Economic Society and Humanities Vol. 3 No. 2, 2026

Academic Education Publishing House

Disease, Health Insurance, and Poverty Reversion in Rural China

Yongyi Li
Department of Economic Statistics, Northwest Normal University, Lanzhou, Gansu, China

Abstract: Using data from the 2018 China Household Income Project (CHIP), this study constructs an identification strategy based on a Recursive Bivariate Probit model to estimate the poverty-inducing effect of illness and the poverty-reducing effect of medical insurance. The results show that, without accounting for medical insurance reimbursement, households experiencing an illness shock have a 0.0824 higher probability of falling into dynamic poverty compared to a counterfactual state without such a shock. When medical reimbursement is included, this elevated probability is 0.0639. The compensation effect of medical reimbursement for illness-induced poverty is thus 22.45%. Furthermore, the effects of illness on poverty and of insurance in reducing poverty exhibit regional heterogeneity. Under comparable conditions, households in western China experience a relatively higher poverty-inducing effect from illness and a lower compensation ratio from medical insurance.

Keywords: Dynamic Return to Poverty; Return to Poverty Caused by Disease; Medical Insurance; Recursive Bivariate Probit Model

1. Introduction

On the morning of February 25, 2021, China declared comprehensive victory in its poverty eradication campaign, marking the end of absolute poverty nationwide. However, the elimination of absolute poverty does not signify the disappearance of poverty altogether, with the risk of returning to poverty among vulnerable groups demanding continued attention. As early as March 20, 2020, authorities established a monitoring and assistance mechanism to prevent relapse into poverty. The monitoring scope explicitly includes households whose per capita disposable income falls below approximately 1.5 times the national poverty line, with a focus on those experiencing "a significant increase in rigid expenditures or a sharp decline in income due to illness, disability, disasters, or the COVID-19 pandemic." Consequently, poverty induced or exacerbated by illness has become a critical basis for policy formulation.

Illness-induced poverty is also a prominent topic in academic research. Health shocks are widely recognized as a primary driver of poverty and relapse into poverty for rural households in China [1-3], exerting a far greater impact than other factors [4]. The deprivation of health undermines "capabilities," ultimately reducing patients' future income—an effect particularly pronounced among low-income groups [5]. More concerningly, low-income individuals may forgo treatment due to unaffordable costs, worsening their health and creating a vicious cycle between illness and poverty [6,7], trapping them in a "poverty-disease" trap [8,9]. From a societal perspective, illness has become a significant contributor to widening income disparities among households [10].

Regarding interventions, health insurance is undeniably effective in mitigating illness-induced poverty [11-14]. Public health insurance programs demonstrate superior targeting efficacy. For instance, the New Rural Cooperative Medical Scheme (NRCMS) enhances the health human capital of rural residents, thereby increasing labor force participation and working hours [15] and strengthening the capacity of low-income rural households, particularly middle- and low-income groups, to prevent health-related poverty [16].

Discussions on health insurance mechanisms inevitably address two issues. First, moral hazard arises from information asymmetry, where insurance coverage, reimbursement methods, and limits may incentivize reduced ex-ante investment in disease prevention and unhealthy behaviors, undermining insurance effectiveness [17-19]. It may also lower patients' price sensitivity, leading to overutilization of medical services [20,21]. Second, regarding redistributive effects, while China's health insurance system has alleviated income inequality [22], the potential for "the poor subsidizing the rich" cannot be ruled out, potentially harming certain groups and exacerbating health inequities [23-25]. Some studies suggest a negative income redistribution effect within China's basic health insurance system, where lower-income groups with poorer health incur significantly lower medical expenditures and receive less reimbursement than higher-income groups, indicating serious benefit inequity [26]. Factors such as disease heterogeneity [27] and household registration restrictions [28] may render an overly uniform compensation mechanism disadvantageous to low-income enrollees [29].

The implication of existing research is that within the analytical framework of illness, health insurance, and poverty dynamics, when undifferentiated policies face heterogeneous needs, information asymmetry can amplify moral hazard and self-selection endogeneity, complicating policy effect identification. Furthermore, average policy treatment effects are insufficient to capture true impacts, offering limited guidance for policy refinement.

Therefore, this study constructs a representative "natural person" framework to identify the impact of varying illness severity on poverty probability. It compares the treatment effects of illness on poverty under two scenarios—with and without health insurance reimbursement—to assess insurance's role in preventing poverty. Additionally, using predicted poverty probability as a proxy for individual heterogeneity, it examines illness's poverty-inducing effects across different risk groups and whether insurance exerts a significant poverty-reduction effect.

This study contributes in two ways. First, it estimates simultaneous equations for poverty and illness to control for endogeneity using Full Information Maximum Likelihood (FIML). By comparing poverty probability changes with and without illness shocks, it identifies the Average Treatment Effect (ATE) and the Average Treatment Effect on the Treated (ATET) of illness, while testing insurance's role—offering a novel methodological approach. Second, it compares illness treatment effects on poverty across samples with different poverty probabilities under reimbursement and non-reimbursement scenarios, providing more relevant insights for designing heterogeneous policies.

2. Research Design

2.1 Model Specification

Building upon the established identification framework [30], both the poverty risk and the illness risk faced by a household are treated as latent variables determined by household characteristics. Under this premise, this paper jointly estimates the poverty risk equation and the illness risk equation using a recursive bivariate Probit model with endogeneity—henceforth abbreviated as the Rec-Biprobit model—to identify the treatment effect of illness on a household's risk of falling into poverty. The model is specified as follows:

$$Y_i^* = \beta T_i + \alpha X_i + \pi Z_i + \varepsilon_i, Y_i = 1[Y_i^* > 0] \quad (1)$$

$$T_i^* = \gamma X_i + \theta X_i^{'} + u_i, T_i = 1[T_i^* > 0] \quad (2)$$

Equation (1) is the poverty risk equation of the sample. Where, Y_i^* is the latent variable representing the poverty risk faced by sample i. If $Y_i^* > 0$ and $Y_i = 1$ sample i is identified as poor, otherwise it is non-poor. X_i represents the characteristics of the family and the head of the household where sample i is located. T_i indicates whether any member of the family where sample i is located suffers a disease shock. If any member of the family is affected by the disease shock, all members of the whole family take $T_i = 1$, otherwise $T_i = 0$. Z_i is other control variables, ε_i is the random disturbance term. Equation (2) is the sample disease risk equation, T_i^* s the latent variable representing the disease risk faced by sample i. If $T_i^* > 0$ and $T_i = 1$, sample i is identified as being affected by disease shock, otherwise it is not affected. $X_i^{'}$ is an exogenous variable, u_i is the random disturbance term.

Equations (1)-(2) show that if the household's exposure to disease shock is strictly exogenous, and $E(\varepsilon_i|T_i^*) = 0$, the random disturbance term ε_i and u_i are independent. However, the impact of disease on poverty is not unidirectional [7-9]. Family poverty risk and disease risk not only have a reverse causal relationship, but also have simultaneity, so the hypothesis that $Cov(\varepsilon_i, u_i) = 0$ cannot be ruled out. Assume that ε_i and u_i follow a bivariate normal distribution:

$$\begin{pmatrix}\varepsilon_i\\ u_i\end{pmatrix} \sim N_2\left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right] \quad (3)$$

The parameter ρ is the correlation coefficient between the random disturbance term ε_i and u_i. According to Equation (3), construct the joint log-likelihood function of Equations (1)-(2):

$$lnL = \sum_{i|Y=1,T=1} ln\,\emptyset_2(\gamma X_i + \theta X_i^{'}, \beta + \alpha X_i + \pi Z_i, \rho) + \sum_{i|Y=1,T=0} ln\,\emptyset_2(-\gamma X_i - \theta X_i^{'}, \alpha X_i + \pi Z_i, -\rho) + \sum_{i|Y=0,T=1} ln\,\emptyset_2(\gamma X_i + \theta X_i^{'}, -\beta - \alpha X_i - \pi Z_i, -\rho) + \sum_{i|Y=0,T=0} ln\,\emptyset_2(-\gamma X_i - \theta X_i^{'}, -\alpha X_i - \pi Z_i, \rho) \quad (4)$$

$\emptyset_2(\cdot)$ is the probability distribution function of the two−dimensional normal distribution. Full information maximum likelihood estimation is used to estimate the parameters of the simultaneous equations, and we obtain Y_i and T_i the two−dimensional joint Probit distribution function $Prob(Y_i = 1, T_i = 1) = \emptyset_2(T_i^*, Y_i^*, \rho)$.

2.2 Partial Effects of Variables

Within the framework of the nonlinear latent variable model, parameter estimates cannot represent the influence degree of covariates on the latent output variable. According to equations (1) to (2), the latent output variable of poverty risk Y_i has a conditional expectation as follows:

$$E\{Y_i|X_i, Z_i, T_i\} = \emptyset(\alpha X_i + \beta T_i + \pi Z_i) \quad (5)$$

$\emptyset(\cdot)$ is the cumulative probability distribution function of the standard normal distribution. Equation (5) shows that household characteristics X_i directly affect Y_i , but indirectly affect Y_i through Equation (2), then the conditional expectation of the potential output variable Y_i can be further written as:

$$E\{Y_i|X_i, Z_i, X_i^{'}\} = E_T\{E\{Y_i|X_i, Z_i, T_i\}\} = \emptyset(-\gamma X_i - \theta X_i^{'})\emptyset(\alpha X_i + \pi Z_i) + \emptyset(\gamma X_i + \theta X_i^{'})\emptyset(\alpha X_i + \pi Z_i + \beta) \quad (6)$$

The partial effect of a variable can be composed of the direct partial effect and the indirect partial effect, where the household covariate X_i the direct partial effect is written as:

$$\frac{\partial E\{Y_i|X_i, Z_i, X_i^{'}\}}{\partial X_i} = \left\{ \begin{array}{l} \emptyset(-\gamma X_i - \theta X_i^{'}) f(\alpha X_i + \pi Z_i) \\ +\emptyset(\gamma X_i + \theta X_i^{'}) f(\alpha X_i + \pi Z_i + \beta) \end{array} \right\} \times \alpha \quad (7)$$

Control variable Z_i the direct partial effect is written as:

$$\frac{\partial E\{Y_i|X_i, Z_i, X_i^{'}\}}{\partial Z_i} = \{\emptyset(-\gamma X_i - \theta X_i^{'}) f(\alpha X_i + \pi Z_i) + \emptyset(\gamma X_i + \theta X_i^{'}) f(\alpha X_i + \pi Z_i + \beta)\} \times \pi \quad (8)$$

Family covariate X_i the indirect partial effect is:

$$\frac{\partial E\{Y_i|X_i, Z_i, X_i^{'}\}}{\partial X_i} = f(\gamma X_i + \theta X_i^{'})[\emptyset(\alpha X_i + \pi Z_i + \beta) + \emptyset(\alpha X_i + \pi Z_i)] \times \gamma \quad (9)$$

Exogenous variable $X_i^{'}$ indirect effect is:

$$\frac{\partial E\{Y_i|X_i, Z_i, X_i^{'}\}}{\partial X_i^{'}} = f(\gamma X_i + \theta X_i^{'})[\emptyset(\alpha X_i + \pi Z_i + \beta) + \emptyset(\alpha X_i + \pi Z_i)] \times \theta \quad (10)$$

$f(\cdot)$ is the probability density function of standard normal distribution. The partial effects at the mean of each variable will be given separately below.

2.3 Identification and Statistical Inference of Health Shock-Induced Poverty Effects

In this paper, the disease proxy variable T_i is a binary treatment variable. Then the treatment effect of T_i on Y_i is:

$$TE_i = Prob(Y_i = 1)_{T_i=1} - Prob(Y_i = 1)_{T_i=0} = \emptyset(\alpha X_i + \pi Z_i + \beta) - \emptyset(\alpha X_i + \pi Z_i) \quad (11)$$

The treatment effect is defined as the difference in the probability of being identified as poor between households experiencing health shocks and their counterfactual scenario without such shocks, expressed as:

$$TET_i = Prob(Y_i = 1|T_i = 1)_{T_i=1} - Prob(Y_i = 1|T_i = 1)_{T_i=0} = \{\emptyset_2(\gamma X_i + \theta X_i^{'}, \beta + \alpha X_i + \pi Z_i, \rho) - \emptyset_2(\gamma X_i + \theta X_i^{'}, \alpha X_i + \pi Z_i, \rho)\} / \emptyset(\gamma X_i + \theta X_i^{'}) \quad (12)$$

When variables are set to their sample means, the Average Treatment Effect (ATE) and the Treatment Effect on the Treated (TET) are obtained. This study applies the delta method to conduct statistical inference and construct confidence intervals. The procedure is as follows:

For each observation i, the identified effect TET_i is a function of the estimated parameter vector $\widehat{\theta}$ from Equations (1)–(2) and the individual's characteristics C_i（including X_i， $X_i^{'}$ and Z_i）. Let D denote the gradient vector of TET_i with respect to the parameter estimates $\widehat{\theta}$:

$$G = \left\{ \frac{\partial TET_i(\theta, C_i)}{\partial \theta} \bigg|_{\theta=\widehat{\theta}} \right\} \quad (13)$$

The variance of the effect TET_i is:

$$\widehat{Var}[TET_i(\theta, C_i)] = G \Sigma_{\widehat{\theta}} G^{'} \quad (14)$$

Where $\Sigma_{\widehat{\theta}}$ denotes the variance-covariance matrix of the coefficient estimates $\widehat{\theta}$. The z-statistic and confidence interval for its effect are then given by:

$$z_i = \frac{TET_i(\theta, C_i)}{\sqrt{G \Sigma_{\widehat{\theta}} G^{'}}} \quad (15)$$

$$TET_i(\theta, C_i) \pm z_{\alpha/2} \sqrt{G \Sigma_{\widehat{\theta}} G^{'}} \quad (16)$$

2.4 Testing the Existence of a Poverty Reduction Effect from Medical Insurance

As part of household transfer income, reimbursement of medical expenses can alleviate poverty caused by illness to some extent. If the rural medical insurance system is well-established and effective, medical reimbursements should further reduce the probability of falling into poverty due to disease. Based on this reasoning, we define: Y_i^b as the poverty status when income excludes medical reimbursement, Y_i^a as the poverty status when income includes medical reimbursement. The poverty status Y_i^a should be no worse than Y_i^b. By comparing the treatment effect of illness on these two poverty measures, we can identify the poverty reduction effect of medical insurance.

If the treatment effect of illness on poverty—and the associated treatment-on-the-treated effect—under the scenario where household income excludes medical reimbursement is significantly larger than the corresponding average treatment effect under the counterfactual scenario where income includes reimbursement, then medical insurance can be considered effective in mitigating medical poverty in rural areas. Following this identification strategy, the hypothesis is as follows:

$$H_0: \overline{TE}_b = \overline{TE}_a, H_1: \overline{TE}_b \neq \overline{TE}_a \quad (17)$$

$$H_0: \overline{TET}_b = \overline{TET}_a, H_1: \overline{TET}_b \neq \overline{TET}_a \quad (18)$$

Here, $\overline{TE}_b$ and $\overline{TET}_b$ represent the average treatment effect and the average treatment-on-the-treated effect of illness under the income-excluding-medical-reimbursement scenario, while $\overline{TE}_a$ and $\overline{TET}_a$ denote the corresponding effects under the income-including-medical-reimbursement scenario. This study estimates Equations (1)–(2) under both scenarios, calculates the effects using Equations (11)–(12), and employs the Bootstrap method with 200 replications to obtain the Chi-square statistic and p-value for testing the above hypothesis.

Table 1. Descriptive Statistics and Group Difference Tests for Household and Household Head Variables

	Variable	Full Sample	Poverty	Non-poverty	t-statistic	p-value
Household Head	Age (years)	54.0506	54.1991	54.0387	-0.67	0.504
	Age squared / 100	30.459	30.7299	30.4372	-1.10	0.272
	Gender (%)	92.68	92.15	92.73	1.05	0.295
	Marital Status (%)	93.61	90.44	93.86	5.69	0.000
	Years of Education (years)	7.3428	6.8728	7.38	9.24	0.000
	Cadre Status (%)	10.41	9.36	10.49	1.86	0.063
Household	Average Age (years)	40.38	39.05	40.49	5.87	0.000
	Average Age Squared	18.74	16.56	18.91	5.78	0.000
	Household Size (persons)	4.40	4.80	4.37	-13.30	0.000
	Number of Children (persons)	0.88	1.01	0.87	-7.04	0.000
	Number of Elderly (persons)	0.77	0.76	0.77	0.47	0.641
	Number of Working Members (persons)	2.50	2.71	2.49	-8.24	0.000
	Years of Education of Labor Force (years)	5.81	5.76	5.82	0.98	0.328
	Per Capita Income (yuan)	12669.13	1800.44	13539.13	181.23	0.000
	Per Capita Medical Reimbursement (yuan)	217.10	355.27	206.04	-1.22	0.222
Living Standard	Life Dissatisfaction (%)	3.87	5.28	3.76	-3.31	0.001
	Expected Income Decline (%)	4.69	4.24	4.73	1.15	0.249

3. Data and Variables

3.1 Data Sources and Variable Characteristics

This study utilizes data from the rural household survey of the China Household Income Project (CHIP 2018). The sample is a subsample of the National Bureau of Statistics survey, ensuring strong representativeness and reliability. After data cleaning to remove households with missing key variables, the final sample consists of 8,825 households and 33,088 individuals. The main variables are constructed as follows: (1) Potential poverty dummy variable. Using the 2018 poverty line of 2,995 yuan per capita, households are classified as potentially poor (coded as 1) or non-poor (coded as 0). (2) Health shock dummy variable. This is based on the reported number of days household members were unable to work or conduct normal activities due to illness/injury. A primary threshold of 60 days (the 95th percentile) is used to define a major health shock; a household is coded as 1 if any member meets or exceeds this threshold. For robustness checks, alternative thresholds of 40, 80, 100, and 12 days are also employed. (3) Household and household head characteristics. The analysis also incorporates additional control variables, including characteristics of the household head (age, gender, marital status, and education level) as well as household characteristics (average age and household size). Basic descriptive statistics for these variables are presented in Table 1. (4) Model identification variables. To aid model identification, two variables related to living standards are defined: a happiness dummy (1 if the respondent reports being "not very happy," "very unhappy," or "uncertain about happiness"; 0 otherwise) and an income expectation dummy (1 if the respondent expects household income to remain unchanged or decrease over the next five years; 0 otherwise).Comparative analysis reveals statistically significant differences in most variables between poor and non-poor households.

3.2 Basic Characteristics of Poverty

Table 2 compares the poverty incidence under two scenarios: income excluding medical reimbursement expenses and income including medical reimbursement expenses. The comparison finds that medical reimbursement can reduce the poverty incidence of the sample, and the magnitude of the reduction is not sensitive to changes in the poverty line. In 2018, the poverty line was 2995 yuan. If medical reimbursement expenses are deducted from the sample income, the poverty incidence is 7.41%; when income includes medical reimbursement expenses, the poverty incidence drops by 0.54% to 6.87%. If the poverty line is raised to 3215 yuan in 2019, the poverty incidence under the two scenarios of income excluding medical reimbursement and income including medical reimbursement are 8.30% and 7.72% respectively, with a difference of 0.58%. After lowering the poverty line, the poverty incidence under both scenarios also decreases, and the difference between the two scenarios still changes little. It can be considered that medical reimbursement is relatively uniform in terms of sample coverage and reimbursement amount, and the medical security system is more latively exogenous in institutional design.

Table 2. Medical Reimbursement and Poverty Incidence

Year	Poverty line	Income after medical reimbursement deduction	Income including medical reimbursement
2018	2995 yuan	7.41%	6.87%
2019	3216 yuan	8.30%	7.72%
2016/2017	2952 yuan	7.24%	6.72%
2015	2855 yuan	6.93%	6.38%

4. Empirical Study

4.1 Parameter Estimation

This paper examines how household illness, defined as a family member being unable to work for at least 60 days, affects the likelihood of falling into poverty. Two income measures are compared: income excluding medical reimbursement and income including medical reimbursement. After controlling for household and household-head characteristics, Probit, IV-Probit, and Recursive Bivariate Probit models are estimated.

The simple Probit results, presented in Table 3, show a positive but statistically insignificant effect of illness with a coefficient of 0.046 when reimbursement is excluded. When reimbursement is included, the effect becomes negative at negative 0.029, though still insignificant. This suggests medical reimbursement may offset poverty induced by illness, but these estimates are likely biased by endogeneity. The endogeneity arises because individuals face a trade-off between continuing to work while ill to avoid immediate income loss, which risks long-term health, and seeking treatment to protect health, which may reduce current income and increase poverty risk.

Using self-reported happiness and future income expectations as instrumental variables, the IV-Probit model yields larger and more significant coefficients: 0.876 for income excluding reimbursement and 0.682 for income including it. This indicates that illness significantly increases the risk of poverty and that medical reimbursement mitigates this effect. However, these instrumental variables may not be fully exogenous.

The Recursive Bivariate Probit model addresses endogeneity by allowing the error terms of the poverty and illness equations to correlate. It produces statistically significant estimates at the 1 percent level: a coefficient of 0.928 when reimbursement is excluded and 0.813 when it is included. These results confirm two key findings: first, illness raises the probability of entering poverty, and second, medical reimbursement alleviates this impact. The coefficients from the Recursive Bivariate Probit model are larger than those from both the Probit and IV-Probit models, indicating that the simpler models underestimate the true effect of illness on poverty.

Table 3. Parameter Estimation Results

	Probit Model		IV-Probit Model		Rec-Biprobit Model			
	M1	M2	M3	M4	M5		M6	
	State 1	State 2	State 1	State 2	State 1		State 2	
Disease	0.046	-0.029	0.876***	0.682**	0.928***		0.813***	
	(1.61)	(-0.96)	(2.80)	(2.02)	(5.30)		(4.07)	
rho	—	—	—	—	-0.463***		-0.440***	
	—	—	—	—	(-5.56)		(-4.65)	
Constant Term	-1.067***	-1.383***	-0.869***	-1.211***	-1.006***	-2.217***	-1.316***	-2.220***
	(-5.00)	(-6.29)	(-3.83)	(-5.08)	(-4.89)	(-21.41)	(-6.17)	(-21.41)
N	33088	33088	33088	33088	33088		33088	
chi2²	678	597	767	642	1722		1719	

Table 4 results indicate that household characteristics significantly affect poverty risk. Larger household size significantly increases the probability of falling into poverty, while having a married household head is the strongest poverty-reducing factor. Interestingly, a higher number of elderly members in the household reduces poverty risk, likely due to transfers like pensions or family support. Furthermore, the instruments for identification—self-reported unhappiness and expected income decline—both show significant positive direct effects on poverty probability, aligning with theoretical expectations. These partial effect estimates provide a reliable parametric basis for the subsequent calculation of counterfactual poverty probabilities under health shock scenarios.

4.2 Estimation of Poverty-Inducing Effect of Illness and Test of Poverty Reduction Effect of Medical Insurance

Table 5 presents the estimated Average Treatment Effects (ATE) of illness on poverty using different econometric models. When household income excludes medical insurance reimbursement (State 1), the Probit model yields an ATE of 0.0077 and an ATET of -0.0352, which are statistically insignificant. When income includes reimbursement (State 2), the ATE is -0.0044 and the ATET is -0.0036, also insignificant. However, the differences in effects between the two income states are significant.

The IV-Probit model, using respondents' current happiness and future income expectations as instrumental variables for illness, estimates the illness ATE on poverty at 0.2184 in State 1 and 0.1491 in State 2, with a significant difference. The corresponding ATET values are 0.1295 and 0.1182, respectively, though their difference is not significant.

After controlling for the correlation between illness and poverty through unobservable factors using the Recursive Bivariate Probit (Rec-Biprobit) model, the illness ATE and ATET on poverty are 0.2335 and 0.0824, respectively, in State 1 (without reimbursement), both significant at least at the 5% level. In State 2 (with reimbursement), these values decrease to 0.1868 and 0.0639, significant at the 1% level. The differences in ATE (0.0467) and ATET (0.0185) between the two income states are also significant at least at the 5% level.

The results show that for an average household, if an adult member suffers from an illness lasting over 60 days and is unable to work, the probability of falling into poverty increases by 0.2335 without medical reimbursement, compared to a non-ill household. With reimbursement, this increase is reduced to 0.1868. In other words, medical reimbursement reduces the poverty probability induced by illness by 0.0467, a decrease of approximately 20%. For households that already have a patient—a more policy-relevant group—the probability of poverty is 0.0824 higher than the counterfactual non-ill scenario without reimbursement. With reimbursement, this value drops to 0.0639, a reduction of 0.0185, implying a compensation effect of 22.45%.

Table 4. Estimation of Partial Effects of Parameters in the Recursive Biprobit Model

		Status 1			Status 2		
		Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect	Total Effect
Household	Average Age	0.0012***	0.0007***	0.0018	0.0015***	0.0005***	0.0020
		(3.20)	(3.25)		(4.05)	(2.64)	
	Age Squared	-0.0003**	-0.0001***	-0.0004	-0.0003**	-0.0001***	-0.0004
		(-2.22)	(-3.15)		(-2.17)	(-2.58)	
	Family Size	0.0121***	0.0111***	0.0232	0.0118***	0.0090***	0.0208
		(4.88)	(3.43)		(4.79)	(2.72)	
	Number of Children in the Family	0.0041	-0.0079***	-0.0038	0.0066*	-0.0065***	0.0001
		(1.09)	(-3.32)		(1.76)	(-2.66)	
	Number of Elderly in the Family	-0.0220***	0.0002	-0.0218	-0.0211***	0.0001	-0.0210
		(-6.86)	(0.31)		(-6.66)	(0.25)	
	Number of Working Members	0.0069***	-0.0072***	-0.0003	0.0061***	-0.0058***	0.0003
		(3.50)	(-3.41)		(3.08)	(-2.71)	
	Education of Labor Force	-0.0003	-0.0008***	-0.0011	0.0003	-0.0006***	-0.0004
		(-0.34)	(-2.69)		(0.31)	(-2.29)	
	Happiness Status		0.0146***	0.0146		0.0119***	0.0119
			(3.27)			(2.63)	
	Income Expectation		0.0118	0.0118		0.0096***	0.0096
			(3.40)			(2.76)	
Household Head	Age	-0.0045***		-0.0045	-0.0031***		-0.0031
		(-4.30)			(-2.94)		
	Age Squared	0.0042***		0.0042	0.0026***		0.0026
		(4.33)			(2.70)		
	Gender	0.0111*		0.0111	0.0091		0.0091
		(1.79)			(1.51)		
	Marital Status	-0.0390***		-0.0390	-0.0395***		-0.0395
		(-6.27)			(-6.49)		
	Years of Education	-0.0024***		-0.0024	-0.0026***		-0.0026
		(-3.44)			(-3.79)		
	Whether a Cadre	0.0035		0.0035	0.0041		0.0041
		(0.60)			(0.71)		

It is important to clarify that the ATE measures the reduction in the potential poverty probability for a typical household, influenced by all poverty-inducing factors including illness, if the illness effect were removed. The ATET for the treated describes the reduction for a typical illness-affected household if the illness effect were eliminated. For example, in State 1, the average potential poverty probability for a general household under all factors is 0.3136, which would drop to 0.0801 without illness. For a household where poverty is induced specifically by illness, the potential probability is 0.09615, which would fall to 0.01375 without illness. Consequently, both the treatment effects and their corresponding potential probabilities are higher for ATE than for ATET. The study also finds that the magnitude of reduction in both ATE and ATET due to medical reimbursement is similar, indicating that the poverty-reduction effect of health insurance does not differ substantially between the general household population and those impoverished by illness

Table 5. Estimation Results of Average Treatment Effect

Model	Effect	State 1			State 2			Medical insurance effect
		Estimated value	95% confidence interval		Estimated value	95% confidence interval		Chi2 [p-value]
Probit	ATE	0.0077	-0.0025	0.0179	-0.0044	-0.0140	0.0053	0.0121***

		(1.48)			(-0.88)			33.92[0.000]
	ATET	-0.0352	-0.0975	0.0271	-0.0036	-0.0497	0.0425	-0.0316**
		(-1.11)			(-0.15)			5.52[0.019]
IV-Probit	ATE	0.2184*	-0.0069	0.4437	0.1491	-0.0592	0.3574	0.0693**
		(1.90)			(1.40)			4.60[0.032]
	ATET	0.1295	-0.0608	0.3198	0.1182	-0.0571	0.2934	0.0113
		(1.33)			(1.32)			0.15[0.700]
Rec-Biprobit	ATE	0.2335***	0.0975	0.3696	0.1868**	0.0364	0.3373	0.0467*
		(3.36)			(2.43)			3.80[0.051]
	ATET	0.0824***	0.0612	0.1035	0.0639***	0.0424	0.0853	0.0185***
		(7.63)			(5.84)			14.90[0.000]

4.3 Heterogeneity Analysis

Table 6 presents the estimation results of effects in different regions. Regional heterogeneity shows two characteristics. First, the poverty-inducing effect of disease increases sequentially from the eastern, central to western regions. When income excludes medical reimbursement, the average treatment effect of disease-induced poverty in the eastern, central and western regions is 0.2291, 0.2340 and 0.2371 respectively, and the treatment effect on the treated is 0.0788, 0.0827 and 0.0854 respectively, all of which are statistically significant at the 1% level. The same disease has a greater poverty-inducing impact on sample households in the western region than on those in the eastern region. Second, the poverty reduction effect of medical insurance is significant, and the gap in the absolute poverty reduction effect among regions is small. For households falling into poverty due to illness, the average poverty reduction effect of medical insurance in sample households of the eastern, central and western regions is 0.0180, 0.0188 and 0.0187 respectively. However, due to the different poverty-inducing effects of disease in the three regions, the relative effect of medical insurance on alleviating disease-induced poverty shows a decreasing trend from the eastern to central and western regions. The compensation ratio of medical insurance in sample households of the eastern, central and western regions is 22.84%, 22.73% and 21.90% respectively.

Table 6. Regional Heterogeneity of Average Treatment Effect

Region	Eeffect	State 1			State 2			Medical insurance effect
		Estimated value	95% confidence interval		Estimated value	95% confidence interval		Chi2 [p-value]
Eastern	ATE	0.2291***	0.0942	0.3640	0.1826**	0.0339	0.3312	0.0465*
		(3.33)			(2.41)			3.67[0.055]
	ATET	0.0788***	0.0586	0.0989	0.0608***	0.0406	0.0809	0.0180***
		(7.66)			(5.91)			13.79[0.000]
Central	ATE	0.2340***	0.0978	0.3703	0.1870**	0.0364	0.3376	0.0470*
		(3.37)			(2.43)			3.66[0.056]
	ATET	0.0827***	0.0613	0.1041	0.0639***	0.0425	0.0854	0.0188***
		(7.58)			(5.84)			13.98[0.000]
Western	ATE	0.2371***	0.1003	0.3738	0.1906**	0.0386	0.3426	0.0465*
		(3.40)			(2.46)			3.53[0.060]
	ATET	0.0854***	0.0634	0.1073	0.0667***	0.0444	0.0890	0.0187***
		(7.63)			(5.86)			13.61[0.000]

5. Robustness Check

Based on the Rec-Biprobit model, this paper tests the robustness of the above conclusions by adjusting the poverty line to change the dependent variable and adjusting the disease severity to change the treatment variable, and the test can also verify the sensitivity of the effect to changes in poverty status and disease severity.

5.1 Poverty Line Adjustment

With reference to the adjustment range of poverty standards in poverty alleviation practices, this paper uses the poverty standards of 2855 yuan in 2015, 2952 yuan in 2016 and

2017, and 3216 yuan at the end of 2019 to replace the 2018 current standard respectively, to simulate the dynamic changes of poverty caused by illness and the poverty reduction effect of medical insurance as poverty standards change. The parameter estimates of the model are shown in the appendix. Table 7 reports the changing characteristics of the effect estimates after the poverty line adjustment. When the poverty standard increases to 3216 yuan, the treatment effect of disease on poverty and the treatment effect of medical insurance (treatment) in state 1 are 0.2306 and 0.0979 respectively, which are statistically significant at the 1% level. The treatment effect of disease and the treatment effect of medical insurance (treatment) in state 2 are 0.2074 and 0.0768 respectively, which are also statistically significant at the 1% level. When the poverty standard drops to 2952 yuan, the estimation result of the treatment effect of disease increases, while the estimation result of the treatment effect of medical insurance (treatment) decreases. When the poverty standard further drops to 2855 yuan, both the treatment effect of poverty caused by disease and the treatment effect of medical insurance (treatment) show a decreasing trend. It is also found that for general households, there is no obvious relationship between the poverty reduction effect of medical insurance and the poverty line, but for households affected by disease shocks, the poverty reduction effect of medical insurance increases with the increase of poverty standards. For example, under the poverty standard of 3216 yuan, the estimated poverty reduction effect of medical insurance is 0.0211, and the compensation proportion for poverty caused by illness is 21.55% (0.0211/0.0979). Both the absolute value of the effect estimate and the compensation degree are greater than the corresponding results under lower poverty standards.

Combined with the estimation results in Table 5, it can be concluded that if medical reimbursement is excluded from income, the treatment effect of poverty caused by illness does not necessarily increase with the rise of the poverty threshold, while the poverty-reduction effect of medical treatment (illness itself) increases with the rise of the poverty threshold. If medical reimbursement is included in income, the treatment effect of poverty caused by illness decreases with the rise of the poverty threshold, and the poverty-reduction effect of disease treatment increases with the rise of the poverty threshold. It can be seen that for ordinary families, regardless of whether the income is supplemented by medical reimbursement, disease is not necessarily the main cause of poverty for families. The higher the poverty threshold, the less important the effect of poverty caused by illness may be. In contrast, for families affected by disease shocks, disease is an important cause of poverty. The increase of the poverty threshold will amplify the effect of poverty caused by illness in such families, and meanwhile, the mitigation degree of medical reimbursement on poverty caused by illness also increases with the rise of the poverty threshold.

Table 7. Poverty Line Adjustment and Changes in Effect Estimates

Poverty Line	Effect	State 1			State 2			Medical insurance effect
		Estimated value	95% confidence interval		Estimated value	95% confidence interval		Chi2 [p-value]
3216 Yuan	ATE	0.2306***	0.0967	0.3644	0.2074***	0.0593	0.3555	0.0232
		(3.38)			(2.74)			0.86[0.354]
	ATET	0.0979***	0.0711	0.1247	0.0768***	0.0551	0.0985	0.0211***
		(7.16)			(6.94)			11.47[0.000]
2952 Yuan	ATE	0.2357***	0.1023	0.3692	0.2095***	0.0691	0.3499	0.0262
		(3.46)			(2.92)			1.30[0.254]
	ATET	0.0785***	0.0588	0.0981	0.0630***	0.0464	0.0795	0.0155***
		(7.83)			(7.47)			13.05[0.000]
2855 Yuan	ATE	0.2124***	0.0576	0.3672	0.2359***	0.0883	0.3836	-0.0235
		(2.69)			(3.13)			0.28[0.597]
	ATET	0.0761***	0.0466	0.1056	0.0609***	0.0447	0.0770	0.0152**
		(5.06)			(7.38)			1.55[0.214]

5.2 Adjustment for Disease Severity

Table 8 presents the estimation results of the treatment effect of poverty caused by illness and the poverty-reduction effect of medical insurance under four different levels of disease severity. Consistent with the above, disease severity is measured by the minimum number of days unable to work due to illness. Combined with the estimation results in Table 8, it is found that for families affected by disease shocks, regardless of whether they receive medical reimbursement, the poverty-causing effect of disease tends to increase as the disease severity increases. For example, when a disease shock is defined as being unable to work for more than 120 days due to illness, the estimated treatment effect of poverty caused by illness is 0.0882 when family income excludes medical insurance reimbursement, and 0.0684 when it includes the reimbursement, both of which are larger than the estimated values under other definitions of disease severity. It should be noted that the poverty-reduction effect of medical insurance does not increase with the aggravation of disease severity, and the level of protection for different diseases is similar. The poverty-reduction effect of medical insurance ranges from 0.0183 to 0.0204, and the compensation ratio for the poverty-causing effect of illness ranges from 22.17% to 25.12%. The implicit implication of this result is that the social welfare effect of China's current medical insurance system is similar for different levels of disease protection.

Table 8. Disease Severity and Effect Estimation Results

Disease Severity	Effect	State 1			State 2			Medical insurance effect
		Estimated value	95% confidence interval		Estimated value	95% confidence interval		Chi2 [p-value]
40 Days	ATE	0.2241***	0.1003	0.3478	0.1805***	0.0472	0.3137	0.0436**
		(3.55)			(2.65)			4.75[0.029]
	ATET	0.0818***	0.0608	0.1028	0.0635***	0.0431	0.0840	0.0183***
		(7.63)			(6.09)			16.47[0.000]
80 Days	ATE	0.1664**	0.0143	0.3184	0.1230	-0.0432	0.2892	0.0434*
		(2.14)			(1.45)			3.03[0.082]
	ATET	0.0812***	0.0404	0.1219	0.0608**	0.0082	0.1134	0.0204**
		(3.91)			(2.26)			3.90[0.048]
100 Days	ATE	0.1718***	0.0439	0.2998	0.1423**	0.0013	0.2834	0.0295
		(2.63)			(1.98)			1.27[0.260]
	ATET	0.0848***	0.0503	0.1194	0.0660***	0.0284	0.1036	0.0188**
		(4.82)			(3.44)			4.79[0.029]
120 Days	ATE	0.1603**	0.0373	0.2832	0.1324*	-0.0043	0.2691	0.0279
		(2.55)			(1.90)			0.94[0.332]
	ATET	0.0882***	0.0511	0.1254	0.0684***	0.0281	0.1088	0.0198**
		(4.65)			(3.33)			3.85[0.050]

6. Summary

Based on the data from the Chinese Household Income Project (CHIP2018), this paper constructs an identification strategy using the recursive bivariate Probit model, and attempts to incorporate the poverty-causing effect of illness and the poverty-reduction effect of medical insurance into the same framework for estimation. The main research conclusions are as follows: (1) Disease shock is an important poverty-inducing factor. Before medical insurance reimbursement, the average treatment effect of disease on poverty and the average treatment effect on the treated are 0.2335 and 0.0824 respectively. After medical insurance reimbursement, the two indicators drop to 0.1868 and 0.0639 respectively. (2) Medical insurance can effectively alleviate the poverty-causing effect of illness. Compared with that before medical insurance reimbursement, the average treatment effect of disease on poverty decreases by 0.0467 after reimbursement, with a compensation ratio of 20%; the average treatment effect on the treated decreases by 0.0185, and the compensation ratio reaches 22.45%. (3) There exists regional heterogeneity in the poverty-causing effect of disease and the poverty-reduction effect of medical insurance. The poverty-causing effect of disease increases sequentially from the eastern, central to western regions of China, while the compensation ratio of medical insurance decreases in that order. (4) Changes in poverty standards and disease severity will affect the poverty-causing effect of illness and the poverty-reduction effect of medical insurance for households hit by disease. For households suffering from disease shocks, the poverty-causing effect of disease increases with the improvement of poverty standards and the aggravation of disease severity. The poverty-reduction effect of medical insurance increases with the improvement of poverty standards, but does not increase with the aggravation of disease severity. The poverty-reduction effect varies little across different disease severities, ranging from 0.0183 to 0.0204, and the compensation ratio for the poverty-causing effect of illness ranges from 22.17% to 25.12%.

The above conclusions have threefold policy implications. First, falling into and returning to poverty due to illness is the biggest obstacle to consolidating poverty alleviation achievements. Although the current medical security system has a good targeting performance for poor households, the compensation ratio of medical insurance reimbursement for the poverty-causing effect of illness is less than 30%, which leaves a large gap for the economic recovery of households that have fallen into or returned to poverty due to illness. Raising the compensation level may be an important part of the future reform of the medical security system. Second, an absolutely fair medical insurance reimbursement system may not be conducive to improving the overall social welfare. Homogeneous disease shocks generate heterogeneous poverty-causing effects on different households, and the same reimbursement level may not be able to compensate for the welfare loss caused by disease to specific groups. Therefore, constructing a pro-poor, precise and differentiated medical security service system may be another important direction for the future reform of the medical security system. Third, regional coordination of macro policies is also an important aspect for improving the medical security system and addressing the problem of falling into and returning to poverty due to illness. The governance of falling into and returning to poverty due to illness in less developed areas has landmark and bottom-line significance. We should increase support for less developed regions through preferential policies, start from the reform of specific disease insurance mechanisms, and explore a long-term mechanism for medical insurance to reduce poverty.

References

[1] Tao J, Zhang S. Can the rural medical assistance system alleviate multidimensional poverty? —An analysis based on china household finance survey data. China Economic Studies 2025(02):132-47.
[2] Yu D, Liao X, Zhan R, Ge J. Can rural health insurance alleviate farmers' multidimensional relative poverty? —Micro-evidence from cfps data. Social Security Studies 2025(04):61-77.
[3] Tao C, Zhou H, Wu Z, Xu D. Research on the impact mechanism of basic medical insurance on the economic resilience of rural households—Based on cfps micro-survey data. Statistical Research 2025;42(10):119-30.
[4] Wang H, Wang Z, Ma P. Analysis and reflection on illness-induced poverty in rural areas—Based on survey data of 1214 illness-induced poor households in 9 western provinces and cities. Economist 2016(10):71-81.
[5] Yu X, Shangguan Y, Shen Y, Yu W. Illness-induced poverty: How do health shocks affect income levels?—Also on the poverty alleviation effect of medical insurance. Comparative Economic & Social Systems 2020(04):30-40.
[6] Huang W. Study on the effect of targeted poverty alleviation by medical insurance policy—Based on the pilot evaluation household survey data of URBMI. Economic Research Journal 2017;52(09):117-32.
[7] Das J, Hammer J, Leonard K. The quality of medical advice in low-income countries. J Econ Perspect 2008;22(2):93-114.
[8] Xu X. Analysis of the formation of the "poverty-disease" trap for rural residents. Shandong Social Sciences 2018(08):66-72.
[9] Zhang H, Zhu W. Catastrophic illness shocks, medical insurance, and poverty vulnerability of urban households in china—A study based on micro-simulation methods. Insurance Studies 2026(01):84-101.
[10] Zeng G, Yang P. The impact of catastrophic medical expenditure on income disparity among Chinese households—A study based on CFPS survey data. Jianghan Tribune 2018(05):25-33.
[11] Huang W. Insurance policy and Chinese-style poverty reduction: Experience, dilemma and path optimization. Management World 2019;35(01):135-50.

[12]Korenman S, Remler DK, Hyson RT. Health insurance and poverty of the older population in the united states: the importance of a health inclusive poverty measure. J Econ Ageing 2021;18:100297.
[13]Gao J, Li H, Xu Y. Can commercial health insurance alleviate "illness-induced poverty" for families covered by urban-rural resident basic medical insurance?—Empirical evidence under catastrophic illness shocks. Journal of Jiangxi University of Finance and Economics 2019(05):81-91.
[14]Xie Y, Yang J. The policy effect of full medical insurance coverage on curbing illness-induced poverty and relapse. Journal of Beijing Normal University (Social Sciences Edition) 2018(04):141-56.
[15]Wang Y, Liu L. The impact of the new rural cooperative medical scheme and health human capital on rural residents' labor participation. Chinese Rural Economy 2016(11):68-81.
[16]Qin L, Chen C, Li Y, Sun Y, Chen H. The impact of the new rural cooperative medical scheme on the “health poverty alleviation” of rural households in china. J Integr Agric 2021;20(4):1068-79.
[17]He W, Shen S. Medical insurance payment methods and supplier moral hazard in medical services—An empirical analysis based on medical insurance reimbursement data. Statistical Research 2020;37(08):64-76.
[18]Zhao S, Zang W, Yin Q. The welfare effects of medical insurance coverage levels. Economic Research Journal 2015;50(08):130-45.
[19]Peng X, Qin X. Will medical insurance induce ex-ante moral hazard? Theoretical analysis and empirical evidence. China Economic Quarterly 2015;14(01):159-84.
[20]Liu J, Liu K, Zeng Y. The mechanism of continuous growth in medical costs—An analysis based on historical data and field materials. Social Sciences in China 2015(08):104-25.
[21]Feng J, Wang Z, Song H. Self-selection and medical costs in China's medical insurance system—A study based on the enrollment behavior of flexible employees. Journal of Financial Research 2018(08):85-101.
[22]Liu N, Wu Y. Study on the household income distribution effect of China's medical insurance system reform—Based on CHNS 2000–2011 analysis. Journal of Xiangtan University (Philosophy and Social Sciences) 2018;42(01):100-6.
[23]Hong H, Ning M, Luo Y. Has the integration of urban and rural resident medical insurance inhibited the health depreciation of the rural middle-aged and elderly? Chinese Rural Economy 2021(06):128-44.
[24]Fan H, Wang Y, Qi R. Integrated urban-rural medical insurance and substantive health equity—Leaping over the rural "health poverty" trap. Chinese Rural Economy 2021(04):69-84.
[25]Wang T, Peng X. The impact of social security on fertility intentions: Evidence from the new rural cooperative medical scheme. Economic Research Journal 2015;50(02):103-17.
[26]Jin S, Yu J, Tian R. Does China's basic medical insurance system promote benefit equity?—An empirical analysis based on China household finance survey. China Economic Quarterly 2020;19(04):1291-314.
[27]Zang W, Chen C, Zhao S. Social health insurance, disease heterogeneity and medical costs. Economic Research Journal 2020;55(12):64-79.
[28]Ma C, Qu Z, Song Z. Opportunity inequality in medical care for the floating population under the background of integrated urban-rural health insurance—The paradox of ex-ante compensation principle and ex-post compensation principle. China Industrial Economics 2018(02):100-17.
[29]Zhou Q, Tian S, Pan J. Inequity under equality—A theoretical and empirical study on the benefit equity of urban resident basic medical insurance. Economic Research Journal 2016;51(06):172-85.
[30]Scott A, Schurer S, Jensen PH, Sivey P. The effects of an incentive program on quality of care in diabetes management. Health Econ 2010;18(9):1091-108.