-income ratio (mean)	1.67	2.46	2.91	3.48	3.01	4.25
OOP-income ratio (median)	0.04	0.05	0.06	0.065	0.03	0.045
Insurance	0.35	0.29	0.35	0.30	0.33	0.56
Annual premium	0.24	0.39	0.57	0.69	4.72	2.10
Total income	445.36	467.64	502.33	773.63	1,213.94	871.97
Durables	644.52	622.39	523.01	666.94	687.41	752.59
Education	5.50	5.57	5.65	5.68	7.49	7.38
Working	0.72	0.76	0.71	0.65	0.58	0.59
Work for state	0.37	0.30	0.24	0.17	0.09	0.09
Rural	0.68	0.68	0.56	0.62	0.60	0.62
Age	43.90	44.00	47.39	47.76	49.28	49.88
Female	0.54	0.56	0.56	0.59	0.57	0.58
Family size	4.35	4.15	3.98	3.63	3.61	3.66
Family with children	0.46	0.43	0.34	0.35	0.20	0.16
Married	0.85	0.85	0.81	0.80	0.85	0.86
Health is good	0.35	0.34	0.28	0.33	0.28	0.28
Health is fair	0.38	0.32	0.43	0.43	0.48	0.46
Health is poor	0.22	0.28	0.25	0.20	0.19	0.21
Observations	670	387	556	502	1,378	1,096

Standard deviations, year and province dummies are not reported.

Table 4: Summary statistics: Uninsured vs. insured when medical expenditure is positive

	1991	1993	1997	2000	2004	2006
SUBSAMPLE: Uninsured whose medical expenditure is positive						
Medical expenditure	91.10	114.24	93.07	167.67	189.92	103.68
MedExp is positive	1.00	1.00	1.00	1.00	1.00	1.00
OOP expenditure	91.10	114.24	93.07	167.67	189.92	103.68
OOPExp is positive	0.99	1.00	1.00	1.00	1.00	1.00
Catastrophic Exp 5%	0.58	0.58	0.61	0.62	0.47	0.52
Catastrophic Exp 15%	0.41	0.44	0.43	0.47	0.31	0.36
OOP-Income ratio (mean)	2.22	3.41	4.18	4.92	4.26	7.64
OOP-Income ratio (median)	0.04	0.05	0.06	0.07	0.03	0.05
Total Income	307.34	301.70	381.64	510.72	833.33	696.58
Durables	356.25	335.42	306.82	419.57	499.30	631.84
Education	4.61	4.55	4.66	5.19	6.73	6.69
Working	0.79	0.78	0.75	0.70	0.60	0.58
Work for state	0.12	0.11	0.10	0.09	0.04	0.03
Rural areas	0.81	0.80	0.69	0.74	0.69	0.63
Age	42.03	43.17	46.92	46.10	48.70	48.82
Female	0.55	0.61	0.57	0.60	0.59	0.58
Family size	4.59	4.31	4.16	3.76	3.75	3.80
Family with children	0.50	0.48	0.37	0.39	0.23	0.16
Married	0.82	0.83	0.79	0.78	0.85	0.83
Health is good	0.36	0.36	0.26	0.33	0.25	0.28
Health is fair	0.36	0.30	0.43	0.40	0.49	0.45
Health is poor	0.22	0.29	0.29	0.22	0.21	0.20
Observations	437	276	361	351	923	486
SUBSAMPLE: Insured whose medical expenditure is positive						
Medical expenditure	122.50	267.17	236.12	435.59	375.95	208.81
MedExp is positive	1.00	1.00	1.00	1.00	1.00	1.00
OOP expenditure	39.17	108.97	89.92	184.80	230.05	151.61
OOPExp is positive	0.78	0.93	0.95	0.95	0.97	0.98
Catastrophic Exp 5%	0.28	0.41	0.44	0.42	0.32	0.48
Catastrophic Exp 15%	0.19	0.27	0.27	0.26	0.20	0.31
OOP-income ratio (mean)	0.70	0.20	0.61	0.28	0.54	1.62
OOP-income ratio (median)	0.01	.02	0.03	0.03	0.02	0.04
Insurance	1.00	1.00	1.00	1.00	1.00	1.00
Annual premium	0.24	0.39	0.57	0.69	4.72	2.10
Total income	704.21	880.25	725.76	1384.77	1986.03	1011.70
Durables	1185.18	1335.95	923.23	1241.93	1069.01	848.80
Education	7.15	8.11	7.47	6.82	9.01	7.92
Working	0.60	0.70	0.64	0.54	0.54	0.60
Work for state	0.83	0.76	0.51	0.38	0.21	0.13
Rural	0.43	0.37	0.32	0.34	0.42	0.61
Age	47.41	46.07	48.26	51.64	50.45	50.72
Female	0.53	0.45	0.54	0.57	0.54	0.58
Family size	3.90	3.74	3.64	3.34	3.33	3.55
Family with children	0.39	0.31	0.29	0.27	0.14	0.16
Married	0.91	0.88	0.86	0.83	0.87	0.88
Health is good	0.35	0.30	0.32	0.33	0.33	0.28
Health is fair	0.41	0.35	0.45	0.48	0.44	0.47
Health is poor	0.21	0.25	0.17	0.15	0.16	0.22
Observations	233	111	195	151	455	610

Standard deviations, year and province dummies are not reported.

Table 5: Replicating exercise: Probit and IV-Probit (marginal probability)

	5% threshold			15% threshold		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	No IV	IV-1	IV-2	No IV	IV-1	IV-2
Insurance	0.004	0.72**	0.13**	0.003	0.85**	0.13*
	(0.12)	(0.03)	(0.04)	(0.14)	(0.02)	(0.05)
log(Total Income)	-0.01***	-0.11***	-0.09***	-0.01***	-0.13***	-0.11***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
log(Durables)	-0.002**	-0.04***	-0.02***	-0.001**	-0.04***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Education (Years)	0.001***	-0.01	0.01**	0.001***	-0.01	0.01*
	(0.00)	(0.48)	(0.05)	(0.01)	(0.33)	(0.08)
Working	-0.008***	-0.10**	-0.09***	-0.01***	-0.14***	-0.12***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Rural	-0.01***	0.01	-0.08***	-0.01***	0.01*	-0.11***
	(0.00)	(0.827)	(0.00)	(0.00)	(0.93)	(0.00)
Age	-0.0005	-0.008	-0.01	-0.00	-0.01	-0.00
	(0.37)	(0.16)	(0.33)	(0.62)	(0.29)	(0.58)
Age^2	0.00	0.0001*	0.0001*	0.00	0.00	0.00
	(0.05)	(0.06)	(0.06)	(0.24)	(0.27)	(0.24)
Female	0.005***	0.06***	0.06***	0.00**	0.04*	0.04*
	(0.00)	(0.01)	(0.01)	(0.05)	(0.07)	(0.06)
Family size	-0.0007***	0.006	-0.01***	0.00	0.02	0.00
	(0.00)	(0.52)	(0.047)	(0.95)	(0.13)	(0.77)
Family with children	-0.001	0.002	-0.01	-0.00	-0.00	-0.01
	(0.68)	(0.93)	(0.73)	(0.54)	(0.94)	(0.57)
Married	-0.004	-0.05*	-0.05	-0.01***	-0.09***	-0.1***
	(0.16)	(0.10)	(0.15)	(0.00)	(0.0)	(0.00)
Health is Good	0.012***	0.13***	0.13***	0.01***	0.12***	0.13***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Health is Fair	0.055***	0.57***	0.60***	0.04***	0.52***	0.56***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Health is Poor	0.135***	1.39***	1.46***	0.10***	1.29***	1.39***
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
First stage R^2		0.30	0.41		0.30	0.41
First stage partial R^2		0.005	0.17		0.005	0.17
DWH test F-statistic		2.66*	5.83***		4.34**	4.75**
		(0.1)	(0.01)		(0.03)	(0.02)
Cragg-Donald test F-statistic		152.11***	8020.37***		152.11***	8020.37**
Observations	50,591	50,591	50,591	50,591	50,591	50,591

p-values in parentheses. *** p<0.01, ** p<0.05, * p<0.10
All regressions include time and region dummy variables (results omitted) and standard errors are clustered by individual.
Instruments: IV-1 uses government official; IV-2 uses mean fraction of insured in one's community
The first stage partial R^2 suggest some need for caution with $IV-1$ as the instrument only explains a small part of the variation of *Insurance* after controlling for the remaining regressors.
The Durbin-Wu-Hausmann test rejects exogeneity of variable *Insurance.*
The Cragg-Donald test rejects that the instruments are weak.

Table 6: Regression results

	TWO-PART MODEL				SELECTION TWO-STEP MODEL					
	No IV		IV		No IV + No Excl.		No IV + Excl.		IV + Excl.	
	(1) Prob.	(2) Level	(3) Prob.	(4) Level	(5) Prob.	(6) Level	(7) Prob.	(8) Level	(9) Prob.	(10) Level
Insurance	0.15***	-0.89***	0.46***	-1.29**	0.15***	-0.98***	0.13**	-0.90***	0.17	-1.32***
	(0.03)	(0.11)	(0.12)	(0.46)	(0.03)	(0.17)	(0.05)	(0.10)	(0.16)	(0.45)
After 2000 × Insurance	0	0.77***	-0.26*	1.12**	0	0.78***	-0.09	0.77***	-0.13	1.13***
	(0.04)	(0.12)	(0.10)	(0.40)	(0.04)	(0.12)	(0.05)	(0.11)	(0.14)	(0.39)
log (Total Income)	0.01**	0.01	0.01*	0.01	0.01**	0	0.01	0	0.01*	0.004
	(0.00)	(0.02)	(0.00)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.01)
log(Durables)	0	0.02	0	0.03	0	0.02	0.02**	0.02	0.02**	0.03
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.02)
Education (Years)	0.004*	0.01*	0.003	0.02*	0.012*	0.01	0.016**	0.02**	0.003	0.02**
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Working	-0.05*	-0.20**	-0.03	-0.23**	-0.05*	-0.17*	-0.01	-0.20***	-0.01	-0.23***
	(0.02)	(0.07)	(0.02)	(0.07)	(0.02)	(0.08)	(0.03)	(0.06)	(0.03)	(0.07)
Official	0.11*	0.18	0.10*	0.2	0.11*	0.12	0.04	0.19	0.04	0.20
	(0.05)	(0.15)	(0.05)	(0.15)	(0.05)	(0.17)	(0.07)	(0.14)	(0.06)	(0.14)
Work for state	-0.07*	-0.01	-0.17***	0.12	-0.07*	0.03	-0.05	0.02	-0.06	0.15
	(0.03)	(0.09)	(0.05)	(0.17)	(0.03)	(0.11)	(0.04)	(0.09)	(0.06)	(0.17)
Rural	-0.10***	-0.23***	-0.08***	-0.25***	-0.10***	-0.17	-0.01	-0.24***	-0.01	-0.25***
	(0.02)	(0.06)	(0.02)	(0.06)	(0.02)	(0.11)	(0.03)	(0.06)	(0.01)	(0.06)
Age	0	0.06***	0	0.06***	0	0.06***	-0.02***	0.05***	-0.02***	0.05***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
Age^2	0	-0.00***	0	-0.00***	0	-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)
Female	0.09***	-0.02	0.09***	-0.01	0.09***	-0.07	0.13***	-0.02	0.13***	-0.01
	(0.02)	(0.05)	(0.02)	(0.05)	(0.02)	(0.10)	(0.02)	(0.05)	(0.00)	(0.05)
Family size	-0.02**	-0.03	-0.02**	-0.03	-0.02**	-0.01	-0.01	-0.03	-0.00	-0.03
	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)
Family with children	0.03	0.05	0.03	0.05	0.03	0.03	0.01	0.04	0.01	0.04
	(0.02)	(0.06)	(0.02)	(0.06)	(0.02)	(0.06)	(0.03)	(0.06)	(0.02)	0.06
Married	0.04	0.09	0.03	0.1	0.04	0.07	0.06	0.09	0.06	0.09
	(0.03)	(0.08)	(0.03)	(0.08)	(0.03)	(0.09)	(0.04)	(0.08)	(0.04)	(0.07)
Health is good	0.18***	-0.01	0.17***	-0.02	0.18***	-0.13	0.09**	-0.07	0.08**	-0.08
	(0.03)	(0.11)	(0.03)	(0.11)	(0.03)	(0.21)	(0.04)	(0.12)	(0.04)	(0.1)
Health is fair	0.72***	0.39***	0.71***	0.39***	0.72***	-0.07	0.30***	0.24*	0.30***	0.23**
	(0.03)	(0.11)	(0.03)	(0.11)	(0.03)	(0.69)	(0.04)	(0.12)	(0.04)	(0.11)
Health is poor	1.58***	1.23***	1.57***	1.23***	1.58***	0.29	0.40***	1.04***	0.39***	1.03***
	(0.04)	(0.12)	(0.04)	(0.12)	(0.04)	(1.40)	(0.06)	(0.13)	(0.06)	(0.02)
Mean Fraction of Insured			0.44***	0.35***					0.44***	0.35***
(*IV-2*)			(0.01)	(0.00)					(0.01)	(0.02)
Ill or Injured							2.78***		0.01***	
(*Excl. Restriction*)							(0.03)		(0.00)	
rho						-0.42		-0.19***		-0.17***
						(1.13)		(0.03)		(0.03)
Observations	50,591	4,786	50,591	4,786	50,591	4,786	50,591	4,786	50,591	4,786

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10
All regressions are run with time and region fixed effects (results omitted) and standard errors clustered by individual.

Table 7: Marginal effects of selection model

	Selection model		**Selection model + IV**		**Sel. model + IV:Year**< 2000		**Sel. model + IV:Year**>= 2000	
	Marginal effects in percent		Marginal effects in percent		Marginal effects in percent		Marginal effects in percent	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Unconditional	Conditional	Unconditional	Conditional	Unconditional	Conditional	Unconditional	Conditional
Insurance	-16.91**	-28.76***	-29.42**	-44.38***	-73.60***	-74.02***	7.70	-17.11
	(8.09)	(4.55)	(14.54)	(11.46)	(9.04)	(8.89)	(27.59)	(21.23)
log(Total income)	3.81	0.96	1.25	1.24	1.06	1.05	0.49	0.48
	(2.39)	(1.55)	(0.45)	(0.45)	(2.68)	(2.66)	(2.08)	(2.07)
log(Durables)	6.36**	1.31	1.30	1.30	3.76	3.72	3.19	3.19
	(2.57)	(1.60)	(0.44)	(0.44)	(2.46)	(2.44)	(2.43)	(2.43)
Education	2.64**	1.93***	2.19***	2.18***	0.66	0.65	2.10**	2.10**
	(1.08)	(0.70)	(0.79)	(0.78)	(1.26)	(1.24)	(1.01)	(1.00)
Working	-18.83**	-15.77***	-19.89***	-16.50***	-27.95***	-15.20	-21.99***	-24.19***
	(7.88)	(5.29)	(5.24)	(5.46)	(8.93)	(10.51)	(6.05)	(5.88)
Official	30.28	18.30	29.95	17.06	38.70	23.70	32.27	33.06
	(27.13)	(16.28)	(19.10)	(17.20)	(26.96)	(24.05)	(26.41)	(26.57)
Work for state	-20.50**	-14.62**	-15.68*	-6.40	30.06	8.81	3.50	33.80*
	(9.88)	(7.04)	(9.30)	(10.33)	(25.19)	(21.07)	(14.11)	(18.24)
Rural areas	-22.86***	-20.96***	-23.68***	-22.24***	-10.84	-13.06	-29.86***	-25.61***
	(6.70)	(4.48)	(4.65)	(4.73)	(9.31)	(9.08)	(5.19)	(5.50)
Age	-1.16	4.40***	5.24***	5.21***	5.22**	5.17**	5.08***	5.07***
	(2.07)	(1.43)	(1.40)	(1.40)	(2.37)	(2.35)	(1.78)	(1.78)
Female	37.87***	2.22	30.38***	-6.19	28.35**	-13.54*	30.61***	-3.70
	(11.04)	(5.37)	(7.14)	(5.14)	(11.07)	(7.46)	(8.91)	(6.57)
Family size	-4.00	-2.38	-2.64	-2.62	-1.24	-1.22	-4.23**	-4.22**
	(2.68)	(1.78)	(1.85)	(1.84)	(2.99)	(2.96)	(2.28)	(2.28)
Family with children	7.92	4.76	7.55	4.21	17.44*	.294725	6.80	8.31
	(9.78)	(6.27)	(6.57)	(6.37)	(10.42)	(8.90)	(8.94)	(9.07)
Married	27.64*	10.39	28.13***	9.16	37.58**	6.54	18.50*	12.19
	(15.64)	(8.64)	(9.65)	(8.22)	(18.53)	(14.35)	(10.63)	(10.07)
Health is good	19.82	-2.87	19.04	-5.23	-20.19	-23.48*	40.82**	1.69
	(18.86)	(11.57)	(12.40)	(9.87)	(14.22)	(13.64)	(17.78)	(12.84)
Health is fair	173.27***	39.83***	166.78***	28.04**	-9.76	7.42	241.84***	37.22**
	(43.00)	(16.67)	(28.90)	(13.87)	(16.98)	(20.21)	(44.55)	(17.88)
Health is poor	665.35***	221.07***	634.04***	184.88***	72.33**	171.14**	853.15***	184.73***
	(133.73)	(40.87)	(90.65)	(35.18)	(35.21)	(55.40)	(146.79)	(43.85)
Observations	50,591	4,786	50,591	4,786	26,420	1,785	24,165	2,996

Robust standard errors are in parenthesis. *** p<0.01, ** p<0.05, * p<0.10
All regressions include time and region dummy variables (results omitted) and standard errors are clustered by individuals.

6.2 Figures

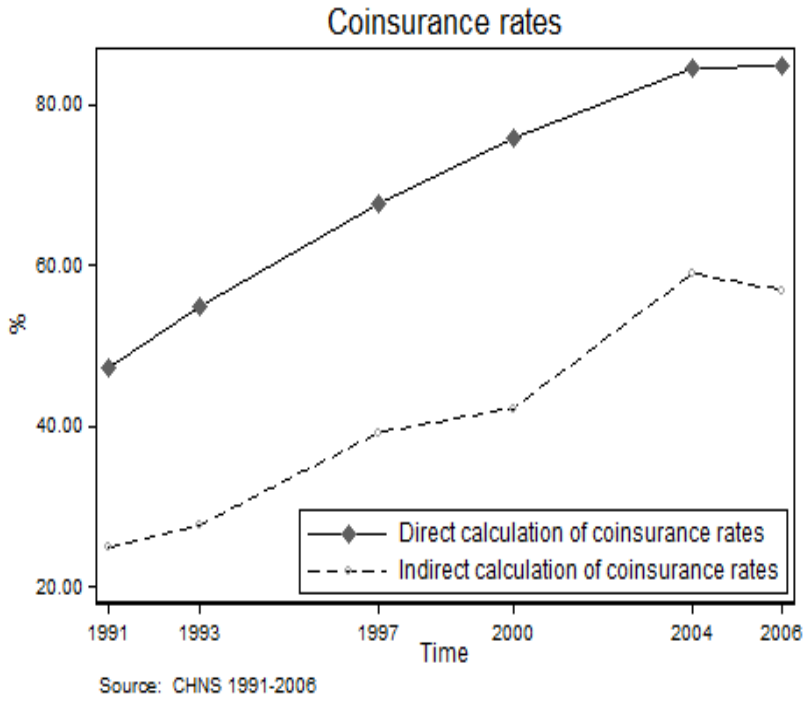

Figure 1: Average coinsurance rates over time.

Notes: (1) Directly calculated coinsurance uses data from individuals who have incurred positive medical spending, in which case the coinsurance rate equals (out-of-pocket spending/total medical costs)× 100%.
(2) Indirectly calculated coinsurance also includes individuals who have not accessed health care and therefore report zero medical spending, in which case the coinsurance rate equals the coinsurance rates stated in the insurance policy.
As shown in the figure, the directly calculated coinsurance rates tend to be higher than the ones stated in the health insurance policy as they include deductibles and copayments.