Risk Management and Healthcare Policy

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

# Adverse Selection as a Barrier to Achieving Universal Public Health Insurance Coverage in China

Panxu Yang[1], Siqi Zhong[2], Xiangping Wang[1], Renyao Zhong[1]

[1]School of Public Management, East China Normal University, Shanghai, People's Republic of China; [2]Faculty of Business and Economics, University of Melbourne, Melbourne, Australia

Correspondence: Renyao Zhong, School of Public Management, East China Normal University, No. 3663, Zhongshan North Road, Putuo District, Shanghai, 200062, People's Republic of China, Email ryzhong297@hotmail.com



**Introduction:** A significant presence of adverse selection in the health insurance market will pose a problem to achieving universal coverage. Public health insurance (PHI) in China is currently facing the challenge of declining enrollments. This situation aligns with the market failure scenario predicted by adverse selection theory.

**Methods:** This study's research sample comprises individuals who are freelancers, self-employed, those who are not actively employed, elderly persons not engaged in employment, and students aged 16 and above. Data from the 2020 wave of the China Family Panel Studies (CFPS) was used to investigate the presence of adverse selection in China's PHI. Logit models were used to analyze the relationship between hospitalization and the decision to enroll in PHI while adopting Bivariate Probit model to address potential bidirectional causality issues arising from "moral hazard."

**Results:** The correlation between coverage and health risk is significantly positive, indicating that individuals who exhibit hospitalization behavior are more likely to access PHI. The heterogeneity analysis reveals that adverse selection behavior is more pronounced among individuals characterized by younger age groups and those with better self-rated health. Furthermore, the mechanism analysis found that previously occurring health risks were positively related to the related risks that could occur after enrolling in PHI, with people using past private health risk information to achieve adverse selection.

**Implication:** The unrestricted enrollment of individuals in PHI may result in adverse selection. Insurers engage in introducing risk-adjusted premiums, and designing PHI as a long-term benefit-oriented product may mitigate the likelihood of adverse selection.

**Keywords:** public health insurance, adverse selection, moral hazard, China

## Introduction

Establishing public health insurance (PHI) has always been a crucial obligation for modern nations in mitigating the financial risks associated with their citizens' healthcare expenses. The inclusion of a mandate in PHI is widely regarded as enhancing efficiency by mitigating the adverse selection problem that undermines the effectiveness of the insurance market. In China, PHI is offered through Urban Employee Basic Medical Insurance (UEBMI) and Urban–Rural Resident Basic Medical Insurance (URBMI). The distinctive feature of China's PHI is the voluntary nature of participation for freelancers, self-employed persons, individuals who are not actively employed, elderly citizens not engaged in employment, and students. The existence of private risk information necessitates careful consideration of adverse selection when allowing for free choice.[1] Examining the presence of adverse selection in the health insurance market holds significant importance. The presence of significant adverse selection in PHI poses a formidable challenge to achieving universal coverage.

As outlined in "Healthy China 2030",[2] the Chinese government has long been committed to attaining universal coverage for PHI. However, granting unrestricted participation to specific individuals may result in adverse selection and ultimately impede the achievement of the above objective. For instance, the population in China lacking coverage under any PHI remains substantial. In 2020, 1,361.31 million people were covered by PHI while China's total population for the year was 1,412.12 million.[3,4] About 50.81 million Chinese were not covered by PHI. The observed phenomena align with the predictions of asymmetric information theory,[1,5,6] which posits that markets will ultimately fail in the presence



Received: 6 December 2024
Accepted: 26 February 2025
Published: 6 March 2025

of individuals possessing private information about their health risks and unrestricted choice to enroll in health insurance. The predominant market failure in the insurance market is adverse selection, which undermines the objective of achieving universal insurance coverage. Some studies have shown that adverse selection exists in China's PHI. For example, some scholars have found that young migrant workers in China have adverse selection into PHI.[7] The absence of underwriting procedures in URBMI has been argued by some scholars as a susceptibility to adverse selection.[8]

The strategy to demonstrate the presence of adverse selection under conditions of asymmetric risk information involves identifying a positive "coverage–risk" correlation.[9–11] However, the research strategy faces the challenge of "moral hazard", a problem that scholars have dedicated their efforts to addressing when examining adverse selection issues in health insurance markets.[12] In the context of health insurance, moral hazard refers to the excessive utilization of healthcare resulting from the presence of insurance coverage.[13] The phenomenon of moral hazard has been consistently observed in numerous studies, demonstrating its potential to contribute to increased medical expenses,[14–17] akin to the concept of adverse selection. Research examined the issues of adverse selection and moral hazard in the US Medigap market.[18] By analyzing these two factors separately, this research discovered that while adverse selection minimally impacts Medigap, moral hazard plays a significant role. Scholars addressed the issue of moral hazard in research by examining the impact of previous periods of utilization of medical services on insurance selection within the current period while investigating adverse selection in the Colombian Managed Care System.[19] The presence of adverse selection has been confirmed. A study revealed no evidence of adverse selection in China's supplementary health insurance market.[20] This study sought to mitigate the influence of moral hazard on healthcare utilization by employing an adjusting risk approach. Previous studies had primarily focused on adverse selection in the private health insurance or community-based health insurance market.[21,22] The law typically mandates individuals to partake in PHI, such as the compulsory membership in social health insurance in Germany.[23] As a result, only a few studies had focused on adverse selection in PHI. However, China's PHI differs in that it allows certain individuals the freedom to choose whether they wish to partake in the insurance program. Therefore, studying the adverse selection within this unique context of China's PHI holds significant importance.

This paper will investigate the presence of adverse selection in China's PHI and elucidate the reasons behind its failure to achieve universal coverage. There is a limited body of literature that has examined adverse selection in PHI within a context of unrestricted choice. The potential adverse selection resulting from the presence of PHI, where individuals have the option to voluntarily enroll, remains an important policy implication. The study of whether an adverse selection problem exists in China's PHI can offer insights into the policy-making process in middle-income countries. The inefficiency caused by adverse selection also exists with freedom of choice in the PHI. Essentially, achieving universal coverage is not feasible through a voluntary PHI.

## Policy Background

Requiring universal participation in PHI can mitigate adverse selection among individuals with asymmetric health risk information, as government-managed PHI typically does not allow for individualized selection based on personal needs. Implementing mandatory participation in PHI is an effective form of government intervention in the market. However, it may also encounter challenges related to adverse selection.[24–26]

Achieving the goal of implementing a comprehensive compulsory PHI would launch challenges for low- and middle-income nations. The execution of a formal labor contract and the establishment of a stable employment relationship are essential prerequisites for eligibility to participate in PHI.[27] The absence of stable employment relationships is a prominent characteristic observed in low- and middle-income nations.[28,29] The presence of a significant number of precarious employment arrangements necessitates the establishment of diverse PHI systems that are not contingent upon stable employment relationships.

The Chinese government offers UEBMI and URBMI, catering to the diverse needs of individuals with varying employment statuses or those not employed. The flexible workforce in China comprises approximately 200 million individuals who operate as self-employed or establish partnerships with large corporations, thereby lacking a stable employer.[30] This situation poses a challenge in China to collect premiums from the "employers" of workers engaged in flexible employment. The government allows the flexible workforce to make premium payments towards UEBMI on behalf of their employers. Simultaneously, different regulations have been formulated by the government for the participation of the flexible workforce in UEBMI compared to regular employees. These regulations enable the flexible workforce to opt for participation in PHI based on their individual preferences.[31] The data indicates that in 2020, a total of 344.55 million individuals were enrolled in UEBMI, with flexible employment accounting for 47.51 million people or approximately 13.8%.[3]

Additionally, there is a segment of the Chinese population engaged in flexible employment or without active employment who are not covered by UEBMI. The implementation of URBMI aims to address this issue.[32,33] The URBMI is open to residents who are freelancers, self-employed, those who are not actively employed, elderly not engaged in employment, and students, allowing for flexibility in participation.[34] The premium associated with URBMI is comparatively lower than that of UEBMI, thus providing a lesser extent of coverage for medical expenses.

The establishment of diverse PHI in middle-income countries like China aims to ensure universal coverage while alleviating the financial burden on low-income individuals. Thus, the government provides protection for individuals with low incomes and unstable employment, offering them the option to participate in PHI voluntarily.

The Social Insurance Law of the People's Republic of China, which was promulgated on October 28, 2010, and came into effect on July 1, 2011 (https://www.gov.cn/jrzg/2010-10/28/content_1732870.htm), emphasizes universal coverage. It neither prohibits individuals from participating in PHI nor imposes restrictive enrollment criteria. Instead, the law explicitly permits certain groups - including residents who are freelancers, self-employed, those who are not actively employed, elderly not engaged in employment - to freely enroll in UEBMI or URBMI. Crucially, the administrative agencies overseeing these programs do not reject applicants, thereby abstaining from risk-based selection. This policy leads to an information imbalance: potential participants retain superior knowledge of their own health risks compared to PHI administrator in China. Consequently, individuals may strategically exploit this asymmetry by enrolling only when they anticipate net financial benefits. These systemic conditions, inherent in China's current social insurance framework, provide the foundational context for our investigation into adverse selection dynamics.

## Data

This paper uses data provided by China Family Panel Studies (CFPS). CFPS is hosted by the Institute of Social Science Survey of Peking University. The CFPS is a nearly nationwide social survey that is intended to serve research needs on a large variety of social phenomena in contemporary China.[35] The CFPS sampling design can be seen in Xie and Lu's study.[35] The CFPS used multistage probability proportional to size sampling with implicit stratification to reduce the operational cost of the survey and better represent Chinese society.[36] CFPS collects data at individual, family, and community levels to track changes in Chinese society, including economy, population, education, and public health. Researchers can apply for access to the CFPS data (http://www.isss.pku.edu.cn). Since the 2010 baseline survey, the CFPS project has successfully conducted multiple rounds of continuous follow-up surveys, with biennial visits. For instance, the first nearly nationwide full-sample follow-up survey was carried out in 2012, while subsequent full-sample surveys were conducted in 2014, 2016, 2018, and 2020.

The decision to utilize 2020 wave CFPS data is based on two key reasons. First, the utilization of cross-sectional data can effectively demonstrate the presence of adverse selection within China's PHI. Second, the coverage of PHI in China has undergone significant transformations in recent years, and the up-to-date survey data can accurately reflect these changes.

According to this paper's research objective, the sample selection adheres to the following principles. First, it includes freelancers, those who are self-employed, persons not actively employed, the elderly not engaged in employment, and students aged 16 and above. Second, it consists of individuals who have complete data and valid information for the variables used in this paper.

The 2020 wave CFPS encompassed a total of 28,530 participants. Based on age information, 2,143 individuals were deleted. Furthermore, the paper excluded 9,577 individuals with employers based on employment-type information. The inclusion of older individuals who had an employer before retirement should also be omitted, and we excluded 60 individuals that had an employer before retirement based on the information regarding retirement status and the type of pension. It is not logical for individuals who are not employees to avail themselves of Free Medical Care, and we excluded 178 individuals of those who did so. Due to missing values and invalid information present in the variables for

5,783 individuals used in this study, we have excluded these individuals from the analysis. Finally, this paper includes a total of 10,789 research samples. The procedure for selecting these individuals is illustrated in Figure 1.

It is important to clarify that the 2018 wave CFPS data was employed for both the robustness and mechanism analyses conducted in this study. Firstly, we utilized variable related to risk aversion from the 2018 wave CFPS data in the robustness test section. Controlling for individual's relative risk aversion is of great significance in studying adverse selection. The 2018 wave CFPS included questions specifically designed to measure risk aversion; however, corresponding questions were not included in the 2020 wave. This omission was due to the restrictions imposed on the CFPS process in 2020, which was the first year of the Covid19 pandemic. As a result, certain information that required face-to-face interviews was not collected in the 2020 wave, with relative risk aversion being one such piece of information that was not gathered. Secondly, we employed a variable from the 2018 wave CFPS data, which indicates whether individuals had experienced hospitalization within the preceding year. The historical hospitalization records of insured individuals can potentially serve as predictors for their future medical expenditures. In order to elucidate the mechanism by which insured individuals manage to enroll in PHI through adverse selection, it is imperative to examine the interplay between individuals' past hospitalization history, their current hospitalization status, and their current participation in PHI.

In our analysis, we primarily utilize data from the 2020 wave CFPS. By employing unique identifier, we link individuals' data to their corresponding records of relative risk aversion and hospitalization status from the 2018 wave CFPS. It is important to acknowledge that, due to the limitations of identifier matching, not all participants can be successfully matched to their 2018 wave data on these variables, resulting in a reduced sample size for our empirical

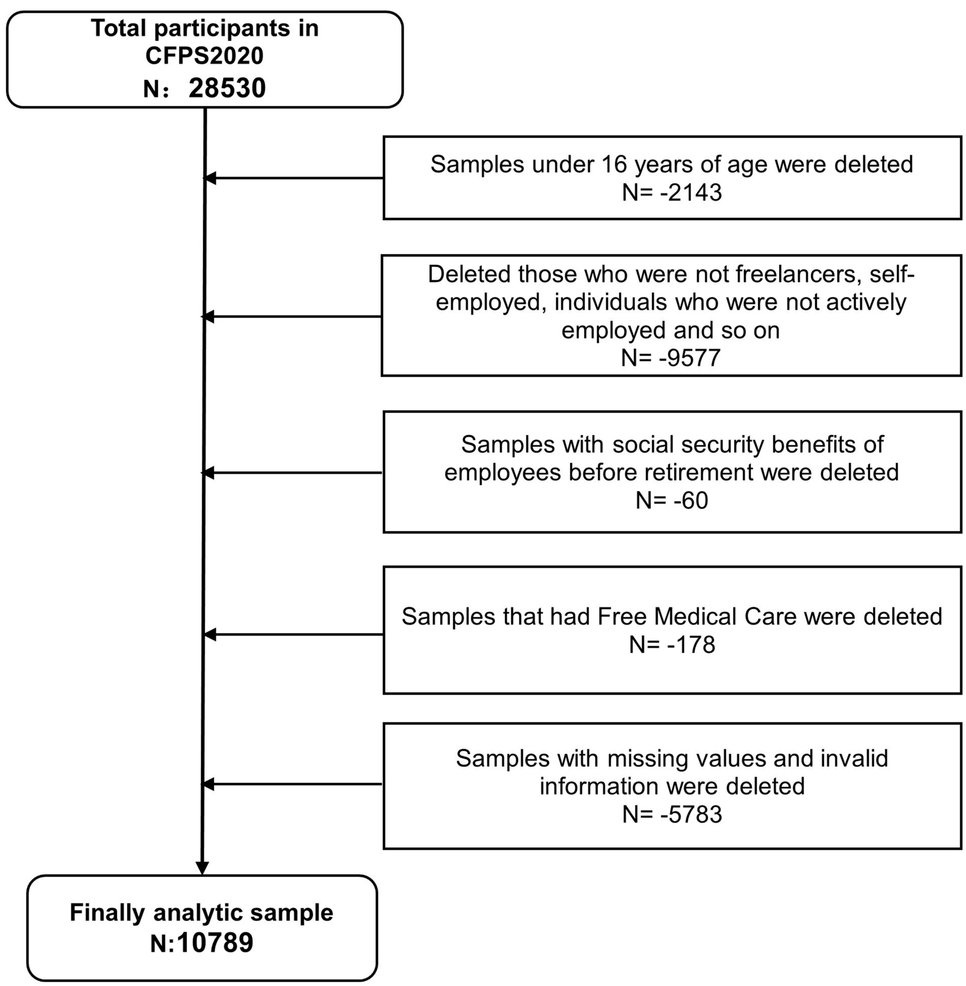


**Figure 1** Sample selection procedure.

model that incorporates both. The rationale for this methodological choice will be elaborated in subsequent sections, along with the presentation of descriptive statistics for these two key variables.

# Variables and Empirical Methods

## Dependent Variable

The dependent variable (PHI) relates to whether the individual participated in UEBMI or URBMI. A value of "1" is assigned if the individual was enrolled in UEBMI or URBMI; otherwise, the value is "0."

## Independent Variable

The independent variable pertains to whether the individual experienced a hospitalization (Hosp) within the past year. In this regard, a value of "1" is assigned if there had been a hospitalization, and "0" otherwise. The presence of adverse selection in medical insurance is typically demonstrated by establishing a positive relationship between coverage and health risk.[37] China's UEBMI and URBMI do not exclude individuals with preexisting health conditions, thereby resulting in the presence of asymmetric health risk information. The presence of adverse selection can be demonstrated when insured individuals exhibit a higher propensity to consume healthcare compared to their uninsured counterparts. The UEBMI and URBMI in China have specific thresholds and ceilings, yet they do provide coverage for a portion of the expenses incurred during hospitalization.[33] The hospitalization experience is utilized as a proxy to assess genuine health risks.

## Control Variables

In the classic review literature, Cohen and Siegelman argue that the control variables in the "coverage-risk" correlation empirical model should therefore be the risk information of the insured, which is used by the insurer for pricing or risk classification.[11] The insurers of UEBMI and URBMI do not engage in risk selection. Thus, the impact of omitting control variables on the empirical model is limited. To achieve a more accurate "coverage-risk" correlation that captures adverse selection in China's PHI, we have incorporated a set of control variables into our empirical model.

The objective variable of hospitalization is utilized in this paper to quantify individual health risks. However, individuals' subjective judgment regarding their health risk information can also influence the realization of adverse selection.[38] A study posited that self-rated health constitutes a reliable indicator of an individual's subjective perception regarding their health risks.[20] The control variables were expanded to include self-rated health (SR-Heal) and self-rated health change (SR-Heal-Chan).

The adverse selection models assumed individuals' risk aversion.[9] Relative risk aversion may affect both an individual's hospitalization and participation in PHI. The habits of smoking,[39,40] drinking,[41] and engagement in physical activity exhibit a correlation with an individual's risk aversion.[42] Therefore, we also incorporated certain control variables that can reflect the extent of individuals' risk aversion. The control variables included smokers (Smoker), the number of cigarettes smoked (Num-Ciga), drinkers (Drinker), the number of exercise sessions (Num-Exer), and the duration of exercise (Dur-Exer).

The occurrence of a disease accompanied by uncertainty necessitates hospitalization, while economic factors can either impede or facilitate the decision to seek hospitalization.[43] Therefore, we included the individual's income (Income) and per capita household income (H-Income) as control variables. After controlling for these economically related variables, the regression model can eliminate the influence of certain factors other than health risks on the dependent variable.

The model included demographic indicators as control variables, such as age (Age), gender (Male), marital status (Married), household registration type (Hukou), urban or rural residence (Urban), education level (Edu), body mass index (BMI), and the presence of diagnosed chronic diseases (Chro-Dise). Economic development varies across different regions in China, leading to potential disparities in UEBMI and URBMI. Additionally, we incorporated dummy variables for provinces (Province) as part of the control variables.

The definitions of all the above variables are presented in Table 1 and the results of the descriptive statistics are presented in Table 2.

**Table 1** Definitions of Variables

| **Variables** | **Symbols** | **Definitions** |
|---|---|---|
| Dependent variable | | |
| Public Health Insurance | PHI | The value is "1" if the individual is enrolled in UEBMI or URBMI, and otherwise, it is "0." |
| Independent variable | | |
| Hospitalization | Hosp | Hosp equals "1", representing the individual who has been hospitalized in the past 12 months, and otherwise, it is "0." |
| Control Variables | | |
| Self-rated Health | SR-Heal | SR-Heal is equal to 1–5, representing excellent, very good, good, fair, poor. |
| Self-rated health change | SR-Heal-Chan | SR-Heal-Chan is equal to "1", representing self-rated health is worse compared to the past year, and otherwise, it is "0." |
| Smoker | Smoker | Smoker is equal to "1", representing the individual that has never smoked, and otherwise, it is "0." |
| Cigarettes | Num-Ciga | Num-Ciga is the average number of cigarettes smoked per day. The unit is a cigarette. |
| Drinker | Drinker | Drinker is equal to "1", representing the individual who drink alcohol three times per week in the past month, and otherwise, it is "0." |
| Amount of exercise | Num-Exer | Num-Exer is equal to 0–7, indicating the frequency of participation in sports, fitness, and leisure activities in the past 12 months. Zero means "never", "1" means "less than once a month on average", and so on, and "7" means "twice a day or more." |
| Duration of exercise | Dur-Exer | Dur-Exer is the duration of each Num-Exer in the past 12 months. The unit is minutes. |
| Income | Income | Income is the individual's self-rated income compared to that of locals. Income is equal to 1–5, representing very low, low, average, high, and very high. |
| Household income | H-Income | H-Income is the household income per capita. The unit is thousands of Yuan. |
| Age | Age | The age range is 16–95. |
| Gender | Male | Male is equal to "1", representing male, and otherwise, it is "0." |
| Marital status | Married | Married is equal to "1", representing married, and otherwise, it is "0." |
| Household registration type | Hukou | Hukou is equal to "1", representing agricultural household registration, and otherwise, it is "0." |
| Urban or rural residence | Urban | Urban is equal to "1", representing residence in urban, and otherwise, it is "0." |
| Education | Edu | Edu is equal to "1", representing the individual who has a bachelor's degree or above, and otherwise, it is "0." |
| Body mass index | BMI | BMI is body mass index. |
| Chronic diseases | Chro-Dise | Chro-Dise is equal to "1", representing the individual who has diagnosed a chronic disease in the past six months, and otherwise, it is "0." |
| Risk aversion | Risk-Aversion | The risk aversion variable takes the value of 1 if the individual chooses "Receive a direct payment of 100 yuan" in the "experiment on risk aversion tendency", and the value of 0 if the individual chooses "Flip a coin and receive 200 yuan if it lands on heads and receive nothing if it lands on tails." |
| Previous hospitalization experience | $Hosp_{t-1}$ | Hospitalization status of the individual during the last wave (2018) CFPS survey. $Hosp_{t-1}$ equals "1", representing the individual who has been hospitalized in the past 12 months in 2018, and otherwise, it is "0." |
| Province | i.Province | Dummy variable for the province |

**Table 2** Descriptive Statistics

| Variables | Total | | Hosp=0 | Hosp=1 | Mean Diff |
|---|---|---|---|---|---|
| | Mean(N=10789) | SD | Mean(N=9390) | Mean(N=1399) | |
| PHI | 0.904 | 0.295 | 0.898 | 0.944 | −0.047*** |
| Age | 51.924 | 15.612 | 50.875 | 58.967 | −8.092*** |
| Male | 0.465 | 0.499 | 0.469 | 0.437 | 0.032** |
| Married | 0.909 | 0.287 | 0.903 | 0.953 | −0.050*** |
| Hukou | 0.795 | 0.404 | 0.796 | 0.793 | 0.003 |
| Urban | 0.421 | 0.494 | 0.422 | 0.414 | 0.008 |
| Edu | 0.017 | 0.128 | 0.018 | 0.006 | 0.012*** |
| BMI | 23.279 | 3.546 | 23.292 | 23.196 | 0.095 |
| Income | 2.997 | 1.116 | 2.992 | 3.034 | −0.043 |
| H-Income | 23.026 | 38.477 | 23.275 | 21.352 | 1.923* |
| Smoker | 0.265 | 0.441 | 0.270 | 0.228 | 0.042*** |
| Num-Ciga | 3.952 | 8.207 | 4.067 | 3.182 | 0.885*** |
| Drinker | 0.127 | 0.333 | 0.133 | 0.087 | 0.046*** |
| Dur-Exer | 15.988 | 33.095 | 15.786 | 17.342 | −1.555 |
| Num-Exer | 1.436 | 2.386 | 1.394 | 1.711 | −0.317*** |
| Chro-Dise | 0.195 | 0.396 | 0.150 | 0.497 | −0.348*** |
| SR-Heal | 3.077 | 1.262 | 2.944 | 3.967 | −1.023*** |
| SR-Heal-Chan | 0.337 | 0.473 | 0.296 | 0.612 | −0.316*** |

**Notes**: SD represents the standard deviation. Mean Diff represents the difference in means. The significance of the t value is determined through either a Chi-square test or a *t*-test. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

## Empirical Model

The objective of this paper is to establish a positive "coverage-risk" correlation model to demonstrate the presence of adverse selection.[11] The utilization of cross-sectional data in the empirical model fulfills the empirical requirements of this study. The dependent variable is a dummy variable, estimated in this paper using a Logit model. The regression model is specified as follows:

$$Logit(P(PHI = 1)) = \beta_0 + \beta_1 Hosp + \beta_2 Controls + \varepsilon \tag{1}$$

The PHI indicates whether the individual has enrolled in PHI. The variable Hosp indicates the hospitalization of the individual in the previous year. Controls refers to the control variables. Logit(P(PHI=1)) represents the log odds of the probability that the dependent variable PHI equals 1. $\beta_0$ is the intercept term, indicating the log odds when the independent variable Hosp and the control variables are all zero. $\beta_1$ is the coefficient for the independent variable, indicating the change in the log odds for a one-unit increase in Hosp holding the control variables constant. $\beta_2$ represents the coefficients for the control variables (Controls). These coefficients indicate the change in the log odds for a one-unit increase in each control variable, holding the other variables constant. $\varepsilon$ is the error term, which captures the random variation that is not explained by the model. It represents the difference between the observed value of the PHI and the value predicted by the model. Using robust standard errors in the Logit model is an effective method to address

 https://doi.org/10.2147/RMHP.S508930 

heteroscedasticity and enhance the robustness of statistical inferences. This paper estimates the Logit model using robust standard errors.

In this paper,$\beta_1$ represents the coefficient of interest. The significant $\beta_1>0$ indicates the presence of a positive correlation between PHI and health risk, highlighting the issue of adverse selection. Certainly, it is also plausible that $\beta_1$ exhibits a significant positive deviation from zero, owing to the presence of moral hazard. In the section dedicated to robustness testing, we thoroughly examine the influence of moral hazard on the empirical findings presented in this paper.

# Empirical Results

## Basic Model Results

The Logit model is employed in this section to estimate the correlation between PHI and health risk. The estimation results are displayed in Table 3. The coefficients of Hosp are all statistically significant at the 1% level in all three models, indicating a positive relationship. The results imply a positive association between coverage and health risk within China's PHI. The empirical model presented in Column (1) of Table 3 lacks any control variables. Compared to

**Table 3** Benchmark Regression Results

| Variables | (1) | (2) | (3) |
|---|---|---|---|
| Hosp | 1.933*** | 1.757*** | 1.570*** |
| | (0.235) | (0.228) | (0.206) |
| Age | | 1.095*** | 1.097*** |
| | | (0.014) | (0.014) |
| Age$^2$ | | 0.999*** | 0.999*** |
| | | (0.000) | (0.000) |
| Male | | 1.300*** | 1.279*** |
| | | (0.112) | (0.111) |
| Married | | 2.060*** | 1.994*** |
| | | (0.236) | (0.236) |
| Hukou | | 1.414*** | 1.266*** |
| | | (0.120) | (0.115) |
| Urban | | 0.816*** | 0.898 |
| | | (0.059) | (0.069) |
| Edu | | 0.549*** | 0.570*** |
| | | (0.105) | (0.112) |
| BMI | | 1.006 | 1.006 |
| | | (0.010) | (0.010) |
| Income | | 1.101*** | 1.070** |
| | | (0.036) | (0.036) |
| H-Income | | 0.999 | 1.000 |
| | | (0.001) | (0.001) |

(*Continued*)

**Table 3** (Continued).

| Variables | (1) | (2) | (3) |
|---|---|---|---|
| Smoker | | 0.836 | 0.802 |
| | | (0.117) | (0.113) |
| Num-Ciga | | 1.002 | 1.003 |
| | | (0.007) | (0.007) |
| Drinker | | 0.899 | 0.968 |
| | | (0.100) | (0.111) |
| Dur-Exer | | 1.003** | 1.004** |
| | | (0.001) | (0.002) |
| Num-Exer | | 1.043** | 1.041* |
| | | (0.021) | (0.022) |
| Chro-Dise | | 1.136 | 1.105 |
| | | (0.117) | (0.115) |
| SR-Heal | | 0.979 | 0.984 |
| | | (0.034) | (0.034) |
| SR-Heal-Chan | | 0.897 | 0.911 |
| | | (0.071) | (0.073) |
| Constant | 8.761*** | 0.296*** | 0.174*** |
| | (0.298) | (0.093) | (0.095) |
| i.Province | No | No | Yes |
| Observations | 10,789 | 10,789 | 10,784 |
| Pseudo $R^2$ | 0.00507 | 0.0491 | 0.0806 |

**Notes**: The dependent variable is PHI. Columns (1)–(3) all report the odds ratio estimated by Logit model. The robust standard error is in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

the empirical model displayed in Column (2) of Table 3, the inclusion of dummy variables for provinces as control variables is introduced in the empirical model reported in Column (3) of Table 3. The results presented in Column 3 of Table 3 indicate that the odds ratio for the effect of Hosp on PHI is 1.570. This means that for a one-unit increase in Hosp, the odds of PHI occurring are expected to be 1.570 times higher, holding all other control variables (Controls) constant. Adverse selection occurs in PHI among individuals with the freedom to opt for participation.

The odds ratio of individuals participating in PHI can be significantly influenced by certain control variables. The variables Age, Male, Married, Hukou, Income, Dur-Exer, and Num-Exer have a significant and positive impact on the odds ratio of PHI occurring. Thus, it is implied that individuals who are older, male, married, possess an agricultural Hukou, have a high income, and have engaged in frequent and long-term exercise exhibit a higher odds ratio of PHI occurring. Additionally, Edu has a significantly negative impact on the odds ratio of PHI occurring.

To comprehensively evaluate the predictive performance of the model, we adopted Receiver Operating Characteristic (ROC) analysis for Table 3 Column (3) model. Figure 2 illustrates the ROC curve, where the horizontal axis represents the False Positive Rate (FPR), and the vertical axis represents the True Positive Rate (TPR). In Figure 2, the solid blue

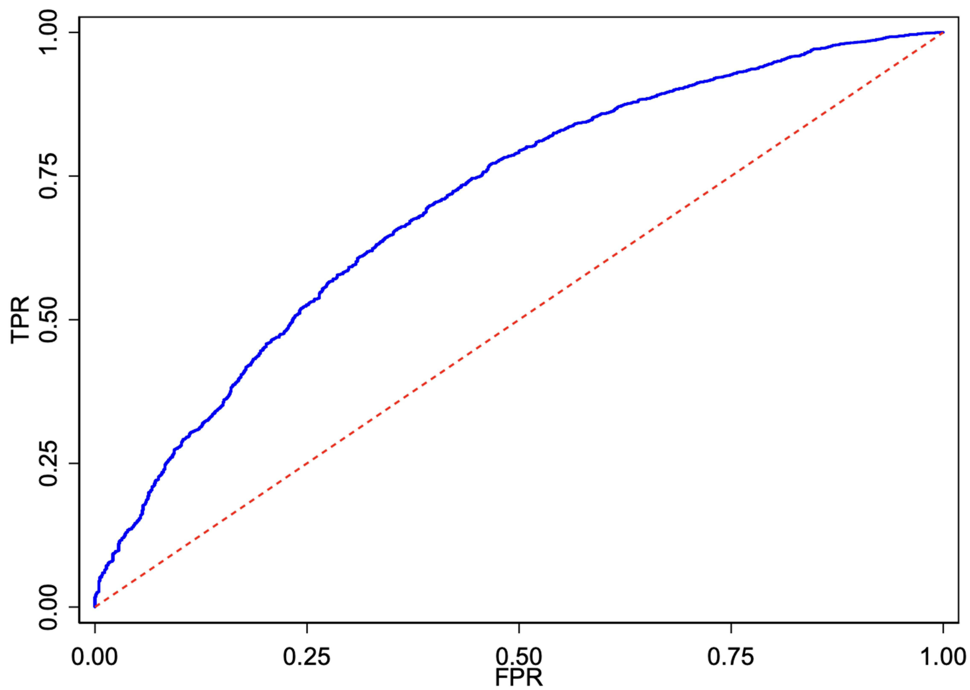


**Figure 2** ROC analysis result.
**Notes**: the solid blue line represents the ROC, and the dashed red line represents random guess.
**Abbreviations**: ROC represents Receiver Operating Characteristic (ROC). FPR represents False Positive Rate (FPR). TPR represents True Positive Rate (TPR).

line represents the ROC, and the dashed red line represents random guess. Through logistic regression (Logit) analysis, combined with the evaluation of the ROC curve, the Area Under the Curve (AUC) of this model reached 0.7038. This result indicates that the model demonstrates moderately high accuracy in distinguishing between PHI=1 and PHI=0, effectively utilizing Hosp and control variables (Controls) to predict the PHI. Therefore, we can infer that the Table 3 Column (3) model possesses certain predictive value in practical applications.

## Robustness Test

### Substituting Dependent Variable and Excluding Student Individuals

In China, the "five insurances" refer to a combination of UEBMI and four other social insurances. Self-employed persons and freelancers in certain provinces of China are required to simultaneously enroll in the Urban Employee Basic Old-age Insurance (UEBOI) when participating in UEBMI. Therefore, in certain provinces, the correlation between UEBMI and health risk may be on par with that of UEBOI and health risk. The existence of adverse selection cannot be assumed if health risks are associated with both UEBMI and UEBOI, as it implies that the individuals may not have the freedom to participate in UEBMI. The empirical model will be employed to examine whether the positive correlation between UEBOI and health risk persists. If this is the case, this paper's empirical findings lack robustness.

In addition, in-school students may face limitations on their financial autonomy. Either parental or institutional regulations might compel them to obtain PHI. Consequently, whether to participate in PHI may not be entirely voluntary for students. We exclude student individuals from the study to test the robustness of the empirical conclusions. If these conclusions remain valid, it confirms the strength of our findings.

The results of the robustness analysis are presented in Table 4. According to the empirical findings in Column (1) of Table 4, no statistically significant correlation was observed between UEBOI and health risk. As indicated by the empirical results in Column (2) of Table 4, the positive correlation between PHI and health risk remains significant even after excluding the individuals of students.

**Table 4** Robustness Test of Substituting Dependent Variable and Excluding Student Individuals

| Variables | UEBOI | PHI(In-School=0) |
|---|---|---|
| | (1) | (2) |
| Hosp | 1.185 | 1.528*** |
| | (0.269) | (0.201) |
| Constant | 0.000*** | 0.259** |
| | (0.000) | (0.173) |
| Controls | Yes | Yes |
| i.Province | Yes | Yes |
| Observations | 10,537 | 10,414 |
| Pseudo $R^2$ | 0.210 | 0.0665 |

**Notes**: UEBOI represents Urban Employee Basic Old-age Insurance (UEBOI). In-School = 0 represents excluding student individuals. Columns (1) and (2) all report the odds ratio estimated by Logit model. The robust standard error is in parentheses. *** p<0.01, ** p<0.05.

## Adding Experimental Risk Aversion as a Control Variable

The individual's risk aversion can influence their selection of insurance as well as the likelihood of risk events occurring. Individuals with a higher degree of risk aversion are more inclined to engage in insurance and actively mitigate the likelihood of encountering risks.[11] The presence of a substantial number of risk-averse individuals in the insurance market can obscure the manifestation of adverse selection.[11] The exclusion of risk aversion, a crucial control variable, may result in an inaccurate correlation between coverage and risk. The inclusion of variables pertaining to individual risk aversion, such as smoking and drinking status, in this paper's basic model is commendable. However, the issue of insufficient control over individual risk aversion persists. The potential influence of risk aversion on the empirical findings of this paper will be examined in this section.

A "Experiment on Risk Aversion Tendency (ERAT)" was included in the CFPS data questionnaire. The CFPS interviewer presented the respondents with two options: "Receive a direct payment of 100 yuan" or "Flip a coin and receive 200 yuan if it lands on heads, and nothing if it lands on tails." The expected utility of a respondent opting to flip a coin is 100 yuan, which is numerically equivalent to the utility of choosing to receive 100 yuan directly. If the respondent opts to flip a coin, they are subject to risk associated with uncertainty. The present study opts to employ the outcomes of the ERAT for assessing individuals' relative aversion to risk. The risk aversion variable is assigned a value of 1 if the individual opts for "receiving 100 yuan directly" in the ERAT, and 0 otherwise.

The CFPS survey in wave 2020, to clarify, did not include data from the ERAT due to the impact of COVID-19. The findings of previous study have demonstrated that the risk attitude of respondents, as assessed through questionnaires, exhibits high test-retest stability, and there exists a significant association between the risk attitude measured by questionnaires and actual risk-taking behavior.[44] We use a unique identifier to obtain the results of ERAT for individuals in the 2018 wave survey. We identified a total of 9,198 individuals who also had recorded Risk-Aversion in the 2018 wave survey without missing values. The average Risk-Aversion value for individuals who participated in PHI was 0.826, while the average Risk-Aversion value for those who did not participate in PHI was 0.800.

Table 5 presents the empirical findings obtained by incorporating the control variables of relative risk aversion into the basic model. The results presented in Column (3) of Table 5 indicate that the odds ratio for the effect of Hosp on PHI is 1.427. This means that for a one-unit increase in Hosp, the odds of PHI occurring are expected to be 1.427 times higher, holding all other control variables (Controls) constant. According to the findings, the positive correlation between

**Table 5** Robustness Test of Adding Experimental Risk Aversion as a Control Variable

| **Variables** | **(1)** | **(2)** | **(3)** |
|---|---|---|---|
| Hosp | 1.776*** | 1.589*** | 1.427** |
| | (0.229) | (0.220) | (0.200) |
| Risk-Aversion | 1.284*** | 1.193** | 1.183* |
| | (0.110) | (0.106) | (0.108) |
| Constant | 7.628*** | 0.231*** | 0.165*** |
| | (0.582) | (0.083) | (0.101) |
| Controls | No | Yes | Yes |
| i.Province | No | No | Yes |
| Observations | 9,198 | 9,198 | 9,185 |
| Pseudo $R^2$ | 0.00560 | 0.0532 | 0.0824 |

**Notes**: The dependent variable is PHI. Columns (1)–(3) all report the odds ratio estimated by Logit model. The robust standard error is in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

PHI and Hosp remains significant even after controlling for variables associated with risk aversion. Adverse selection is observed in the context of PHI in China.

## Using Bivariate Probit Model

According to the existing literature,[12] it is evident that a bidirectional causality relationship potentially exists between health risk and coverage. Participation in the PHI generally results in a reduction in the cost of medical services. The presence of PHI creates an economic incentive for individuals to utilize a greater number of medical services, leading to the phenomenon known as "moral hazard." The adverse selection and moral hazard may establish a bidirectional causality relationship, giving rise to endogeneity issues. Chiappori and Salanié pioneered the utilization of the Bivariate Probit model in examining asymmetric information within the insurance market.[10] The Bivariate Probit model is commonly employed to address bidirectional causality resulting from moral hazard in empirical studies of adverse selection. The Bivariate Probit model is formulated as follows:

$$PHI^* = X_1^T\beta_1 + \mu_1 \tag{2}$$

$$Hosp^* = X_2^T\beta_2 + \mu_2 \tag{3}$$

$$(\mu_1, \mu_2) \sim N\left(\binom{0}{0}, \begin{pmatrix} \rho & 1 \\ 1 & \rho \end{pmatrix}\right) \tag{4}$$

Let the two latent variables $PHI^*$ and $Hosp^*$ be determined by $X_1^T\beta_1 + \mu_1$ and $X_2^T\beta_2 + \mu_2$, respectively, where $(\mu_1, \mu_2)$ follows a bivariate standard normal distribution with a correlation coefficient $\rho$. The observable variables $PHI^*$ and $Hosp^*$ take values based on the positivity or negativity of the latent variables, and the joint probability $P(PHI^*, Hosp^*)$ is calculated using the cumulative distribution function $\Phi\left(X_1^T\beta_1, X_2^T\beta_2, \rho\right)$ of the bivariate normal distribution. The presence of a significant and positive relationship between $PHI^*$ and $Hosp^*$ can be demonstrated if $\rho$ is both statistically significant and greater than 0.

The variables in vector $X_1^T$ exhibit a correlation with an individual's enrollment in PHI. The variables in vector $X_2^T$ are associated with an individual's utilization of inpatient services. In our Bivariate Probit model, the distinction between vector $X_1^T$ and vector $X_2^T$ lies in the inclusion of several variables in vector $X_1^T$ that reflect individual health risk aversion, such as smoking and drinking status. The variables incorporated into vector $X_1^T$ align with the control variables utilized in our basic model.

The test results of Bivariate Probit model are reported in Table 6. We find that $\rho$ is significant and greater than 0, which indicates a significant positive relationship between $PHI^*$ and $Hosp^*$. After addressing the issue of bidirectional causality resulting from moral hazard, we observe that adverse selection persists within China's PHI.

**Table 6** Results for Estimation of Adverse Selection on the Bivariate Probit Model

| Variables | (1) | (2) |
|---|---|---|
| | PHI | Hosp |
| Age | 0.048*** | 0.001 |
| | (0.007) | (0.008) |
| Age$^2$ | −0.000*** | 0.000 |
| | (0.000) | (0.000) |
| Male | 0.130*** | −0.014 |
| | (0.045) | (0.035) |
| Married | 0.380*** | −0.038 |
| | (0.064) | (0.081) |
| Hukou | 0.135*** | 0.067 |
| | (0.047) | (0.048) |
| Urban | −0.048 | 0.045 |
| | (0.039) | (0.038) |
| Edu | −0.324*** | −0.132 |
| | (0.111) | (0.178) |
| BMI | 0.003 | −0.003 |
| | (0.005) | (0.005) |
| Chro-Dise | 0.085* | 0.619*** |
| | (0.051) | (0.039) |
| SR-Heal | −0.001 | 0.229*** |
| | (0.017) | (0.017) |
| SR-Heal-Chan | −0.042 | 0.321*** |
| | (0.041) | (0.038) |
| Income | 0.032* | 0.047*** |
| | (0.017) | (0.016) |
| H-Income | 0.000 | −0.000 |
| | (0.001) | (0.000) |

(*Continued*)

**Table 6** (Continued).

| **Variables** | **(1)** | **(2)** |
|---|---|---|
| | **PHI** | **Hosp** |
| Smoker | −0.117 | |
| | (0.072) | |
| Num-Ciga | 0.002 | |
| | (0.004) | |
| Drinker | −0.026 | |
| | (0.058) | |
| Dur-Exer | 0.002** | |
| | (0.001) | |
| Num-Exer | 0.020* | |
| | (0.011) | |
| Constant | −0.814*** | −3.003*** |
| | (0.304) | (0.403) |
| i.Province | YES | YES |
| $\rho$ | 0.129*** | |
| | (0.033) | |
| Observations | 10789 | |

**Notes**: $\rho$ represents athrho. Columns (1)–(2) all report the coefficients estimated by Bivariate Probit model. The standard error is in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

## Analysis of Heterogeneity

The aim of our research is to provide evidence for the presence of adverse selection in PHI in China. The rationale behind conducting heterogeneity analysis lies in the following factors: First, the analysis of heterogeneity should be capable of elucidating potential variations in adverse selection across diverse individuals. Second, heterogeneity must be accounted for by considering individuals' capacity to effectively utilize asymmetric health risk information.

Age plays a crucial role in influencing an individual's decision to participate in PHI, as it is closely associated with their health risk. The higher odds of a health risk occurring in advanced age groups and the lower odds of a health risk occurring in younger age groups are evident. The longer individuals live and the more life experiences they accumulate, the greater their likelihood of providing accurate assessments of their health. The question of whether the adverse selection into PHI varies across different age groups is of particular interest.

The connection between individuals' subjective assessments of their own health and their adverse selection behavior is a significant aspect that warrants consideration. Some researchers argue that self-rated health may serve as a proxy for asymmetric information.[20] Therefore, examining the combined impact of subjective and objective health risks on the decisions of individuals to participate in PHI can enhance our understanding of adverse selection.

The results of the heterogeneity analysis are shown in Table 7. We examine adverse selection across different populations by incorporating interaction terms between various variables and Hosp into the model. The interaction effect of Age, SR-Heal, and Hosp will all significantly decrease the odds of individuals enrolling in PHI.

We further examine the positive correlation between PHI and health risk, considering variations in Age and SR-Heal, through interaction plot. The horizontal axis in Figures 3 and 4 respectively represents Age and SR-Heal while

**Table 7** Analysis of Heterogeneity

| Variables | (1) | (2) |
|---|---|---|
| Age*Hosp | −0.021** | |
| | (0.009) | |
| SR-Heal*Hosp | | −0.192* |
| | | (0.108) |
| Age | 0.009*** | 0.008*** |
| | (0.003) | (0.003) |
| Hosp | 1.653*** | 1.194*** |
| | (0.561) | (0.454) |
| SR-Heal | 0.001 | 0.015 |
| | (0.034) | (0.036) |
| Constant | −0.512 | −0.514 |
| | (0.524) | (0.524) |
| Controls | Yes | Yes |
| i.Province | Yes | Yes |
| Observations | 10,784 | 10,784 |
| Pseudo $R^2$ | 0.0750 | 0.0747 |

**Notes**: The dependent variable is PHI. Columns (1)–(2) all report the coefficients estimated by Logit model. The robust standard error is in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

the vertical axis represents the marginal effect indicating the impact of Hosp on PHI at given levels of Age or SR-Heal. Adverse selection into PHI gradually weakens with the increase of age, as depicted in Figure 3. In Figures 3 and 4, the solid blue line represents the marginal effects, and the dashed red line represents 95% confidence interval. According to the findings depicted in Figure 4, there exists a slight inverse relationship between SR-Heal and adverse selection.

## Analysis of Mechanism

The mediating effect model can assist in analyzing the occurrence mechanism of adverse selection in PHI in China. We utilize Hayes' identification method and testing procedure to examine the presence of a mediating effect.[45] The mediating effect model is structured as follows:

$$Logit(P(Hosp_t = 1)) = \alpha_0 + \alpha_1 Hosp_{t-1} + \alpha_2 Controls_t + \varepsilon \tag{5}$$

$$Logit(P(PHI_t = 1)) = \beta_0 + \beta_1 Hosp_t + \beta_2 Controls_t + \varepsilon \tag{6}$$

$$Logit(P(PHI_t = 1)) = \gamma_0 + \gamma_1 Hosp_t + \gamma_2 Hosp_{t-1} + \gamma_3 Controls + \varepsilon \tag{7}$$

The mediating effect model examines the relationship between previous hospitalization experience ($Hosp_{t-1}$) as the main independent variable, current participation in PHI ($PHI_t$) as the dependent variable, and current hospitalization experience ($Hosp_t$) as the mediating variable. The control variables remain consistent. The significance of both $\gamma_1$ and $\gamma_2$ indicates that the $Hosp_t$ serves as a partial mediating variable. The $Hosp_t$ acts as

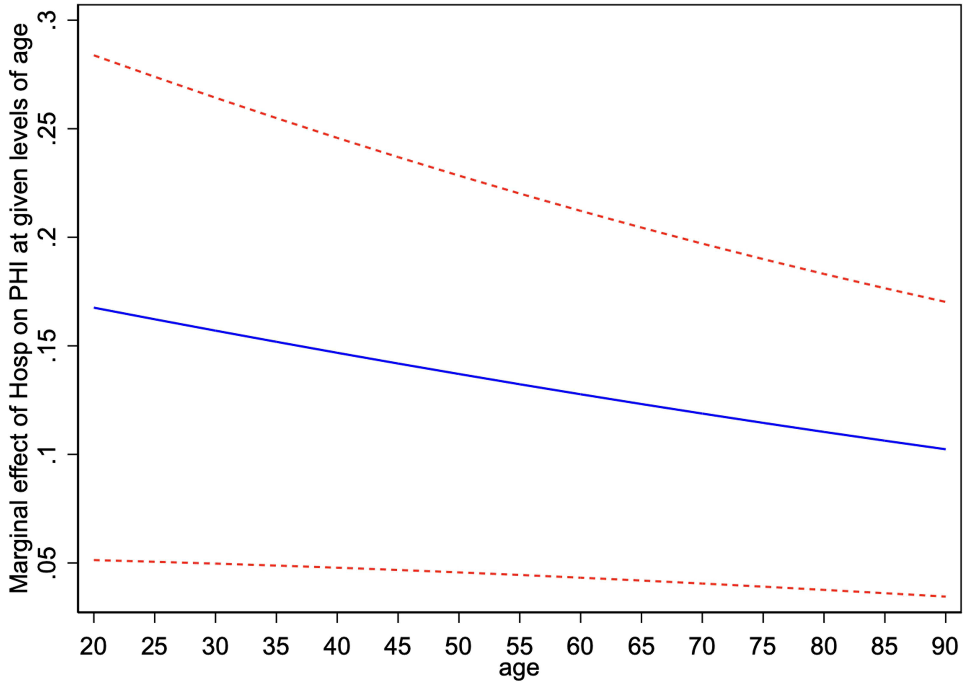


**Figure 3** Marginal effect of Hosp on PHI at given levels of age.
**Notes**: the solid blue line represents the marginal effects, and the dashed red line represents 95% confidence interval.

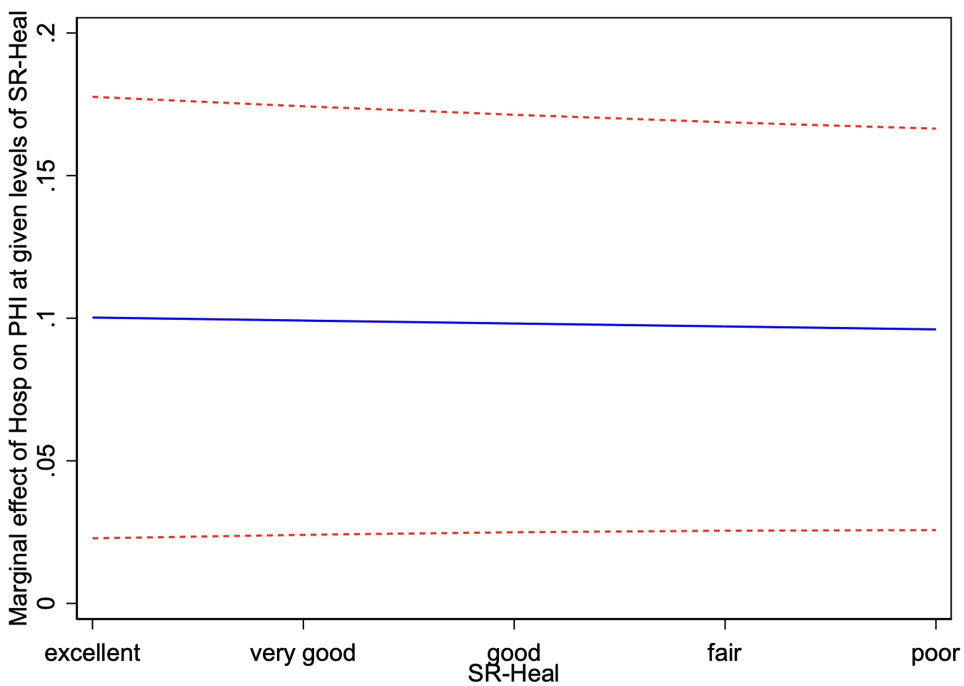


**Figure 4** Marginal effect of Hosp on PHI at given levels of SR-Heal.
**Notes**: the solid blue line represents the marginal effects, and the dashed red line represents 95% confidence interval.

a complete mediating variable if $\gamma_1$ is significant while $\gamma_2$ is not. We identified a total of 9,198 individuals who also had recorded $Hosp_{t-1}$ in the 2018 wave CFPS data without missing values. The average $Hosp_{t-1}$ value for

**Table 8** Results of Mechanism Analysis

| **Variables** | $Hosp_t$ | $PHI_t$ | $PHI_t$ |
|---|---|---|---|
| | **(1)** | **(2)** | **(3)** |
| $Hosp_{t-1}$ | 1.000*** | | 0.298** |
| | (0.080) | | (0.127) |
| $Hosp_t$ | | 0.435*** | 0.306** |
| | | (0.132) | (0.141) |
| Constant | −6.200*** | −1.711*** | −1.692*** |
| | (0.891) | (0.546) | (0.609) |
| Controls | Yes | Yes | Yes |
| i.Province | Yes | Yes | Yes |
| Observations | 9,194 | 10,639 | 9,185 |
| Pseudo $R^2$ | 0.214 | 0.0816 | 0.0828 |

**Notes**: $Hosp_{t-1}$ represents the hospitalization status of the individual during the 2018 wave of the CFPS survey, while $Hosp_t$ and $PHI_t$ represent the hospitalization status and PHI status, respectively, of the individual during the 2020 wave of the CFPS survey. Columns (1)–(3) all report the coefficients estimated by Logit model. The robust standard error is in parentheses. *** p<0.01, ** p<0.05.

individuals who participated in PHI was 0.379, while the average $Hosp_{t-1}$ value for those who did not participate in PHI was 0.114.

The classical literature posits that adverse selection arises due to the asymmetric information,[1] as potential enrollee possess more private knowledge regarding their own health risk. The private health information that can be identified by potential enrollees may contribute to the occurrence of adverse selection.[11] The fact that the potential enrollee had been hospitalized in the previous period indicates a higher likelihood of hospitalization in the current period, noting that hospitalization is a highly objective and readily applicable indicator of private health risk.

The results of the mechanism analysis are presented in Table 8. According to the findings presented in Column (1) of Table 8, it can be observed that there is a significant positive association between $Hosp_{t-1}$ and $Hosp_t$. Furthermore, the findings in Column (3) of Table 8 indicate a statistically significant positive relationship between $Hosp_{t-1}$, $Hosp_t$ and $PHI_t$. Based on the results of the mechanism analysis, it can be inferred that individuals with information on their previous hospitalization experience are more likely to opt for PHI due to adverse selection. The occurrence mechanism of adverse selection lies in individuals making rational decisions based on their private health risk information prior to enrolling in PHI.

## Discussion and Enlightenment

UEBMI and URBMI in China play a crucial role in providing essential risk protection against medical costs for most of the population. The Chinese government allows certain individuals, ie, freelancers, self-employed individuals, unemployed persons, and students, to have the freedom to opt for coverage. The government's intention to allow certain individuals the freedom to choose their participation in PHI is aimed at achieving universal coverage. However, this study reveals that such policy design will result in adverse selection, leading to a gradual decline in PHI coverage.

This study aims to examine the relationship between hospitalization and the likelihood of participating in PHI, finding that individuals with a hospitalization experience are more likely to enroll compared to those without such experience. The finding aligns with the prediction of adverse selection theory,[9] which posits that individuals only choose insurance when it is financially advantageous to do so, given their superior private information regarding health risk compared to insurers. This finding aligns with the current predicament confronting China's PHI. In the presence of adverse selection, an increase in insurance prices leads to a corresponding rise in the marginal cost of new participants.[46] Adverse selection will lead to a "death spiral" in China's PHI. The prerequisite for attaining universal coverage in China's PHI is to prevent adverse selection among specific demographic groups.

The patterns of adverse selection into PHI vary among different age cohorts. In this regard, as age increases, the probability of adverse selection of the individual to participate in PHI gradually decreases. The implication of this discovery is that the ability to use asymmetric private health risk information differs across various age groups. Even when older individuals possess private health risk information, they exhibit a relatively low ability to exploit this informational advantage for the purpose of engaging in adverse selection. There are two reasons accounting for this phenomenon. Firstly, younger individuals exhibit a higher propensity to recognize their health risk information. The condition for the occurrence of adverse selection, according to scholars' perspective, lies in individuals' accurate identification and utilization of such asymmetric information.[11] Most younger individuals enjoy good health and are less likely to experience higher healthcare expenses even in the presence of health risk, indicating that they possess a greater likelihood of accurately predicting their future low healthcare costs. Secondly, the level of risk aversion tends to be lower among individuals in younger age groups. Relevant research has indicated that young individuals exhibit a relatively lower level of risk aversion,[47] and there exists a positive correlation between risk aversion and participation in insurance.[48,49] The belief that they have a low health risk and do not require PHI is more prevalent among younger individuals, implying that adverse selection among this demographic is tantamount to opting out of PHI.

The findings of this study indicate that the likelihood of adverse selection is comparatively low among individuals with poor self-rated health. The distinction between subjective and objective health risk arises from individual variations in the accuracy of self-perceived health risk identification. People with poor self-rated health have a relatively lower likelihood of engaging in adverse selection, indicating their limited ability to utilize their estimated probability of hospitalization for the purpose of selectively obtaining PHI. The findings further suggest that individuals with better self-rated health are more likely to accurately anticipate their future non-hospitalization, whereas those with poorer self-rated health face greater challenges in making such predictions. Consequently, individuals with lower self-rated health are comparatively less capable of utilizing asymmetric health risk information for the purpose of adverse selection.

This study reveals partial mediating effects between prior hospitalization, current hospitalization, and current participation in PHI. The external manifestation of adverse selection is that individuals who enroll in PHI exhibit a heightened likelihood of triggering healthcare coverage payments. However, the underlying cause for adverse selection lies in individuals' ability to anticipate potential medical expenses based on their prior private health risk information and consequently choose to participate in PHI only when they perceive it as financially advantageous. China's PHI lacks the ability to adjust fees based on preexisting health conditions, resulting in the absence of tools to prevent or mitigate adverse selection.

Achieving China's goal of universal PHI will be challenging without addressing adverse selection among freelancers, self-employed individuals, unemployed persons, and students. According to this paper's research findings, we are of the opinion that implementing the following institutional design in China's PHI will mitigate the adverse effects of adverse selection on efficiency:

Introducing risk-adjusted premiums. The PHI insurers in China should establish risk adjustment factors to apply varying premiums based on individuals' health risks, such as their hospitalization status during the previous contributory year and age. The findings of this study indicate that adverse selection occurs when individuals make inferences about the potential future benefits associated with enrolling in PHI based on their prior private health risk information. In addition, the findings of this study indicate that younger age groups exhibit a relatively higher odds for experiencing adverse selection, thereby providing empirical support for the policy recommendations. In practical operation, it is quite difficult to set a premium for each specific age. Therefore, we propose to divide ages into several intervals and establish a corresponding premium standard for each interval. Our suggestion is highly practical. Firstly, there is a certain technical practicality in introducing risk-adjusted premiums. In China's PHI system, where the insurer is managed by the government, they can legally and reasonably obtain individuals' age and hospitalization information to adjust premiums accordingly. Secondly, introducing risk-adjusted premiums also holds economic practicality. Adverse selection can impede the effective functioning of insurance,[1,5,6] as individuals have more information about their own hospitalization risks. The larger the gap between an individual's self-assessed hospitalization costs and the premium, the stronger the economic incentive for them to engage in adverse selection when it comes to participating in PHI. Introducing risk-adjusted premiums can help bridge this gap, thereby improving economic efficiency.

Associating the act of enrolling in PHI with people's long-term gains — If URBMI stipulates that individuals will be exempted from paying premiums once they have accumulated certain years of participation and have reached the legal retirement age, this policy design would effectively transform the decision to enroll in PHI into long-term gains. Consequently, it would mitigate people's adverse selection behavior driven solely by short-term gains.

## Conclusion

This study utilizes 2020 wave CFPS data to examine the presence of adverse selection in PHI within China, employing individuals comprising freelancers, self-employed individuals, unemployed persons, and students. The findings yield several noteworthy conclusions. First, the adverse selection in China's PHI is evident as the odds of participating in PHI is significantly higher for individuals with hospitalization compared to those without such experience. Second, adverse selection behavior is more noticeable among younger individuals and those with better self-rated health. Our empirical findings indicate a significant negative association between the interaction term and enrollment in PHI. Third, the occurrence of adverse selection is attributed to individuals' enrollment decisions being influenced by their preexisting health conditions. The health status prior to enrollment exhibits a strong correlation with the subsequent health status. Utilizing previous private health risk information to project anticipated medical expenses enables adverse selection. This paper proposes a solution for mitigating adverse selection effects, which is that PHI should introduce risk-adjusted premiums. Simultaneously, the decision to participate in PHI should align with individuals' long-term interests.

## Abbreviations

PHI, Public Health Insurance; CFPS, China Family Panel Studies; UEBMI, Urban Employee Basic Medical Insurance; URBMI, Urban–Rural Resident Basic Medical Insurance; SD, standard deviation; ROC, Receiver Operating Characteristic; FPR, False Positive Rate; TPR, True Positive Rate; AUC, Area Under the Curve; UEBOI, Urban Employee Basic Old-age Insurance; ERAT, Experiment on Risk Aversion Tendency.

## Data Sharing Statement

The data that support the findings of this study are available from: http://www.isss.pku.edu.cn

## Ethics Approval and Consent to Participate

This study utilizes publicly available data from the China Family Panel Studies (CFPS), which involved human participants. The original data collection process, including participant recruitment and survey administration, was reviewed and approved by the Peking University Biomedical Ethics Review Committee (approval number: IRB00001052-14010), and informed consent was obtained from all participants at that stage.

For the current analysis of these de-identified, publicly accessible data, ethical approval was exempted in accordance with Article 32 (Items 1 & 2) of China's Measures for Ethical Review of Life Science and Medical Research Involving Human Subjects (2023), which stipulates that secondary use of anonymized public datasets does not require additional ethical review.

## Acknowledgments

We thank the Institute of Social Science Survey of Peking University for providing the original data used in this study.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This study was funded by the ECNU Academic Innovation Promotion Program for Excellent Doctoral Students (grant number: YBNLTS2024-025).


## Disclosure

The authors report no conflicts of interest in this work.

## References


1. Akerlof GA. The market for "lemons": quality uncertainty and the market mechanism. *Q J Econ*. 1970;84:488–500. doi:10.1016/B978-0-12-214850-7.50022-X
2. Li L, Fu HQ. China's health care system reform: progress and prospects. *Int J Health Plann Manage*. 2017;32:240–253. doi:10.1002/hpm.2424
3. National Healthcare Security Administration. Statistical Bulletin on the Development of National Medical Security. 2020. Available from: http://www.nhsa.gov.cn/art/2021/6/8/art_7_5232.html. Accessed December 9th, 2024.
4. National Bureau of Statistics. Yearbook of Statistics. 2021. Available from: https://www.stats.gov.cn/sj/ndsj/2021/indexch.htm. Accessed December 9th, 2024.
5. Spence M. Job market signaling. *Q J Econ*. 1973;87:355–374. doi:10.2307/1882010
6. Stiglitz JE. The theory of "screening," education, and the distribution of income. *Am Econ Rev*. 1975;65:283–300.
7. Wang HB, Gong X. Adverse selection and health insurance decisions of young migrant workers: an empirical study in China. *Front Public Health*. 2023;11:1084133. doi:10.3389/fpubh.2023.1084133
8. Chen BZ, Feng FY, Powers MR, et al. Risk-revealing contracts for government-sponsored microinsurance. *Pac-Basin Financ J*. 2019;57:101199. doi:10.1016/j.pacfin.2019.101199
9. Rothschild M, Stiglitz J. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *Q J Econ*. 1976;90:629–649. doi:10.2307/1885326
10. Chiappori PA, Salanié B. Testing for asymmetric information in insurance markets. *J Polit Econ*. 2000;108:56–78. doi:10.1086/262111
11. Cohen A, Siegelman P. Testing for adverse selection in insurance markets. *J Risk Insur*. 2010;77:39–84. doi:10.1111/j.1539-6975.2009.01337.x
12. Powell D, Goldman D. Disentangling moral hazard and adverse selection in private health insurance. *J Econom*. 2021;222:141–160. doi:10.1016/j.jeconom.2020.07.030
13. Konetzka RT, He DF, Dong J, et al. Moral hazard and long-term care insurance. *Geneva Pap Risk Insur-Issues Pract*. 2019;44:231–251. doi:10.1057/s41288-018-00119-1
14. Syafrawati S, Machmud R, Aljunid SM, et al. Incidence of moral hazards among health care providers in the implementation of social health insurance toward universal health coverage: evidence from rural province hospitals in Indonesia. *Front Public Health*. 2023;11:1147709. doi:10.3389/fpubh.2023.1147709
15. Nguyen L, Worthington AC. Moral hazard in Australian private health insurance: the case of dental care services and extras cover. *Geneva Pap Risk Insur-Issues Pract*. 2023;48:157–176. doi:10.1057/s41288-021-00245-3
16. Li Y, Li L, Liu JX. The efficient moral hazard effect of health insurance: evidence from the consolidation of urban and rural resident health insurance in China. *Soc Sci Med*. 2023;324:115884. doi:10.1016/j.socscimed.2023.115884
17. Ko H. Moral hazard effects of supplemental private health insurance in Korea. *Soc Sci Med*. 2020;265:113325. doi:10.1016/j.socscimed.2020.113325
18. Keane M, Stavrunova O. Adverse selection, moral hazard and the demand for medigap insurance. *J Econom*. 2016;190:62–78. doi:10.1016/j.jeconom.2015.08.002
19. Bardey D, Buitrago G. Supplemental health insurance in the Colombian managed care system: adverse or advantageous selection? *J Health Econ*. 2017;56:317–329. doi:10.1016/j.jhealeco.2017.02.008
20. Jiang YW, Ni WY. Risk selection into supplemental private health insurance in China. *Health Econ Rev*. 2019;9:1–11. doi:10.1186/s13561-019-0252-8
21. Olivella P, Vera-Hernández M. Testing for asymmetric information in private health insurance. *Econ J*. 2013;123:96–130. doi:10.1111/j.1468-0297.2012.02520.x
22. Parmar D, Souares A, De Allegri M, et al. Adverse selection in a community-based health insurance scheme in rural Africa: implications for introducing targeted subsidies. *BMC Health Serv Res*. 2012;12:1–8. doi:10.1186/1472-6963-12-181
23. von der Schulenburg JMG G, Uber A. Current issues in German healthcare. *Pharmacoeconomics*. 1997;12:517–523. doi:10.2165/00019053-199712050-00002
24. Castano R, Zambrano A. Biased selection within the social health insurance market in Colombia. *Health Policy*. 2006;79:313–324. doi:10.1016/j.healthpol.2006.01.010
25. Sapelli C, Torche A. The mandatory health insurance system in Chile: explaining the choice between public and private insurance. *Int J Health Care Financ Econ*. 2001;1:97–110. doi:10.1023/A:1012886810415
26. Pardo C. Health care reform, adverse selection and health insurance choice. *J Health Econ*. 2019;67:102221. doi:10.1016/j.jhealeco.2019.07.001
27. Gao Q, Yang S, Li S. Labor contracts and social insurance participation among migrant workers in China. *China Econ Rev*. 2012;23:1195–1205. doi:10.1016/j.chieco.2012.09.002
28. Wang FX, Gan B, Cheng YY, et al. China's employment contract law: does it deliver employment security? *Econ Labour Relat Rev*. 2019;30:99–119. doi:10.1177/1035304619827758
29. Sychenko E, Laruccia M, Cusciano D, et al. Non-standard employment in the BRICS countries. *BRICS Law J*. 2020;7:4–44. doi:10.21684/2412-2343-2020-7-4-4-44

30. Wang T, Li SQ, Gao D. What factors have an impact on the employment quality of platform-based flexible workers? an evidence from China. *Heliyon*. 2024;10(e24654). doi:10.1016/j.heliyon.2024.e24654
31. Li Y, Zhang X, Yang X. Analysis of economic effect and mechanism of basic medical insurance for urban employees in China. *Discrete Dyn Nat Soc*. 2021;1:1–9. doi:10.1155/2021/8423836
32. Li QS, Zhang LY, Jian WY. The impact of integrated urban and rural resident basic medical insurance on health service equity: evidence from China. *Front Public Health*. 2023;11:1106166. doi:10.3389/fpubh.2023.1106166
33. Chen R, Li NX, Liu X. Study on the equity of medical services utilization for elderly enrolled in different basic social medical insurance systems in an underdeveloped city of Southwest China. *Int J Equity Health*. 2018;17:1–8. doi:10.1186/s12939-018-0765-5
34. Gao J, Tang SL, Tolhurst R, et al. Changing access to health services in urban China: implications for equity. *Health Policy Plan*. 2001;16:302–312. doi:10.1093/heapol/16.3.302
35. Xie Y, Lu P. The sampling design of the China family panel studies (CFPS). *Chin J Sociol*. 2015;1:471–484. doi:10.1177/2057150X15614535
36. Xie Y, Hu JW. An introduction to the China family panel studies (CFPS). *Chin Sociol Rev*. 2014;47:3–29. doi:10.2753/CSA2162-0555470101
37. Desmond KA, Rice T, Fox PD. Does greater Medicare HMO enrollment cause adverse selection into Medigap? *Health Econ Policy Law*. 2006;1:3–21. doi:10.1017/S1744133105001039
38. Fang HM, Keane MP, Silverman D. Sources of advantageous selection: evidence from the Medigap insurance market. *J Polit Econ*. 2008;116:303–350. doi:10.1086/587623
39. de Oliveira ACM. When risky decisions generate externalities. *J Risk Uncertain*. 2021;63:59–79. doi:10.1007/s11166-021-09357-6
40. Parada-Contzen MV. The value of a statistical life for risk-averse and risk-seeking individuals. *Risk Anal*. 2019;39:2369–2390. doi:10.1111/risa.13329
41. Anderson LR, Mellor JM. Predicting health behaviors with an experimental measure of risk preference. *J Health Econ*. 2008;27:1260–1274. doi:10.1016/j.jhealeco.2008.05.011
42. Bessey D. Loss aversion and health behaviors: results from two incentivized economic experiments. *Healthcare*. 2021;9:1040. doi:10.3390/healthcare9081040
43. Brown PH, Theoharides C. Health-seeking behavior and hospital choice in China's new cooperative medical system. *Health Econ*. 2009;18(S2): S47–S64. doi:10.1002/hec.1508
44. Lönnqvist JE, Verkasalo M, Walkowitz G, et al. Measuring individual risk attitudes in the lab: task or ask? an empirical comparison. *J Econ Behav Organ*. 2015;119:254–266. doi:10.1016/j.jebo.2015.08.003
45. Hayes AF. Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Commun Monogr*. 2009;76:408–420. doi:10.1080/03637750903310360
46. Einav L, Finkelstein A. Selection in insurance markets: theory and empirics in pictures. *J Econ Perspect*. 2011;25:115–138. doi:10.1257/jep.25.1.115
47. Bonem EM, Ellsworth PC, Gonzalez R. Age differences in risk: perceptions, intentions and domains. *J Behav Decis Mak*. 2015;28:317–330. doi:10.1002/bdm.1848
48. Hemenway D. Propitious selection in insurance. *J Risk Uncertain*. 1992;5:247–251. doi:10.1007/BF00057881
49. Hemenway D. Propitious selection. *Q J Econ*. 1990;105:1063–1069. doi:10.2307/2937886