Health Economics Review

RESEARCH

Open Access

The moral dilemma of healthcare service utilization: a perspective from the consolidation of urban and rural resident health insurance policy in China

Hui Yuan[1], Jubo Han[1*] and Ruifeng Luo[1]

Abstract

Patient moral hazard is a significant issue in healthcare system reform and a prominent factor affecting the efficiency of healthcare services in China. Based on the consolidation of urban and rural resident health insurance, this paper employs a staggered DID model to analyze the impact of patient moral hazard on the healthcare service utilization. The findings are as follows. First, the healthcare service utilization significantly increases after the consolidation of urban and rural resident health insurance. This conclusion remains robust when subjected to the placebo test, the mitigation of non-random policy effects, and the exclusion of other insurance type interference. Second, after considering the effects of patient healthcare demand release, supplier-induced demand, and collusion between doctors and patients, we find the evidence of patient moral hazard leading to increased healthcare service utilization. Third, patient moral hazard is mainly manifested in the heightened utilization of patients with general illnesses and middle income, indicating the phenomena of "over-treatment for minor ailments". The study is limited by its focus on expanded reimbursement in urban and rural insurance consolidation, excluding details like fund management changes. Future research should incorporate more policy details and longer time horizons.

Keywords Patient moral hazard, Healthcare service utilization, Consolidation of urban and rural health insurance, Staggered DID

JEL classification G22, I18

*Correspondence:
Jubo Han
hanjubo1997@163.com
[1]School of Finance, Zhongnan University of Economics and Law, Wuhan 430073, China

Introduction

The challenge of "difficult access and high expense of medical treatment" has hindered the development of healthcare system in China. With the improvement of medical insurance coverage leading to a significant reduction in healthcare service costs, the issue of "high expense of medical treatment" has been partially alleviated. However, the problem of "difficult access of medical treatment" still remains. The overcrowding and bed shortages in large hospitals, making it difficult for some severe and emergency patients to access in. Concurrently, there is the overutilization of healthcare services for less severe cases, resulting in the wastage of medical resources and the detriment to the residents' health. In the healthcare market, patients may seek unnecessary treatments out of a sense of prevention, while healthcare suppliers may induce patients to overuse services for financial benefits. Information asymmetry between healthcare suppliers and patients exacerbates the structural shortage in medical resources, generating a significant challenge for healthcare reform in China.

In 2016, the Chinese central government issued "Opinions in Consolidating the Basic Medical Insurance System for Urban and Rural Residents", marking the initiation of nationwide consolidation of urban and rural medical insurance. This policy integrates two different medical insurance systems, namely, the urban resident basic medical insurance and the new rural cooperative medical care, into a comprehensive medical insurance system known as the basic medical insurance for urban and rural residents. The consolidation aims to eliminate disparities in medical insurance for urban and rural residents, promote the equity and accessibility of healthcare services, and represents a substantial government investment to benefit the people. However, with information asymmetry in the healthcare market, is there any patient moral hazard due to the promotion of medical insurance coverage? And would the healthcare suppliers exploit their information advantage seeking for the extra incentives? These are crucial topics regarding the healthcare reform in China. This study employs the consolidation of urban and rural health insurance as an exogenous shock, and applies a staggered DID model to analyze the impact of patient moral hazard on healthcare service utilization, showing the practical significance for reducing unnecessary medical expenditure and alleviating the burden of the healthcare system.

Since Arrow [1] introduced the concept of moral hazard into the healthcare research, the information asymmetry has been the focus in the academic. Information asymmetry in the healthcare insurance market originates from both demand and supply sides. Patient moral hazard indicates that individuals, upon purchasing health insurance, may engage in high-risk behaviors or consume more healthcare services. Moral hazard can be classified into ex-ante and ex-post moral hazard based on different motivations. Ex-ante moral hazard suggests that individuals, after participating in health insurance, may reduce their concerns for their own health. For example, some individuals may increase their unhealthy behaviors, such as smoking and drinking.

This article focuses on ex-post moral hazard, which refers to individuals deliberately causing insurance incidents to claim insurance benefits. The wastage of medical resources and the inefficiency of healthcare services prompted by moral hazard derives significant concerns in healthcare insurance reform [2–6]. Junga and Tran [7] used U.S. data to find that, when the medical insurance coverage grew to a certain level, ex-post moral hazard became the main cause of healthcare costs increase. Wong et al. [8] discovered that patients participating in social medical insurance tended to overuse healthcare services compared to patients who purchased private health insurance. Cheng and Chiang [9] found that the frequency of patient seeking medical advice was doubled after the implementation of health insurance policy. Einav and Finkelstein [10] found that, with increased medical insurance coverage, the total cost of healthcare payments expanded significantly.

Some scholars have simplified healthcare services to analyze them as ordinary goods [11–13]. Nyman et al. [12] proved that the reimbursed portion of medical expenses by individuals who participated in medical insurance was considered a transfer payment from healthy individuals to patients. The increased healthcare expenditures by individuals after participating in medical insurance can be divided into two parts. First, efficient income effect, which means the decrease in self-paid is considered as an increase in income. Second, inefficient substitution effect, which means the decrease of medical products price would not increase the healthcare service utilization. Li et al. [13] analyzed the patient moral hazard of rural residents in China based on this theory and combined it with the consolidation of urban and rural resident health insurance policy. They claimed that the effective moral hazard accounted for 43.33-66.36% of the overall moral hazard.

Other scholars argue that supplier-induced demand is a significant reason for the increased utilization of healthcare services [14–17]. Supplier-induced demand in the healthcare market refers to situations where healthcare suppliers, driven by their own benefits, encourage patients to increase unnecessary medical consumption. Roberts et al. [18] found that in the U.S. health insurance market, larger market share insurance companies could lower prices for medical service reimbursement, confirming the significant role of supplier market power in bargaining. Finkelstein [19] discovered that the

introduction of health insurance led to increased investments by hospitals and higher medical expenses for patients. A major factor of increasing patient medical costs was healthcare suppliers employing advanced medical equipment to enhance treatment intensity. Wu [17] observed that after restricting the percentage of medicine sales, some doctors compensated for reduced medicine sales by increasing non-medicine expenditures, with no significant impact on the overall income of healthcare suppliers. These phenomena are prevalent among insured patients, and further research has revealed that the insurance status of patients plays a crucial role in supplier-induced behavior. Eliason et al. [16] found that healthcare suppliers' decision-making was not mainly driven by delivering effective patient care and treatment, but rather aimed at maximizing their financial compensation. Currie et al. [15] noted that in the usage of antibiotics, patients did not actively choose the treatment. Instead, the decisions were depended on the healthcare suppliers. Lu [20] also claimed that when doctors could gain profit from patients' medical expenses, the treatment expenses in insured patients were 43% higher than those in uninsured patients, indicating that doctors considered patients' insurance status to engage in excessive healthcare, thereby earning extra benefits.

Healthcare policy reform that reduces medical costs for residents could potentially increase their healthcare service utilization. However, there is limited research on whether the increased utilization of healthcare services is associated with moral hazard. Some studies used insurance coverage as a premise for patient moral hazard [8, 10, 12]. Nonetheless, the level of social health insurance coverage in China is almost saturated, such approach may be less suitable currently in China. The increase in healthcare service utilization is mainly influenced by three factors: patient moral hazard, healthcare demand release, and supplier-induced demand. Some studies attributed the increase in healthcare service utilization solely to policy effects without considering the impact of patient moral hazard. Additionally, other studies identified the presence of patient moral hazard but may not effectively account for the interference of supplier-induced demand [11, 12]. Certain literatures found that patient moral hazard could negatively affect the policy effects of health insurance reforms but has not delved into which specific demographic groups are more prone to have moral hazard [3–5].

This paper extends previous literatures in the following ways. Firstly, with the medical insurance coverage approaching saturation, the measurement standard based on whether residents are insured is outdated. Instead, this study employs the increased reimbursement rates in the consolidation of urban and rural health insurance as an exogenous shock, which is more practically significant. In addition, previous research relied on the recall of survey respondents to determine their participation in medical insurance. This method overlooks issues related to respondents' limited awareness, lack of expertise, and recall biases regarding the healthcare reforms. Given that the consolidation of urban and rural health insurance is implemented at the city level, it is more realistic to be accordance with the timing of policy implementation in respondents' cities. Secondly, few studies explore the impact of patient moral hazard, medical demand release, and supply-induced demand factors comprehensively. This paper, considering the characteristics of various information asymmetry factors, offers a more comprehensive analysis of key factors affecting healthcare service utilization to confirm the existence of patient moral hazards. Thirdly, this paper provides a validation of the existence of patient moral hazards by examining the behaviors and motivations of different demographics, indicating that the "over-treatment for minor ailments" is the main pattern of patient moral hazards, especially among patients with general illnesses and middle income.

Institutional background and theoretical framework

Institutional background

The consolidation of the basic medical insurance for urban and rural residents (simplified as "urban-rural health insurance consolidation") is a major reform policy implemented in accordance with documents such as "Opinions on Consolidating the Basic Medical Insurance for Urban and Rural Residents". Its main objective is to establish a more equitable and efficient healthcare coverage system, catering to the diverse medical insurance needs of different regions and groups while reducing the urban-rural disparity in healthcare coverage. The core of this reform involves the consolidation of the previous urban resident basic medical insurance and the new rural cooperative medical insurance into one unified system, known as the urban-rural resident basic medical insurance.

Urban-rural health insurance consolidation was initially piloted in selected cities, such as Chengdu in 2009. In subsequent years, developed cities in eastern China, such as Tianjin, Xiamen, and Suzhou, also implemented this reform. The nationwide implementation began in 2016 after the central government released the "Opinions on Consolidating the Basic Medical Insurance for Urban and Rural Residents". By the end of 2018, 80% of cities in China had completed the consolidation. According to the guideline, consolidated cities are required to achieve the "Six Unifications" within the administrative boundaries of the municipality, encompassing unified coverage, funding policy, benefit provisions, insurable

medicine contents, designated management, and fund administration.

The key components of urban-rural health insurance consolidation can be summarized as follows.

Equalization of urban-rural benefits and improved welfare for rural residents Documents from municipal government claim the unequivocal use of a common medicine list for urban and rural residents engaged in the basic medical insurance. This equalizes the healthcare coverage for urban and rural residents and notably enhances the welfare of rural residents.

Expansion of the insurable medicine contents and insurance coverage With the standardization of insurable medicine contents, the range of medicine eligible for reimbursement by rural residents has significantly expanded. For instance, in Hebei province, prior to the consolidation, the medicine contents for the new rural cooperative medical insurance and urban resident medical insurance included 1,000 and 2,400 types, respectively. After the consolidation, the unified medicine contents comprised 2,900 types.

Increase in reimbursement rates Consolidation documents from various cities include provisions for raising the reimbursement rates for medical services. For instance, in Guangdong province, after the consolidation, the inpatient reimbursement rate increased from 54 to 76%, and the payment cap rose from ¥5,000 to ¥44,000.

These measures represent significant changes in urban-rural health insurance consolidation, with the aim of improving access to medical services for residents, which holds practical significance in our subsequential research.

Theoretical framework

Based on the facts of improved medical benefits due to the consolidation of urban and rural health insurance, this paper analyzes the impact of information asymmetry on healthcare service utilization. As depicted in Fig. 1, the reasons for the increase in healthcare service utilization resulting from the consolidation are complicated. The urban-rural health insurance consolidation affects medical service utilization in following ways:

Ex-post moral hazard The consolidation may influence patient behaviors due to the increase of reimbursement rates. This, in turn, might lead to unnecessary medical expenses incurred by patients.

Ex-ante moral hazard The reduced cost of medical treatment after the consolidation may weaken the importance which some residents attach to their health. This could potentially encourage or exacerbate unhealthy behaviors such as smoking and alcohol consumption. However, the academic has not yet reached a consensus regarding the direct relationship between ex-ante moral hazard and

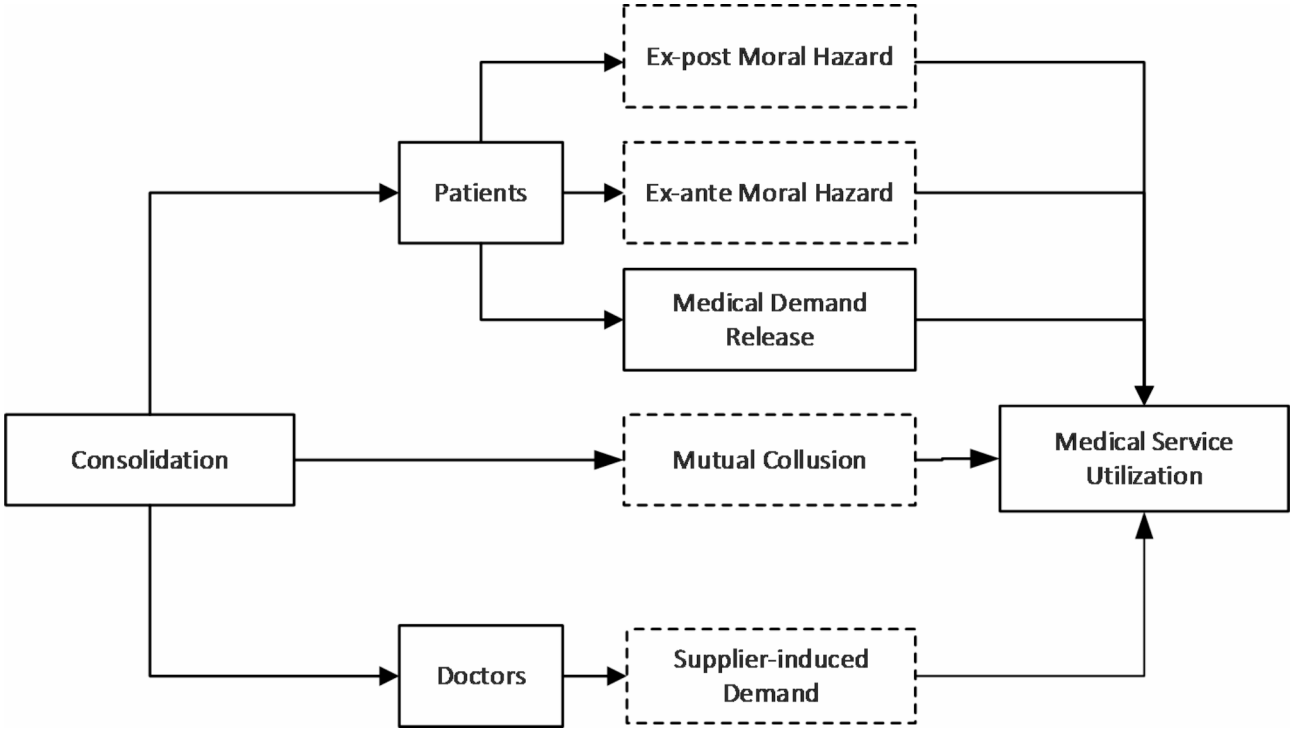

Fig. 1 Theoretical framework

increased healthcare service utilization [21–23]. A comprehensive explanation of this issue requires specialized knowledge from the medical research and clinical experience. Therefore, this paper does not focus on ex-ante moral hazard in our analysis.

Medical demand release and supplier-induced demand The treatment cost for patients is reduced after the consolidation. Consequently, some individuals who previously refrained from seeking medical care due to cost considerations can now afford the necessary treatment. And urban-rural health insurance consolidation also affects healthcare suppliers. These suppliers with specialized skills hold information advantages in the healthcare market. They may be motivated by financial benefits, resulting from the substantial government funding, to induce patients into unnecessary medical consumption.

Collusion between patients and doctors Patients and doctors may conspire to perform unnecessary diagnosis and treatments in order to obtain medical insurance reimbursements or other economic incentives, ultimately resulting in the wastage of healthcare resources.

The urban-rural health insurance consolidation in China aims to equalize healthcare benefits between urban and rural residents and elevate the overall healthcare coverage for residents. However, factors such as patient moral hazard, supplier-induced demand, and mutual collusion can influence the effectiveness of the policy. How can we validate the claim that patient moral hazard may lead to an increase in healthcare service utilization? Do healthcare suppliers encourage or induce patients into overutilizing healthcare services? This paper places a particular focus on addressing these questions above.

Empirical design

Data source

This paper utilizes the China Health and Retirement Longitudinal Study (CHARLS) database from Peking University for the analysis. The selection of the CHARLS database is based on several considerations. Firstly, CHARLS data includes specialized and detailed individual health examination information, which is essential for a comprehensive study of patients' healthcare service utilization behaviors. Secondly, the middle-aged and elderly people face increased health risks and a higher frequency of healthcare service utilization. The main survey subjects in the CHARLS database are individuals aged at least 45, which provides a sufficient sample size that meets the research requirements. Thirdly, the implementation of urban-rural health insurance consolidation is at the city level. And the CHARLS database allows access to the municipal information (unlike other micro-databases only at the province level), which aligns with the research needs. This paper uses survey data from 2011 to 2018, including four merged periods. We remove the samples with missing or abnormal values, and all continuous variables underwent winsorization at the 1% level on both the upper and lower bound.

Model specification

Urban-rural health insurance consolidation is implemented at the city level, and the timing of completion varies across different cities. Based on such characteristics, this paper employs a staggered Differences-in-Differences (DID) model for the analysis.

The model is specified as follows:

$$Utilization_{ict} = \beta_0 + \beta_1 URH_{ct} + \beta_3 X_{ict} + \delta_c + \delta_t + \epsilon_{ict} \quad (1)$$

In Eq. (1), i, c, and t represent the individual identifier, the place of household registration, and the survey year, respectively. $Utilization_{ict}$ denotes the healthcare service utilization, defined as whether the individual was hospitalized during the year. URH_{ct} is the core explanatory variable, representing whether the individual's city implemented urban-rural health insurance consolidation in year t. Detailed explanations for the variable selection in the subsequent sections will provide more information about this indicator. X_{ict} represents the control variables selected in this paper. δ_c and δ_t control for fixed effects at the regional and time levels, and ϵ_{ict} is the random error term.

Variable selection

Dependent variable: *Utilization*. This paper selects whether the respondent was hospitalized within one year of the survey as the dependent variable. If the respondent had a hospitalization experience within one year of the survey, it is denoted as 1; otherwise, it is denoted as 0. The reason of choosing hospitalization representing the healthcare service utilization is that patients are more impressive about their hospitalization experiences, which is beneficial to avoid recall bias in the survey data [13].

Core independent variable: *URH*. The specific time of policy implementation is obtained from formal documents published by the municipal government. If the policy implementation date precedes the year of the respondent's survey, the variable *URH* is assigned 1 for the year t; otherwise, it is assigned 0.

Control variables. The main control variables include gender, age, household registration type, marital status, years of education, self-rated health index, household population size, household annual income, and household wealth.

Table 1 provides descriptive statistics for main variables in this paper.

Table 1 Descriptive statistics for main variables

Variables	Definition	Std	Mean
Utilization	Hospitalization within one year of the survey	0.310	0.108
URH	Urban-rural health insurance consolidation	0.469	0.326
Gender	Male = 1; Female = 0	0.500	0.482
Age	Age	10.00	60.79
Registration	Household registration type: Rural = 1; Urban = 0	0.417	0.776
Marriage	Marital status: Married or cohabited = 1; Otherwise = 0	0.347	0.860
Education	Years of education	4.727	4.544
Health	Self-rated health index	0.963	3.039
Population	Household population size	1.663	3.240
Ln_Income	Household annual income	2.360	9.012
Ln_Wealth	Household wealth	1.882	11.080

Table 2 The baseline regression of URH affecting utilization

	(1)	(2)	(3)
	LPM	Logit Model	Probit Model
URH	0.0124*	0.1711**	0.0941**
	(0.0074)	(0.0818)	(0.0436)
Gender	0.0065	0.0515	0.0310
	(0.0042)	(0.0464)	(0.0250)
Age	0.0026***	0.0292***	0.0157***
	(0.0002)	(0.0027)	(0.0015)
Registration	-0.0306***	-0.3199***	-0.1735***
	(0.0058)	(0.0613)	(0.0334)
Marriage	0.0082	0.1118*	0.0580*
	(0.0059)	(0.0638)	(0.0347)
Education	0.0002	0.0026	0.0010
	(0.0005)	(0.0059)	(0.0032)
Health	-0.0569***	-0.6693***	-0.3523***
	(0.0021)	(0.0256)	(0.0134)
Population	-0.0028**	-0.0339**	-0.0174**
	(0.0014)	(0.0165)	(0.0087)
Ln_Income	0.0016*	0.0145	0.0089
	(0.0010)	(0.0118)	(0.0062)
Ln_Wealth	-0.0011	-0.0023	-0.0013
	(0.0012)	(0.0127)	(0.0070)
Respondents	23,876	23,876	23,876
Time fixed effects	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes

Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Evidence of increased healthcare service utilization after URH

Baseline regression

Table 2 reports the impact of urban-rural health insurance consolidation on residents' healthcare service utilization. Columns (1) to (3) respectively present the results estimated using the Linear Probability Model (LPM), Logit Model, and Probit Model. The results indicate that urban-rural health insurance consolidation significantly increases the healthcare service utilization. Subsequent analysis in this paper is based on the Probit Model. Column (3) shows that the urban-rural health insurance consolidation has significantly increased the probability of patients being hospitalized by 9.41%. Older respondents and elderly individuals living alone exhibit a higher probability of hospitalization. Respondents with larger family size and higher level of health have a lower probability of being hospitalized.

Parallel trends test

The prerequisite for using DID model is passing the parallel trends test. Parallel trends refer that there is

no difference in hospitalization situations between the treated and control groups before the policy implementation. As shown in Fig. 2, the healthcare service utilization of the treated group and the control group does not exhibit significant differences before the policy implementation, thus passing the parallel trends test.

Robustness test

Placebo test

To eliminate the bias caused by random factors, we conduct a placebo test. The procedure involves creating a counterfactual virtual treated group for regression. Specifically, random treated and control groups are selected for urban-rural health insurance consolidation, repeating 500 times to obtain the virtual t-value and plot it on the graph. As shown in Fig. 3, the t-value in the placebo test is mostly concentrated around 0 and is far away from the true t-value, indicating that the hospitalization behavior of residents is not significantly affected by random factors. This suggests that the increase in healthcare service utilization after the consolidation of urban-rural health insurance is not due to random factors.

Randomness of policy implementation

The implementation of urban-rural health insurance consolidation is not entirely random. It began with pilot city and gradually expanded nationwide. Regions with strong financial resources and abundant healthcare resources are more likely to implement the consolidation. To address these issues, this study employs per capita GDP and local fiscal expenditures as indicators of economic development and the number of medical institutions as a measure of healthcare resource abundance. Based on the baseline regression model, the study further controls these variables to mitigate the non-randomness of policy implementation.

In Table 3, column (1) controls macroeconomic factors at the city level, including per capita GDP, local fiscal expenditures, and the number of medical institutions. The column (2) adjusts the value using the base year 2011 to account for the inflation. The results indicate that healthcare service utilization significantly increased after urban-rural health insurance consolidation, validating the robustness of the baseline results.

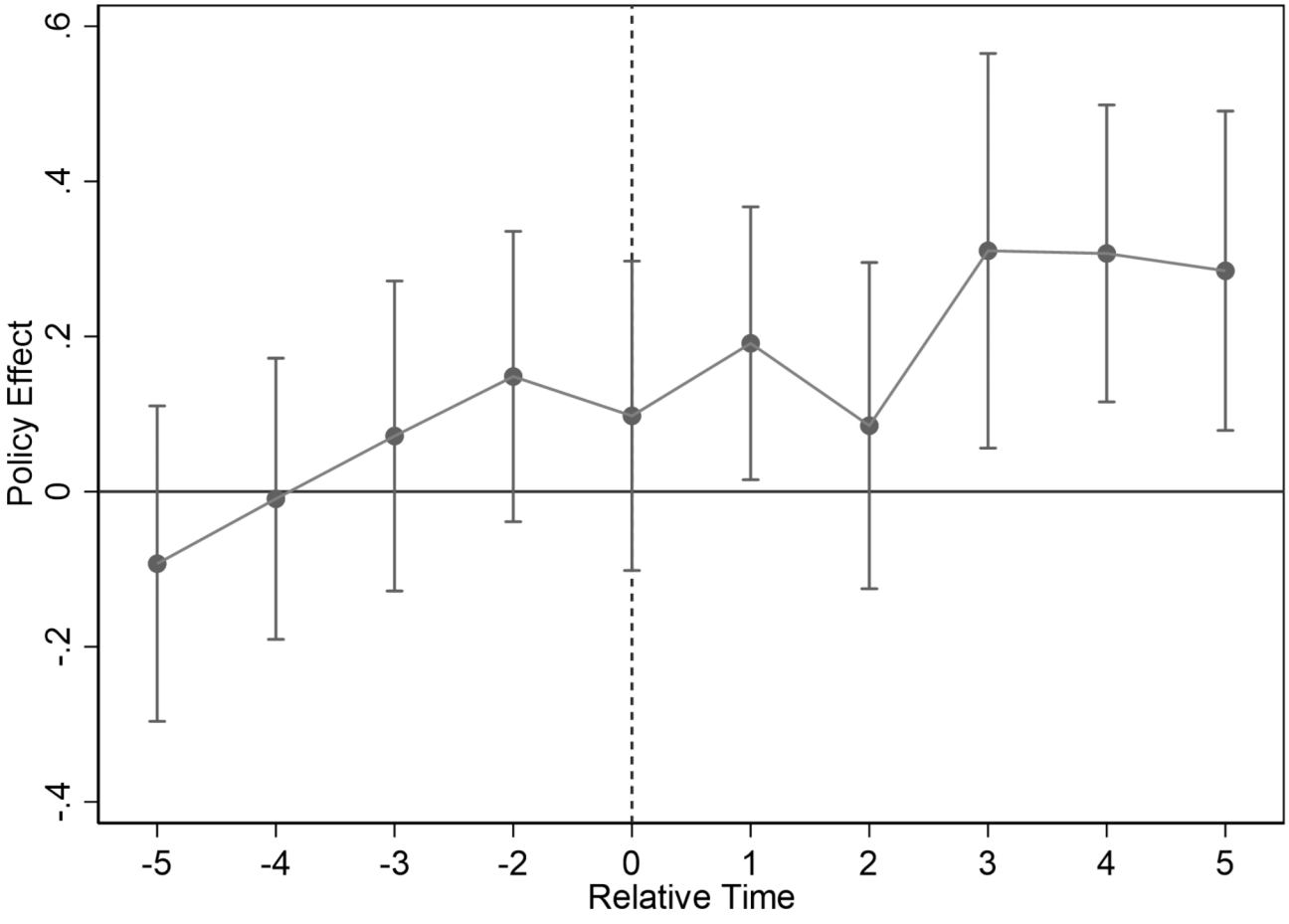

Fig. 2 Parallel trends test

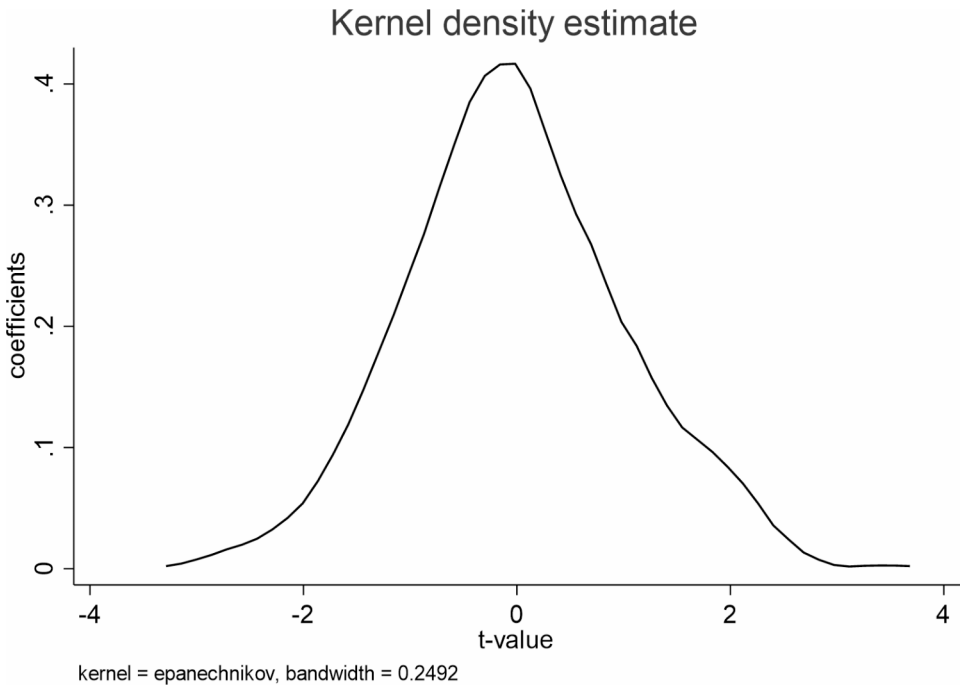

Fig. 3 Placebo test

Table 3 Randomness of policy implementation

	(1) Control macroeconomic factors	(2) Adjust the price level
URH	0.0938*	0.0953*
	(0.0503)	(0.0515)
Per capita GDP	0.0256	0.0241
	(0.0374)	(0.0384)
Fiscal expenditures	0.1093***	0.1072***
	(0.0324)	(0.0330)
Medical institutions	-0.0092	-0.0149
	(0.0163)	(0.0166)
Respondents	20,187	19,336
Controls	Yes	Yes
Time fixed effects	Yes	Yes
Regional fixed effects	Yes	Yes

Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Additional robustness check

Although the samples are mainly middle-aged and elderly individuals, a small number of respondents may have participated in urban employee basic medical insurance or purchased commercial health insurance, both of which are not influenced by the consolidation. To mitigate such potential interference, we remove samples of individuals who participated in those two insurances, to make sure that all respondents could be affected by the consolidation. Furthermore, we use the number of hospitalization days within one year in the survey as a replacement for whether the hospitalization had occurred within one year in the survey as dependent variable. As shown in Table 4, the coefficients remain significantly positive after eliminating interference from other insurance types and replacing the dependent variable, thus confirming the robustness of the conclusion.

Evidence of patient moral hazard leading to healthcare utilization increase

Patient willingness and supplier inducement

Based on the conclusion from previous analysis, we prove that healthcare utilization significantly increases after the consolidation of urban and rural medical insurance. In this section, we will continue to explore the evidence of information asymmetry leading to increased healthcare utilization and figure out that which types of information asymmetry factors are the primary reasons.

Table 4 Eliminate interference from other insurances and replace the dependent variable

	(1)	(2)
	Eliminate interference from other insurances	**Replace the dependent variable**
URH	0.0915**	0.2911*
	(0.0445)	(0.1535)
Respondents	23,107	24,753
Controls	Yes	Yes
Time fixed effects	Yes	Yes
Regional fixed effects	Yes	Yes

Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table 5 Patient willingness and supplier inducement: urban-rural disparities

	(1)	(2)
	Rural	**Urban**
URH	0.1001**	0.0984
	(0.0498)	(0.0937)
Respondents	18,945	4924
Controls	Yes	Yes
Time fixed effects	Yes	Yes
Regional fixed effects	Yes	Yes

Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Patients usually have limited knowledge about their own illnesses and lack sufficient professional expertise, making it difficult to make rational decisions. In contrast, medical suppliers possess extensive medical knowledge and experience, holding more information regarding disease diagnosis and treatment. Information asymmetry results in patients more influenced by healthcare suppliers when choosing medical services. Additionally, patients may choose unnecessary or expensive treatment methods based on personal preferences, which could bring about the wastage of healthcare resources. For instance, some patients might overly rely on advanced medical technologies while neglecting cost-effective treatment options. Therefore, this paper will analyze the impact of both patients and medical suppliers.

Urban-rural disparities

The policy goal of urban and rural medical insurance consolidation is to bridge the gap in healthcare coverage between urban and rural areas, with rural residents being the primary beneficiaries. After the consolidation, rural residents experience a significant increase in the reimbursement rate for their medical expenses, which can lead to higher healthcare resource utilization. Given that healthcare resources are highly concentrated in urban areas in China, if supplier inducement is the primary cause, urban residents are expected to have a more significant increase in healthcare utilization. Conversely, if there is no evidence of supplier inducement in analysis, we can tentatively claim that the increase in healthcare utilization is due to patients' active choices.

Table 5 presents the results of subsample regressions for rural and urban residents in Column (1) and (2) respectively. The results show that healthcare utilization significantly increases for rural residents after urban and rural medical insurance consolidation, while urban residents are not significantly affected. Despite healthcare resources being concentrated in urban areas, healthcare utilization for urban residents does not significantly increase. It suggests that supplier inducement is not the primary reason for the increase in healthcare utilization.

Number of medical visits

If supplier inducement is the primary reason of increased healthcare utilization, patients with a higher number of medical visits are more likely to make hospitalization decisions induced by healthcare suppliers. Patients who seek medical treatment for the first time have not previously contacted with doctors and have not established long-term relationships with doctors, resulting in lower trust in medical suppliers. They tend to be more cautious about doctors' advice and seek various opinions to verify the treatment rationality. Therefore, it is less likely for patients with first visits to be influenced by supplier inducement. Additionally, considering that the number of medical visits is related to the severity of the medical condition, patients of different types cannot be directly compared. Therefore, this study further utilizes the Propensity Score Matching (PSM) method to filter the sample. Using the k-nearest neighbor matching method, the treated group are matched with the control group one by one (caliper within 0.05).

Table 6 Patient willingness and supplier inducement: number of medical visits

	(1)	(2)	(3)	(4)	(5)	(6)
	First-time visit	**Multiple visits**	**First-time visit**	**Multiple visits**	**First-time visit**	**Multiple visits**
URH	0.3338**	0.1727	0.0740*	0.0264	0.2652*	0.2248
	(0.1328)	(0.1220)	(0.0435)	(0.0427)	(0.1497)	(0.1401)
Respondents	2261	2167	1036	1022	1960	1898
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Supply-side Factors	No	No	No	No	Yes	Yes

Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table 7 Patient moral hazard and medical demand release: disease types

	(1)	(2)	(3)	(4)	(5)
	High blood pressure	**Diabetes**	**High blood lipids**	**Malignant tumors**	**General illnesses**
URH	-0.0046	0.0819	-0.0499	-0.8189*	0.2151***
	(0.0699)	(0.1284)	(0.0996)	(0.4538)	(0.0656)
Respondents	8090	2253	3930	315	12,943
Controls	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Column (1) and (2) in Table 6 respectively analyze subsamples of patients with first-time and multiple medical visits. Column (3) and (4) represent the analysis in PSM-DID method. The results show that healthcare utilization significantly increases for patients with first-time medical visit, while there is no significant change in healthcare utilization for patients with multiple medical visits. This result remains valid in PSM-DID analysis, which also confirms that the increase in healthcare utilization is due to patients' active choices rather than the supplier inducement. Columns (5) and (6) further incorporate the number of regional medical institutions to control for supply side factors. The results still approve the conclusions.

Patient moral hazard and medical demand release

Although the previous analysis concludes that the increase in healthcare utilization is due to patients' active choices, the evidence of patient moral hazard is not sufficient. The motivation for increasing healthcare utilization can be divided into medical demand release and moral hazard. After the consolidation of urban and rural health insurance, the increase in reimbursement rates to some extent has addressed the issue of "high expense of medical treatment". Patients who previously couldn't afford medical treatment now have an opportunity to seek the medical care, which can be referred to as medical demand release. However, some individuals may engage in excessive medical behaviors after the consolidation, constituting patient moral hazard. Effectively distinguishing between medical demand release and patient moral hazard is crucial for healthcare system reform.

Medical demand release primarily occurs within patients who could not access to medical treatment in the past, which include two main types. First, those with severe illnesses that require significant medical expenditures. Second, those with poor economic conditions who cannot afford healthcare costs or choose not to accept medical treatment. These two patient groups typically lack the motivation for moral hazard. Therefore, this section will analyze this from disease types and income levels to effectively differentiate between medical demand release and patient moral hazard.

Disease types

This paper categorizes common diseases into chronic diseases (high blood pressure, high blood sugar[1], high blood lipids), malignant tumors, and general illnesses for analysis. As shown in Table 7, the healthcare utilization of patients with chronic diseases does not significantly increase. In contrast, healthcare utilization decreases significantly for patients with malignant tumors at the 10% level. Notably, healthcare utilization significantly increases at the 1% level for patients with general

[1] Since CHARLS does not include specific questions about whether individuals have high blood sugar, we use the existence of diabetes as the proxy. Diabetes is a prevalent condition among those with high blood sugar, but it's important to note that not all individuals with high blood sugar have diabetes. By using diabetes as an indicator, we potentially underestimate the effect of supplier-induced demand.

Table 8 Patient moral hazard and medical demand release: income levels

	(1)	(2)	(3)
	Low-income	**Middle-income**	**High-income**
URH	0.0978	0.1482**	0.0589
	(0.0776)	(0.0754)	(0.0799)
Respondents	7983	8259	7625
Controls	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes
Regional fixed effects	Yes	Yes	Yes

Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table 9 Medical collusion

	(1)	(2)
	Total expenses	**Self-paid expenses**
URH	0.0981	0.0950
	(0.0698)	(0.0633)
Respondents	24,542	24,542
Controls	Yes	Yes
Time fixed effects	Yes	Yes
Regional fixed effects	Yes	Yes

Standard errors in parentheses, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

illnesses. This indicates that patients with general illnesses exhibit excessive medical behaviors, verifying the existence of patient moral hazard. Furthermore, in theory, improved supply-side efficiency should uniformly increase the healthcare utilization of patients. However, our study finds that while healthcare utilization does not significantly rise for patients with severe and chronic illnesses, it increases significantly for those with general illnesses. Therefore, we conclude that the patient moral hazard persists even if accounting for the improved supply-side efficiency.

Income levels

Column (1) to (3) in Table 8 represent regression for subsamples of low, middle, and high-income groups. The results indicate that, after the consolidation of urban and rural health insurance, healthcare utilization significantly increases at the 5% level for the middle-income group. However, the coefficients for the low and high-income groups are not significant. The low-income group exhibits the highest willingness to release medical demand, but their healthcare utilization does not increase. On the other hand, healthcare utilization significantly increases for the middle-income group. This further substantiates the existence of moral hazard. Although urban-rural healthcare consolidation is expected to improve medical coverage for low-income group, this study finds its impact on their hospitalization is limited. First, even with higher reimbursement rates, low-income group struggles to afford self-paid hospitalization costs. Second, multi-tiered payment system in underdeveloped areas leads low-income group to choose the lowest tier of limited coverage. Third, low-income group is less sensitive to policy changes and faces challenges in accessing policy information and adjusting their medical behaviors.

Further analysis: medical collusion

In the preceding sections, we identify the impact of patient moral hazard on healthcare utilization. However, "medical collusion" may occur in the healthcare market where both patients and medical suppliers mutually agree on the treatment that maximizes their gains. For instance, patients and doctors might jointly decide to increase treatment costs to reach the deductible threshold for medical insurance reimbursement. The primary purpose of medical collusion is for patients to reduce their self-paid medical expenses, while medical suppliers can earn higher profits. The profits for medical suppliers are highly correlated with the total medical expenses. In this regard, this paper uses total medical expenses to represent the healthcare suppliers' profit level and uses patients' self-paid medical expenses to examine whether patients benefit from this collusion.

Column (1) and (2) in Table 9 report regressions for both total medical expenses and self-paid expenses. The consolidation policy has no significant impact on either of these factors. Further regression analyses on subsamples of rural residents and hospitalized patients confirm this conclusion (not shown in the table). Patients' self-paid medical expenses don't decrease, and healthcare suppliers' profits don't increase, suggesting that medical collusion is not the primary reason for increased healthcare utilization.

Based on above analysis, total medical expenses have not significantly increased, indicating that supplier-induced demand has not been achieved. Additionally, neither medical suppliers' profits nor patient self-paid expenses have significantly changed, suggesting mutual collusion also has not been realized. These conclusions help mitigate the interference from supplier inducement and mutual collusion. The lack of significant reduction in patient self-paid expenses suggests that excessive medical behaviors are more likely due to patient' health concerns than their intentional waste of resources.

Conclusions and discussions

Conclusions

This paper utilizes the data from the China Health and Retirement Longitudinal Study (CHARLS) from 2011 to 2018 and employs a staggered DID model to analyze the impact of urban and rural medical insurance consolidation on healthcare service utilization. The study reveals the following findings. First, after the consolidation of urban and rural medical insurance, there is a significant increase in the residents' healthcare service utilization. This conclusion remains robust after a series of robustness tests. Second, after excluding the influence of supplier-induced demand, it is evident that the increased utilization of healthcare service is driven by patients' active choices. Furthermore, after accounting for the medical demand release, we show the evidence of patient moral hazard, but we find no evidence of medical collusion between doctors and patients. Third, patient moral hazard is observed mainly in patients with general illnesses and those with middle income, indicating the phenomena of "over-treatment for minor ailments".

Based on the findings, this paper puts forward the following policy implications.

First, the relevant authority should accelerate the implementation of the "gatekeeper system". The "gatekeeper system" stipulates that residents covered by social medical insurance must seek initial diagnosis and treatment at community clinics before being eligible to larger hospitals. By ensuring the ethical conduct and professional qualifications of primary healthcare personnel, community clinics and health stations can effectively screen patients, thus contributing to alleviating "over-treatment for minor ailments" and optimizing the allocation of medical resources.

Second, the government should further narrow the gap in healthcare services between urban and rural areas. The relevant department could strengthen the communication between professional medical experts and primary-level medical institutions, such as regular visits to rural areas. It is beneficial to identifying the issues in the rural healthcare development, thereby improving the quality of healthcare services.

Third, cultivate health concepts and enhance the public health literacy through systematic education and awareness campaigns. Inform the public about disease prevention and management by means of schools, communities, and media channels. Develop community health services to provide accessible primary care and promote healthy lifestyle. In the context of universal healthcare in China, it's essential to strengthen the information transparency, doctor training and supportive policies to prevent the excessive medical treatment.

Lastly, it is essential to continuously enhance the ethical standards of healthcare personnel. Physicians, as the information-advantaged side in the healthcare market, play a crucial role in patient decisions. While we do not find evidence of medical collusion between doctors and patients, there could be scenarios where healthcare suppliers cater to patients' inappropriate demands for the purpose of financial incentives. It is profound to incentivize healthcare personnel to provide transparent and detailed medical information and enable patients to understand the treatment risks and make rational decisions.

Discussions

This study has some limitations. It focuses only on the impact of expanded reimbursement under the urban-rural healthcare consolidation policy and does not address other aspects like fund management changes, which may lead to an incomprehensive analysis. Additionally, current CHARLS dataset may not fully capture the long-term effects of the policy. Future research should include more policy details and use longer-term datasets to provide a comprehensive assessment of the policy's ongoing impact and long-term effects.

Regarding the above limitations, future research can be explored as following. First, adopt a more comprehensive approach and take into account changes in the management department of the medical insurance fund, changes in the supply side of medical services, and details in the policy implementation process. This can be achieved through department interviews, document analysis and field research. Secondly, adopt updated datasets for longer time horizon to better analyze the long-term impact of consolidation policy. This can be obtained from the statistics of relevant authorities and the cooperation with research institutions.

Author contributions

Professor Hui Yuan provided invaluable guidance with his professional insights, significantly influencing the selection and implementation of our research. Dr. Jubo Han played a central role throughout the entire research process, overseeing data collection and analysis, leading the research design, implementing empirical studies, and contributing to the writing process by drafting the paper and integrating various sections. Dr. Ruifeng Luo dedicated considerable efforts to the writing and revision of our paper, making an exceptional contribution to the success of our research.

Funding
Not applicable.

Data availability
The datasets generated and analyzed during the study are available in the CHARLS. (http://charls.pku.edu.cn/en)

Declarations

Ethics approval and consent to participate
All methods were carried out in accordance with relevant guidelines and regulations.

Competing interests
The authors declare no competing interests.

Received: 9 December 2023 / Accepted: 14 January 2025
Published online: 18 February 2025

References

1 Arrow KJ. Uncertainty and the Welfare Economics of Medical Care. Am Econ Rev. 1963;53(5):941–73.
2 Pauly MV. The Economics of Moral Hazard: comment. Am Econ Rev. 1968;58(3):531–7.
3 Hsiao W, Cheng S, Yip W. What can be achieved with a single-payer NHI System: the case of Taiwan. Soc Sci Med. 2019;233(7):265–71.
4 Kwon S. Thirty Years of National Health Insurance in South Korea: lessons for Achieving Universal Health Care Coverage. Health Policy Plann. 2008;24(1):63–71.
5 Yip W, Lee Y, Tsai S, et al. Managing Health expenditure inflation under a single-payer system: Taiwan's National Health Insurance. Soc Sci Med. 2019;233(7):271–80.
6 Boone J. Basic versus supplementary Health insurance: Moral Hazard and adverse selection. J Public Econ. 2015;128:50–8.
7 Junga J, Tran C. Market Inefficiency, Insurance Mandate and Welfare: U. S. Health Care Reform 2010. Rev Econ Dyn. 2016;20(2):132–59.
8 Wong IOL, Lindner MJ, Cowling BJ, et al. Measuring Moral Hazard and adverse selection by Propensity Scoring in the mixed Health Care Economy of Hong Kong. Health Policy. 2010;95(1):24–35.
9 Cheng S, Chiang T. The Effect of Universal Health Insurance on Health CareUtilization in Taiwan. J Am Med Association. 1997;278(2):89–93.
10 Einav L, Finkelstein A. Moral Hazard in Health Insurance: what we know and how we know it. J Eur Econ Assoc. 2018;16(4):957–82.
11 Han H, Lien H, Yang T. Patient cost-sharing and Healthcare Utilization in early childhood: evidence from a regression discontinuity design. Am Economic Journal: Economic Policy. 2020;12(3):238–78.
12 Nyman JA, Cagatay K, Dowd BE, et al. Decomposition of Moral Hazard. J Health Econ. 2018;57:168–78.
13 Li Y, Li L, Liu J. 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.
14 Wolinsky A. Competition in a market for informed experts' services. RAND J Econ. 1993;24(3):380–98.
15 Currie J, Lin W, Zhang W. Patient knowledge and antibiotic abuse: evidence from an audit study in China. J Health Econ. 2011;30(5):933–49.
16 Eliason PJ, Grieco P, McDevitt RC, et al. Strategic Patient Discharge: the case of long-term care hospitals. Am Econ Rev. 2018;108(11):3232–65.
17 Wu B. Physician Agency in China: evidence from a medicine-percentage incentive Scheme. J Dev Econ. 2019;140:72–89.
18 Roberts ET, Chernew ME, McWilliams JM. Market share matters: evidence of Insurer and Supplier Bargaining over prices. Health Aff. 2017;36(1):141–8.
19 Finkelstein A. The Aggregate effects of Health Insurance: evidence from the introduction of Medicare. Quart J Econ. 2007;122(1):1–37.
20 Lu F. Insurance Coverage and Agency problems in Doctor prescriptions: evidence from a field experiment in China. J Dev Econ. 2014;106:156–67.
21 Klick J, Stratmann T. Diabetes treatments and Moral Hazard. J Law Econ. 2007;50(3):519–38.
22 Dave D, Kaestner R. Health Insurance and Ex Ante Moral Hazard: evidence from Medicare. Int J Health Care Finance Econ. 2009;9(4):367–90.
23 Yilma Z, van Kempen L, de Hoop T. A perverse 'net' effect? Health Insurance and Ex-ante Moral Hazard in Ghana. Soc Sci Med. 2012;75(1):138–47.

Publisher's note