The London School of Economics and Political Science

Essays on Evaluating Payment Reforms in China

Xiaozhou Zhao

A Thesis Submitted to the Department of Health Policy of the London School of Economics and Political Science for the Degree of Doctor of Philosophy. London, September 2023

Declaration of Authorship

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 37700 words.

Statement of co-authored work

I confirm that I was the leading author of all the chapters of my thesis. Other than the chapter of empirical studies, Prof. Andrew Street (LSE) and Dr. Rocco Friebel (LSE) provided general supervision, writing assistance, technical editing, language editing, and proofreading for Chapter 1, 2, 3 , 4 and 8. For the chapters of empirical studies, Chapter 5 (i.e., The effects of GB on patient payments and health outcomes in Chengdu) was co-authored with Prof. Andrew Street, Dr. Rocco Friebel and Prof. Shaoyang Zhao (Sichuan university), under the supervision of Prof. Andrew Street and Dr. Rocco Friebel, we jointly developed the research idea, Prof. Shaoyang Zhao provided data for this chapter, and I undertook all the data estimations and wrote majority of sections of the chapter. Overall, I contributed 80% of the work. Chapter 6 (i.e., The effects of PPS on patient payments and health outcomes in Chengdu in 2011) was co-authored with Prof. Andrew Street, Dr. Rocco Friebel and Prof. Shaoyang Zhao, under the supervision of Prof. Andrew Street and Dr. Rocco Friebel, we jointly developed the research idea, Prof. Shaoyang Zhao provided data for this chapter, and I undertook all the data estimations and wrote majority of sections of the chapter. Overall, I contributed 70% of the work. Lastly, Chapter 7 (i.e., The effects of PPS on patient payments and health outcomes among the frail elderly in Chengdu in 2018) was co-authored with Prof. Andrew Street, Dr. Rocco Friebel and Prof. Shaoyang Zhao, under the supervision of Prof. Andrew Street and Dr. Rocco Friebel, we jointly developed the research idea, Prof. Shaoyang Zhao provided data for this chapter, and I undertook all the data estimations and wrote majority of sections of the chapter. Overall, I contributed 70% of the work.

Abstract

Since the New Health Reform in 2009, China initiated the health payment reform, proposing to actively promote new payment methods other than Fee-for-Service (FFS), including Global Budgets (GB) and Prospective Payment Systems (PPS), to control rising health care costs and reduce over treatment. However, there are potential difficulties in developing new payment methods in China. First, the new payment methods are designed with relatively simple payment mechanisms at the initial stage of payment reforms, and may not be the same effective on cost control. Second, cost control may cause a decrease in the quality of health services after payment reforms. It is therefore important to investigate the changes in expenditure as well as the change in patients' health outcomes after payment reforms in China, and provide empirical evidence for improving the new payment methods. This dissertation focuses on investigating the effects of different payment reforms on patients and hospitals in China. The dissertation provides a theoretical framework to predict the changes in bills, length of stay (LoS) for patients and service volumes for hospitals after replacing fee-for-service (FFS) with global budget (GB) and to predict the changes in bills, out-of-pocket payments (OOPs), LoS and patient health outcomes after replacing FFS with prospective payment system (PPS). Then the dissertation conducts three empirical studies on the GB and PPS reforms in Chengdu in China. The first empirical study investigates the GB effect on bills, LoS and the number of patients, and the results indicate that GB reform in Chengdu has no initial effect either on bills, LoS or service volumes. The second empirical study evaluates the PPS reform in Chengdu in 2011, and finds that bills, OOPs and LoS decrease after the reform, while the patients' health outcomes are worsened after the reform. The third empirical study investigates effect of the PPS reform in Chengdu in 2018 on the frail elderly patients, and finds that bills decrease and LoS increase for the frail elderly patients after the reform. However, these findings may not be very reliable due to the limited sample size of the data.

Acknowledgements

I would like to express deep gratitude to my lead supervisor Prof. Andrew Street and my supervisor Dr. Rocco Friebel, who have provided the invaluable and crucial guidance on developing my dissertation, given comprehensive suggestions on improving my dissertation to a higher level, and spared no effort to motivating me to develop my research abilities. I am extremely grateful to their patience and support throughout my doctoral study, especially when I met difficulties with completing my dissertation. I am indebted to Prof. Shaoyang Zhao in Sichuan University, who provided data sources for my dissertation, I could not complete the dissertation without the data from him. I am also thankful to Prof. Alistair McGuire and Dr. Mylene Lagarde for their comments in my Major Review Document.

I would like to thank the Department of Health Policy of LSE, who provided us with the great platform and cherished opportunities to gain our abilities. I would like to thank the organizers of HP500 class, I learned a lot on undertaking health-related studies form this class. I also would like to thank my fellow PhD students in Department of Health Policy, I was greatly inspired from the communication with them.

The special thanks go to my family, they have always supported me through all the aspects of my life, I could not get through the hard moments or complete my doctoral study without their unconditional love and support. And I am thankful for my friends, who have always provided companion and help to me.

Contents

List of tables

List of tables (Appendix)

List of figures

List of figures (Appendix)

List of abbreviations

AA	Acute appendicitis
ADRG	Adjacent Diagnosis Related Groups
AM	Acute mastitis
BA	Before-after
BPH	Benign Prostatic Hyperplasia
BOT	Benign Ovarian Tumor
CBA	Controlled before-after
CCDT	Chinese Clinical Disease Terminology
CCHI	Chinese Classification of Health Interventions
CHS-DRG	China Healthcare Security Diagnosis Related Groups
CI	Confidence interval
CNY	Chinese Yuan
DDD	Tripple differences
DiD	Difference-in-Difference
DIP	Diagnostic Intervention Packet
DRG	Diagnosis Related Group
FFS	Fee-For-Service
GB	Global Budget
HIVD	Herniated intervertebral disc
ICD codes	International Classification of Diseases codes
IR-DRGs	International Refined DRGs
ITP	Idiopathic Thrombocytopenic Purpura
ITS	Interrupted Time Series
LoS	Length of stay
MDC	Major Diagnostic Categories
NRCMS	New Rural Cooperative Medical Scheme
OOP	Out-of-pocket payment
OR	Odds ratio
P4P	Pay-for-Performance
PPS	Prospective payment
PSM	Propensity Score Matching
RW	Relative weight
SPC	Spontaneous pneumothorax (conservative treatment)
SPS	Spontaneous pneumothorax (surgical operation)
SMD	Standard mean difference
TC	Treatment-control
TEH	Thrombotic external hemorrhoids
UC	Ureteral Calculi
UF	Uterine fibroids
UEBMI	Urban Employee Basic Medical Insurance
URBMI	Urban Resident Basic Medical Insurance
US	The United States
USD	United States Dollar

List of characters in equations

b-The bill of the patient

e- The proportion of out-of-pocket payment to bill for the patient (i.e. the OOP:bill ratio under FFS)

f-The fixed amount of OOP paid by the patient (i.e. the OOP under PPS)

v-The utility function of the representative patient

h-The health outcome of the patient

H_a- The average health status of the population in the area

H- The population health outcome of the hospital

LoS-The length of stay of the patient

m-The prospective price paid by the payer per patient to the hospital under PPS (i.e. the PPS payment paid by the payer) which is a fixed amount

OOP-The out-of-pocket payment paid by the patient

p-The per patient reimbursed payment paid by the payer to the hospital

P_a-The average reimbursed payment of the population in the area

P_0- The reimbursed payments for the past year for all the hospitals in an area

p_{GB}-The per-patient GB reimbursed payment paid by the payer to the hospital (i.e. the implicit price under GB)

Q-The number of patients treated by the hospital in the present year

Q_0-The number of patients treated by the hospital in the past year

r-The share of the bill paid by the payer to the hospital (i.e. the reimbursement:bill ratio under FFS)

u-The utility of the hospital for treating the representative patient

u_{FFS}-The utility of the hospital for treating the representative patient under FFS

u_{PPS}-The utility of the hospital for treating the representative patient under PPS

u_{GB}-The utility of the hospital for treating the representative patient under GB

U_{GB}-The utility of the hospital for treating all its patients under GB

W- The utility function of the payer

z- The elements other than LoS (e.g. the medicine on the prescription and the surgical treatments) that have effect on the bill, health outcome and treatment cost for the patient.

η-The payer's degree of altruism

λ- The degree of altruism of the hospital

π-The profit of treating the patient for the hospital

1. General background

Summary

This dissertation evaluates hospital payment reforms in Chengdu, China. This opening chapter provides a general background of payment reforms in China, starting by introducing the New Health Reform and pointing out that one of the key components of this reform is hospital payment reform. The chapter then covers hospital payment methods in China and then, specifically, the payment reforms in China and Chengdu.

1.1 Introduction to health reforms in China

Since 1998, China started to construct the universal health insurance system, which guaranteed all residents in China basic health insurance, by creating three basic health insurance schemes: Urban Employee Basic Medical Insurance (UEBMI), Urban Resident Basic Medical Insurance (URBMI) and New Rural Cooperative Medical Scheme (NRCMS). In 1998[1], the establishment of UEBMI and URBMI started the construction of China's health insurance system, which ended the era of free medical care system developed by planned economy. UEBMI targeted employees of all the sectors in urban areas, while URBMI targeted the unemployed, children, students, and the disabled in urban areas. Subsequently, in 2004, NRCMS was implemented nationwide, covering rural residents.

Then, in 2009, the State Council of China initiated New Health Reform, aiming at establishing a suitable health system for the nation and achieving the goal of universal coverage of health services. New Health Reform included three major reforms: a universal health insurance plan, a basic drug system reform, and a payment reform, designed to improve China's health system from the demand side and supply side. These major reforms in China's health system are listed in Figure 1.

From the demand side, the New Health Reform was aimed to develop the three main insurance schemes and achieve universal insurance coverage. Under the New Medical Reform, the government claimed that the coverage rates of the basic health insurance would exceed 90% by 2011. In 2010, NCRMS covered 97% of the rural population, URBMI covered 93% of the target population, and the coverage of UEBMI reached 92% (Yip et al., 2012, Meng and Tang, 2013, Yu, 2015, Zhang, 2010). Evidence shows that the goal of the universal health insurance plan has almost been achieved. According to the report of United Nations Children's Fund, in 2013, over 95% residents in China have health insurance) and contributes to a higher propensity for households to seek health care service and increases the volume of health care services provided to the

[1] Before 1998, the health security system consisted of the labor insurance supported by public sectors and free medical care supported by fiscal expenditure, as a product of planned economy.

population of China (Wagstaff et al., 2009).

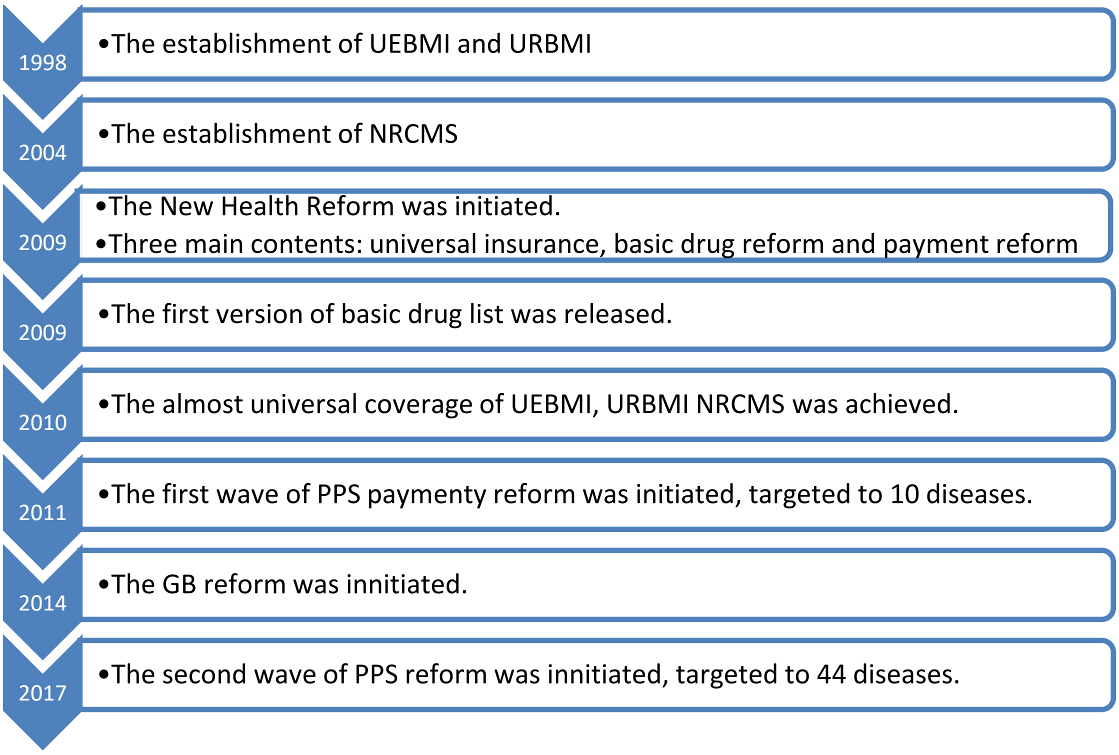


Figure 1. The major reforms in China's health system

At the same time, the expansion of health insurance coverage led to a rapid growth of health expenditure in China. The annual rate of health expenditure per capita in China had risen to 13% from 1990 to 2009 (Feng et al., 2015). To control the rapidly growing health expenditure for patients, two supply side policies, the basic drug system reform, and a payment reform, were introduced.

Basic drugs refer to drugs that can meet the patients' basic health needs, can be supplied at the reasonable price, and have a sufficient supply. In the New Health Reform in 2009, China released the first version of a basic drug list, which included 307 drugs. By 2018, this catalogue had expanded to 685 drugs. After the implementation of the basic drug reform, the drugs on the basic drug list were subject to centralized procurement, unified distribution, quality supervision and performance evaluation. Moreover, these drugs are sold with zero price difference among all the health institutes, with a subsidy provided by the health department.

However, establishing the basic drug list proved insufficient in addressing all root causes of increasing health expenditures in China, with a subsequent payment reform targeting health care providers to promote cost control.

Fee-For-Service (FFS) is a retrospective payment method that had been adopted as the main health payment method for hospitals in China since the 1980s. Under FFS, patients receive medical service and then pay by the unit of the service, which has been linked to rapid increase in health expenses. Specifically, FFS provides an easy way to settle accounts but has the disadvantage that physicians are stimulated to supply over-treatment for higher profit (Baker, 1999). Thus, due to both high market power and this type of payment method, high medical cost became a major problem of China's health system (Feng et al., 2015).

Under the New Health Reform, China began to explore payment methods other than FFS. In 2011, the Ministry of Human Resources and Social Security of the People's Republic of China issued the official document of the payment reform, proposing to actively promote new payment methods, including Global Budgets (GB) and Prospective Payment Systems (PPS), to control rising health care costs and reduce over treatment. Meanwhile, capitation was also adopted in a small range of cities (e.g., Changde, Qingdao and Taizhou) during the payment reform. This dissertation will focus on evaluating the effects of the move from FFS to either GB or PPS.

1.2 The health payment methods in China

This section introduces the payment methods of FFS, GB and PPS, involving a comparison of these three methods across the dimensions of cost control, activity levels and health service quality.

1.2.1 Introduction to FFS, GB and PPS

(1) FFS

Fee For Service (FFS) is a typical method of retrospective payment. Under FFS, payments are made according to the units of the health service consumed at individual patient level. Everyone is described specifically, there is the item for service or per diem payment, with the bill being the sum of prices for all units of services. In effect there is a contract for each individual patient, which can be regarded as a bill that includes itemized list of services and the price per service, and the total payment for the patient equals the price per service multiplied by the quantity of services. Under FFS hospital revenue equals the number of patients multiplied by the fee for each patient. The bill for each patient is largely determined by the hospital, particularly in deciding the volume of services to provide and, sometimes, what fee to charge for each service. Sometimes insurers use a price list, specifying the fees for each service, to impose some cost control.

(2) Global budget

Under the system of global budget, the government estimates the annual funds for a certain area according to the number of insured persons, the expected total number of annual visits and the average cost of visits in that area. Then, at the beginning of a year, the funds are distributed to each designated health institution. The total payment obtained by each designated health institution is generally calculated and determined by the expected number of patients and the expected price for treating a typical patient in that institution in previous years. However, in China, there are a large volume of patients who seek medical treatment across regions (He et al., 2019, Liu and Wang, 2024), and the composition of patients in a hospital varies in each year, thus, the predictions may not match reality.

(3) PPS

Under PPS, the medical insurance department sets the tariffs for each designated hospital according to the tier of the hospital, develops the reimbursement standard for different diseases, and then makes a settlement with the hospitals in accordance with the actual number and type of patients treated by the hospitals. According to the contract under PPS, the activity is described by Diagnosis Related Group (DRG), the price per patient is fixed in advance, and the price is not determined by each hospital but based on costs across all the hospitals. The hospitals become price takers under PPS, and the expenses that exceed the tariffs will be paid by the hospitals. The revenue for a hospital equals the actual number of patients in each DRG multiplied by the prospective price per patient of DRG, summed across all DRGs.

DRGs were first implemented for hospital payment purposes in 1983 for the elderly medical insurance enrollees in the United States (US), and the tariffs were made according to the diagnosis classification. Medical costs of hospitals may be higher or lower than DRG tariffs. If costs exceed the tariffs, hospitals need to cover the difference, but if costs are below the tariff, the hospitals profit from the surplus. In fiscal year 1984, the initial year for implementing PPS in the US, the increase in real payments for inpatient hospital care for Medicare beneficiaries was only 3.8%, compared to an average annual increase of 10.0% from 1973 to 1982 (Guterman and Dobson, 1986).

Limitations that the first version of DRG in the US did not consider were the complexities in costs brought by the individual characteristics of patients (e.g., severity degree of patients and complication status) and that hospitals might refuse treatment of the severe patients due to under-compensation. In response, Yale University grouped the diseases based on the International Classification of Diseases (ICD) codes, categorized the severity degree of diseases and complication status and developed the second generation of DRGs (Averill, 1994). Then being based on the third generation and fifth generation DRGs, 3M information system developed the International Refined DRGs (IR-DRGs) that could be applicable to different countries, with IR-DRGs being implemented in the US for the first time in 2000 (Mullin et al., 2002).

1.2.2 Comparison of FFS, GB and PPS

This section compares FFS, GB and PPS across the dimensions of cost control, increase activity and health service quality and discusses the pros and cons of both methods as shown in Table1.

Table 1. The comparison of FFS, GB and PPS

	FFS	GB	PPS
Cost control	Weak	Strong	Moderate
Increase activity	Strong	Weak	Moderate
Quality	Strong	Weak	Moderate

Cost control

Under FFS, since the price is determined by the amount of service, the marginal cost of treating more severe illness is still reimbursed.. Thus, under FFS, the hospitals can provide as many services as they can build up for patients, which causes over supply and excess expenditure.

Under GB, a fixed amount of money made through retrospective price-setting mechanism, all services delivered during the year are contained within that sum, no matter however many patients are treated in that year. Thus, GB has a strong incentive for hospitals to reduce costs.

Under PPS, being faced with the elimination of the reimbursement for the marginal cost, hospitals might take cost sharing behaviors and reduce services. When it comes to PPS, the service price is the same for everyone in the same DRG, the reimbursement for the marginal cost is eliminated (Ellis and McGuire, 1986). Also, PPS brings Yardstick competition for hospitals (Shleifer, 1985): The prospective DRG price convert the role of hospitals from price maker to price taker, and subject hospitals that previously enjoyed local monopoly market power to competitive pressure, with hospitals competing to treat patients under a given price. Thus, PPS has a stronger power of cost control for each patient compared with FFS. In China, PPS stimulates hospitals to reduce health costs, especially expensive and unnecessary drug prescriptions, to resolve the problem of induced demand for drugs caused by promoting the policy of 'raising hospitals by drug'[2].

Increase activity

Under FFS, the total revenue of the hospital increases with the volume of patients and

[2] In China, for motivating physicians to pay more effort, the government has allowed for the drug mark-up coupled with FFS reimbursement which is at a 15% margin on drug payments (Yip and Hsiao, 2009; Yip et al, 2010). However, the drug mark-up increases the providers' reliance on drug revenue and the financial incentive causes physicians induce patients' demand for drug (Yi et al, 2015). During the New Health Reform in 2009, Chinese government announced that the drug mark-up should be gradually eliminated in public hospitals. Since then, the 'Zero Mark-up Drug Policy' (ZMDP) has been implemented as the pilot project in the primary health care institutions in some provinces. In our study, we do not need to consider the effect of ZMDP: This project was initiated in Sichuan province in 2017, before that, ZMDP was only implemented as the pilot project among several randomly selected primary health institutions in rural areas (Lian, 2015; Yang et al, 2017). And our study focuses on the secondary and tertiary hospitals in Chengdu (the capital city of Sichuan province) from 2010 to 2014, the samples are not affected by ZMDP.

the quantity of services provided. Consequently, hospitals have an incentive to treat more patients and provide more services to increase their revenues. The increasing activities can result either from increasing volume of patients or increasing units of services.

Under GB, there is limited incentive for hospitals to do more work. Due to the lack of inspection to providers' behavior, hospitals may even refuse severe patients whose costs are higher than others to not exceed the total budget. If so, the number of visits in a hospital will be lower than under FFS.

Under PPS, total revenues for the hospital increases with the number of patients, thus hospitals have an incentive to treat more patients. However, under PPS, since the price for each group of patients is the same across all the hospitals, the incentive for hospitals to treat additional patients under PPS is not as strong as under FFS.

Health service quality
If health service quality is defined as 'do more for patients', then quality is high under FFS: by definition, high quality care is achieved by providing as high quantities of services.

Under GB, hospitals are faced with a budget constraint. Thus, if provision of quality is costly, GB provides weak motivation for hospitals to maintain the health service quality when being faced with financial constraints.

Under PPS, because hospitals are paid the same price for patients in each DRG, the hospitals are faced with this budget constraint when treating patients. To avoid the cost of treatment exceeding the DRG price, hospitals might decrease the quantities of services, which consequently may lead to a lower quality compared to FFS.

1.3 The payment models in China

(1) FFS in China

There are three agents in the health payment process: the patient, hospital, and payer. In China, the payer in each city is the local health sector named as Healthcare Security Administration in that city, as the government sector. The local health payers in China play an important role in paying hospitals as well as monitoring the quality of health service provided by hospitals.

If a patient who is enrolled in any basic health insurance (either URBMI, UEBMI or NRCMS) in China, the patient will be provided with a social security card linked to the health insurance account. The discharged patient pays an out-of-pocket payment (OOP) via the patient's personal bank account/cash as a portion of the total cost of treatment to the hospital. Exemptions from the OOH payment are verified by showing the patient's health insurance account in the social security card. Then the hospital prepays the insured payment as the remaining portion of the copayment for the patient, and the payer reimburses the insured payment to the hospital.

In China, the payer set a deductible for discharged patients at approximately CNY1000 for tier 3 hospitals, CNY500 for tier 2 hospitals and CNY200 for tier 1 hospitals. The payment under the deductible should be paid completely by the patient. As to the amount above deductible, the payer pays a fraction while the patient pays the other fraction. This fraction paid by payer is approximately 70% for tier 3 hospitals, 80% for tier 2 hospital and 90% for tier 1 hospitals. The deductibles for each tier of hospitals varies among different cities, the type of health insurance and varies among different groups of patients in China, which are higher for the cities with lower income, higher for the URBMI and NRCMS beneficiaries (i.e., the urban beneficiaries as children, students, the retired and the elderly and the rural beneficiaries) relative to UEBMI beneficiaries (i.e., the beneficiaries who are urban employees of the formal sectors) and higher for some special groups (e.g., the disabled and the households who receive subsistence allowance).

Thus, in China, under FFS, the OOP of a patient includes the sum of deductible and a fraction of the amount that is in excess of the deductible.

(2) GB in China

Under GB, the OOP method between the health insurance enrollee patient and the hospital follows the pattern of FFS, therefor the bill for the patient under GB is the same as the one under FFS, which lists all the services for the patient and the price per service, and with the patient only paying for OOP to the hospital. However, the reimbursing method between the payer and the hospital has changed: under GB, the payer's reimbursement to the hospital no longer covers the full amount of insured payment for every patient, but only reimburses a fixed amount for the total number of services (i.e., the GB payment) prospectively to the hospital by year, and the GB payment equals the average health payment for that type of hospitals in one city in the previous year multiplied the service quantity of this hospital in the previous year.

(3) PPS in China

Under PPS, the OOP method between the health insurance enrollee patient and the hospital also follows the pattern of FFS, while the reimbursing method between the payer and the hospital has changed: under PPS, the payer's reimbursement to the hospital no longer covers the full amount of insured payment for a patient, but only reimburses the hospital a predetermined amount for the service received by a PPS targeted patient, and this predetermined amount for a particular service is derived according to the classification system of that service.

1.4 The payment models in Chengdu

The payment model in Chengdu is the same as the payment model in China Chengdu Healthcare Security Administration and Chengdu Municipal Health Commission are two main sectors of the government in charge of health service supply in Chengdu. The main functions of Chengdu Healthcare Security Administration include: designing and improving the local health insurance schemes of UEBMI, URBMI and NRCMS, managing the local health insurance fundings, implementing the local health insurance and health institution reforms, signing the health insurance payment contract with local health institutions, setting the health service price and monitoring health services quality of the contracted health institutes and procuring the drugs for contracted health institutions pharmaceutical companies. The main functions of Chengdu Municipal Health Commission include coordinating with Chengdu Healthcare Security Administration on implementing local health institution reforms and assessing the performance of local health institutions, designing the local plan for disease prevention and control, designing the local health care policies for aging population and responding to local public health emergencies.

In 2018, there are 2948 health institutions contracted with Chengdu Healthcare Security Administration, including all the tier 3 and tier 2 hospitals and several primary health care institutions. Being based on the health insurance payment contract, the relationship between Chengdu Healthcare Security Administration and the hospitals in Chengdu can be characterized as a principal-agent relationship as will be set out in Chapter 2.

1.5 The payment reforms in China

1.5.1 The GB reforms in China

There are two alternative methods to designing GB payments. The first is by employing an expenditure target, which sets up the goals for the current and, possibly, subsequent year. And the second is an expenditure cap, which sets up a fixed amount of money made through retrospective price-setting mechanisms, with all services delivered during the year being contained within that sum.

In China, the GB reform followed a pattern of expenditure cap. The GB payment to one hospital equals the average health payment for that type of hospitals in one city in previous year multiplied by the service quantity of the hospital in previous year. The GB reforms were initiated in China from 2011, and was introduced to Shanghai, Zhenjiang, Jinhua, Chengdu, and other cities.

Implementation of the GB in China followed a ‘soft’ target, where hospitals could be reimbursed partially if expenditures were incurred above the benchmark under some special conditions: when, 1) the number of ‘expensive’ patients of the hospital in that year was higher than the number in last fiscal year; 2) the price of medical consumables or health services increased; 3) the tier of a hospital increased; 4) the structure of patients changed (e.g., the urban enrollees increased); and, 5) the bed utilization exceeded 100%. Interestingly, there was neither official statistics describing the proportion of patients under the special cases nor how often these special cases occurred. The hospitals were able to claim reimbursement for the expenditures over benchmark, the amount of this reimbursement varied for each hospital (Meng, 2015, Li, 2018, Zhao, 2015). Thus, the 'soft' target may not be an efficient tool for cost containment, since the hospitals do not have strong incentives to decrease costs (Chen and Fan, 2016).

To overcome this limitation of the basic GB payment and to motivate stronger competition among hospitals, China elevated the basic GB payment to the regional GB payment. Compared with the original form GB, regional GB had three main features: 1) the GB funds were no longer paid to each hospital, but paid to a region, and were then shared out among the hospitals in that region; 2) the amount of GB funds to a hospital was decided by the market share of this hospital in the region; 3) there was also a grading system on hospital performance as an assessment packaged with the regional GB, and the amount of GB fund to a hospital was also related with the score of the hospital in this assessment. Then, every hospital needed to provide healthcare service at a lower price and at higher quality, not only for the cost containment pressure from GB adoption, but also for attracting more patients and striving for a higher share of the regional GB fund for itself.

Then, to provide a refined management on hospital performance with considering the patient case-mix groups, being based on the regional GB, China created the Diagnostic Intervention Packet (DIP) system in 2018. Under DIP, a Relative Weight (RW) was calculated for every case-mix group, and a point value for the total RWs of all the case-mix groups in a city was calculated as follows:

$$point\ value = \frac{Total\ \ heath\ payment\ in\ the\ city}{\sum_{i=N} RW_i * Volume_i} \quad (1)$$

In Eq. (1), RW_i was RW for case-mix group i, and $Volume_i$ was the number of patients for case-mix group i in a city, and there were N case-mix groups in the city. Thus, $\sum_{i=N} RW_i * Volume_i$ was the total RWs for patients all the case-mix groups in the city, and the point value could be understood as the average value for every unit of RW.

Then, the DIP payment for hospital j of type k was:

$$DIP_{jk} = point\ value * [\sum_{i=n}(RW_i * Volume_{ijk})] * adjustment_k \quad (2)$$

In Eq. (2), $Volume_{ijk}$ was the number of patients for case-mix group i in hospital k of type k, and the number of case-mix groups in hospital j of type k was n. $\sum_{i=n}(RW_i *$

$Volume_{ijk}$) was the total RWs for patients of all the case-mix groups in hospital j of type k. Under the DIP system, there were two types of hospital: tier 2 and tier 3 hospitals. By considering that the average payments of the tier 3 hospitals were normally higher than tier 2 hospitals, the adjustment factor was included, which equaled 1 for hospitals and equaled 0.7 for tier 2 hospitals.

In 2020, the National Health Care Security Administration officially initiated the DIP reform, with the DIP system adopted among 71 cities in this wave of reform.

So far, there has been a large volume of studies on evaluating GB effects in China, most studies suggested that GB was effective in cost control. The success of cost control was observed among the cardiovascular patients in Shanghai (Dou et al., 2019), UREBMI outpatients of hypertensive in Tianjin (Huang et al., 2016), the patients from a secondary hospital in Yunnan province (Guan et al., 2020), 91 hospitals in Shunyi district in Beijing (Xie et al., 2015), AMI patients in Zhongshan in Guangdong province (Yuan et al., 2019), the UEBMI funding in two districts in Foshan in Guangdong province (Liang et al., 2012), stroke patients randomly collected from the China Health Insurance Research (Yang et al., 2022), NRCMS patients from Gaolan county in Gansu province (Wang, 2015). On the other hand, the evidence from NRCMS inpatient in Weiyuan county in Gansu province (Li et al., 2018) and the patients in a dental hospital in Beijing (Wu and Guo, 2016) found that cost control was ineffective.

In this dissertation, by focusing on the GB reform in Chengdu in 2013, we develop the theoretical framework, describing the hospitals' reaction to PPS reform, and making predictions about bill, OOP and the number of patients. Then we adopt an empirical strategy to evaluate the policy effect on bills, OOP, LoS and the number of patients. So far, the studies that investigate GB effect on service volumes are quite limited, we contribute to evaluate GB effect on the number of patients in Chengdu to provide the empirical evidence of the change in hospital service volume after GB. Moreover, rather than focusing on a certain group of patients, we focus on all the tier 3 and tier 2 hospitals in Chengdu, which makes our empirical findings be more representative.

1.5.2 The PPS reforms in China

1.5.2.1 The implementations of PPS reform in China

During the past decade, the PPS reforms followed the form of simple case-mix PPS payment in China, where each PPS reform targeted serval diseases rather than adopted DRGs. Local governments in different cities in China promoted PPS reforms by starting with simple case-mix DRG funding. Under these simple case-mix PPS reforms, medical insurance departments pay hospitalization expenses to designated health institutions according to the number of discharged patients under each targeted disease. In the first wave of simple case-mix PPS reforms in each city, only selected diseases that are easy for cost accounting were chosen as targeted diseases. After the first wave of simple case-mix PPS reform, with the enriched experience of cost accounting, local governments added more targeted diseases in other waves of DRG reforms and, as a result, the DRG categories gradually expanded. Since there are hundreds of waves of simple case-mix PPS reforms among all the policy implementation cities in China, and it is difficult to document the time point of each policy wave in each city. Even though it is hard to describe the process of increasing the targeted diseases for all the implementation areas, the study of (Peng et al., 2018b) suggests that as a result, simple case-mix PPS reforms have targeted over 100 diseases among 80% of policy implementation cities in China until 2018. According to the statistics of National Health Commission of China in 2019, 294 cities have initiated simple case-mix PPS reform and continued to expand their DRG categories. A large volume of studies focuses on the early stage of implementing PPS, all of which are evaluations based on simple case-mix PPS reform.

Moving on from simple case-mix PPS, BJ-DRG as the first version of DRG grouping system in China was completed in 2008. Since then, a limited number of cities (e.g., Beijing, several cities in Yunnan province) have applied other forms of DRG classifications (including BJ-DRG, C-DRG and CN-DRG) since their first wave of payment reform, while other areas started the reform with simple case-mix PPS, then expanding their DRG categories. Until 2018, according to a survey by Beijing Aden InfoTech Ltd., among 474 tier 2 and tier 3 hospitals in 15 provinces, only 21% of these hospitals have applied or plan to apply DRGs. After the design of China Healthcare Security Diagnosis Related Groups (CHS-DRG) was completed, this national DRG grouping system was planned to be adopted in 30 cities in the form of pilot policy since 2019.

1.5.2.2 The development of DRG grouping system in China

In China, PPS was gradually implemented as a pilot project across different provinces, including Zhenjiang (2001), Shanghai and Changsha (2004), Harbin (2010), Beijing and Chengdu (2011), and Yuxi (2016). A simple case-mix PPS does not consider the differences in costs brought by the individual characteristics of patients (e.g., severity degree of patients and complication status) other than those used in the construction of the DRGs themselves.

Building on simple case-mix PPS reforms, China began developing its own DRGs that are suitable for health service environment of the nation. The first local version of DRG grouping scheme was BJ-DRG, which was commissioned by the Beijing Medical Insurance Association. BJ-DRG included 108 diseases and was first applied to the evaluation of inpatient health service among public hospitals in Beijing in 2011. The characteristics of DRG allocation include the gender, age, complications, and other factors that can impact the patients' health costs.

To promote a nationwide version of DRGs, China established CN-DRG in 2014 and C-DRG in 2016, both of which built upon the BJ-DRG grouping scheme. Although CN-DRG and C-DRG share a similar grouping scheme they have a different grouping basis. C-DRG grouping is based on Chinese Clinical Disease Terminology (CCDT) and Chinese Classification of Health Interventions (CCHI), while CN-DRG grouping is based on ICD-10 and ICD-9.

The main difference among these three grouping schemes is the grouping basis, which may lead to varied grouping results of one disease under these three grouping schemes. There are three main steps of DRG grouping: 1) classifying the cases into different Major Diagnostic Categories (MDC); 2) classifying the same MDC into Adjacent Diagnosis Related Groups (ADRG), according to different procedures; and, 3) classifying ADRG into DRGs according to patients' characteristics (e.g., gender, age, and complications). Since these three grouping schemes use different terminologies, differences might occur in step 2. For instance, varicose veins are classified under the ADRG of peripheral venous disease in BJ-DRG, the ADRG of other venous disorders in CN-DRG, and the ADRG of venous disorders of lower extremities in C-DRG.

In October of 2019, the National Healthcare Security Administration published CHS-DRG, which replaced all previous versions as BJ-DRG, CN-DRG and C-DRG and became the official and standardized DRG grouping scheme across China.

So far, there has been a large volume of studies on evaluating PPS effects in China. However, as the studies focused on PPS reforms in different areas and different groups of patients, the policy effects found by the present studies were diversified. Most studies found PPS incentivize cost control (Jian et al., 2015a, Peng et al., 2021, Wu et al., 2022, Peng et al., 2018a, Yuan et al., 2019, Li et al., 2018, Li et al., 2019, Meng et al., 2022, Hu et al., 2023, Huang et al., 2022, Jian et al., 2015b), and serval studies found that cost control was associated with decrease in LoS (Jian et al., 2015b, Jian et al., 2015a, Li et al., 2018, Wu et al., 2022, Peng et al., 2018a, Meng et al., 2022, Huang et al., 2022). Only the study of Zhang (2010) which was among a random sample of inpatients from a tier 3 hospital in Shanghai and the study of Wei and Feng (2019) which was among patients from 4 hospitals in Guizhou provinces found PPS did not incentivize cost control, however, the samples of these studies were limited, their findings may not be generalisable. The findings on health service quality were more complicated. The studies among Acute Myocardial Infarction (AMI) patients in Beijing (Jian et al., 2019), AMI patients in Zhongshan (Yuan et al., 2019),AMI patients in Tianjin(Wu et al., 2022) found PPS had no effect on health outcomes, while the studies (Jian et al., 2015b, Hu et al., 2014, Zhang and Hu, 2015) among patients from various diseases in Beijing found worsened health outcomes after PPS.

In this dissertation, by focusing on the PPS reforms in Chengdu in 2011 and 2018, we develop the theoretical framework, describing the hospitals' reaction to PPS reform, and making predictions about bill, OOP, LoS and health service quality. Then we adopt an empirical strategy to evaluate the policy effect on bills, OOP, LoS and health service quality. In the study on PPS reform in 2011, we contribute to provide the comprehensive evaluation on PPS adoption in Chengdu, we investigate the changes in bills, OOP, LoS and health service quality at the same time by focusing patients of different diseases from various hospitals, which allow us to draw generalizable insights. In the study on PPS reform in 2018, by focusing on the frail elderly patients, we evaluate PPS effect on a special group. As China is improving the national DRG grouping scheme, an inescapable fact is that the heterogeneity between different patient groups should be considered when setting the PPS price, thus, we provide empirical evidence of PPS effect from a special group of patients.

Conclusion

This chapter introduces the existing payment methods that have been adopted across different geographies in China, namely FFS, GB and PPS, and it briefly illustrates three agents in the health payment process as the patient, the hospital, and the payer. It also highlighted their roles in the reimbursing process. The next chapter will introduce the theoretical framework of payment reforms in China and will describe the three-agent payment model in detail.

2. Theoretical framework

Summary
This chapter set out the theoretical framework of the payment reforms in China. It first introduces a three-agent setting that involves the patient, the hospital and the payer in the reimbursing process. Then it describes the payment models under FFS, GB and PPS. Next it explains the changes in the utility function of each agent when the payment method switches from FFS to GB or PPS. And finally, it predicts how bills, LoS and the number of patients might change following the GB and PPS reforms.

2.1 The general setting of the framework

There are three agents in the theoretical framework: the patient, the hospital, and the payer. The patient pays the out-of-pocket payment (OOP) as a portion of the bill to the hospital, and the payer pays the remaining portion of the bill (reimbursement payment) for this patient to the hospital. The payer is the Chengdu Healthcare Security Administration in China and plays a role in reimbursing the price of health services to hospitals as well as supervising on the quality of health service provided by hospitals. In the section of the comparison between FFS and GB, we focus on the relation between the hospital and the payer. In the section of the comparison between FFS and PPS, we focus on the relation between the patient, the hospital, and the payer.

The utility function of the representative patient (v) is stated in Eq. (3) and suggests that the patient cares about the OOP, the health outcome (h) derived from the hospitalisation and the length of stay (LoS):

$$v = v(OOP, h, LoS) \quad (3)$$

v is negatively related with OOP, while positively related with health outcome, which means a patient would prefer lower OOP and better health status after hospitalization. The relationship between utility and LoS is indeterminate because it influences utility both directly but also indirectly through an influence on OOP and health outcomes, as we shall explain.

Eq. (4) shows the *OOP* for the representative patient:

$$OOP = eb + f \quad (4)$$

The bill (b) contains the composition of all the units of health services that a patient has received and the total price of these health services. Under FFS, the *OOP* for a patient is a share of his total payment on the bill ($0 < e < 1, f = 0$). In other words, the *OOP* is related to the size of the bill which is a function of the number of services (The function will be introduced in Eq.5). Under PPS, the patient only pays the fixed amount of OOP to the hospital ($e = 0, f > 0$), and the OOP is no longer related to the quantity of health services consumed or size of the bill.

In Eq. (5) and Eq. (6), we introduce the bill for treating the patient (b) and the health outcome of the patient (h), both because some elements (e.g. hospitalization fee) of the bill are directly based on LoS and others (e.g. drug fee) are indirectly related to LoS. This means that LoS can be considered a proxy measure of treatment intensity for in-patients (Hayford, 2012, Hua et al., 2018). Thus, we define b as an increasing function of LoS and the other treatments received by the patients (e.g. the medicine on the prescription and the surgical treatments) that are related to the bill (z).

$$b = b(LoS, z) \quad (5)$$

The bill for treating the patient is positively related to the quantity of health services and is also positively related to LoS. We use the bill is an outcome variable, which is used to measure the total health payment in the empirical section of this thesis. Since we use LoS to measure the treatment intensity in the empirical section, we mainly focus on the relation between the bill and LoS.

$$h = h(LoS, z) \quad (6)$$

In the empirical section, we use the 30-day unplanned readmission of a patient as a measure for health outcome, and hence, as a measure of the benefit of treatment.

LoS has a negative and direct effect on the patient's utility since he doesn't want to stay in hospital longer than necessary and has indirect effect on the patient's utility through the relationship with the patient's health outcome because he needs to stay long enough to have recovered. Hence, there should be minimum LoS to guarantee recovery, but once LoS exceeds the minimum required, patient utility will be negatively related with LoS (Clarke and Rosen, 2001).

In Eq. (6), the patient health outcome (h) is also related with LoS and the other treatments received by the patients (z) that are related to the patient health outcome. Since we use LoS to measure the treatment intensity in the empirical section, we mainly focus on the relation between the 30-day readmissions and LoS. In the function of patient health outcome, there is an optimal value of LoS (LoS^*). When LoS is under the optimal value, h is positively related with LoS ($h' > 0$): longer LoS guarantees the patient to receive more treatment and is, therefore, more likely to have higher health outcome (Skinner, 2011). However, when LoS exceeds the optimal value, the patient outcome will not increase in LoS ($h' \leq 0$) and the longer LoS (i.e., over treatment) might have negative effect on patient health outcome.

We consider the utility function (W) of the payer in Eq. (7):

$$W = W(P_a, H_a) \quad (7)$$

Compared with price p and health outcome h of a representative patient, P_a and H_a represent the total health payment and total health status in the area. The payer makes payments to hospitals to settle the bill. The payer reimburses the hospital as a set of prices ($P_a = \sum_{i=1}^{I} p_i$) across all patients (i=1…I) for which it is responsible and monitors the quality of health service provided by the hospital. As part of the public sector of the Chinese government, the payer is interested in both financial budget and social health benefit. In the payer's utility function in Eq. (7), W is negatively related with P_a since the payer faces a constrained fiscal budget so prefers to pay lower prices to keep within the budget. W is positively related with $H_a = \sum_{i=1}^{I} h_i$ since the payer wishes to improve overall health status in the area

In Eq. (8), we linearize the utility function of the payer:

$$W = -(1-\eta)P_a + \eta H_a \quad (8)$$

η measures the payer's degree of altruism. A higher η indicates that the payer cares more about the average health status in the area and less about fiscal budget[3]. For a purely altruist payer, η would equal 1 since it only cares about average health status in the area. For a payer that only cares about fiscal budget, η equals 0. In practice, the payer needs to balance both objectives, suggesting 0< η <1 (Makris and Siciliani, 2013)

To ensure an affordable price and favourable quality of health services for patients, the payer makes the reimbursed payment p for any representative patient to hospitals according to Eq. (9):

$$p = rb + m \quad (9)$$

This can be regarded as a contract between the payer and the hospital. The amount paid by the payer per patient can be expressed as the sum of a fixed per-case payment (m) and a share of the bill (rb) after excluding the OOP. Hereby, m represents the average cost reimbursed (m is the same for the patients under the same diagnosis) and r represents the marginal reimbursement of the bill (Hodgkin and McGuire, 1994, Ellis and McGuire, 1986, Cutler, 1995). Under FFS, the total reimbursement of the bill is the patient's OOP and the payer's contribution, with the payer's proportion being a fraction of the bill ($m = 0, 0 < r < 1$). Under PPS, the payer only pays the PPS payment to the hospital ($m > 0, r = 0$), and the PPS payment is no longer related to the quantity of health services. Compared with PPS tariff which is determined by the average payments for the past years for each diagnostic group, the GB payment is calculated in a relatively imprecise way without considering the patient diagnostics.

[3] To incentivize hospitals to provide high quality care to patients, the payer adds the additional reward to hospitals according to patient's health outcome (which, in the context we consider, is measured by unplanned 30-day readmissions), and the way of providing the reward keeps the same under any type of payment system.

Finally, we switch to the perspective of the hospital. Applying the theories of Hodgkin and McGuire (1994) and Ellis and McGuire (1986), the hospital is not assumed to be interested solely in pure profit maximization, but also in the patient's health outcome. Thus, the utility of the hospital to treat the representative patient can be expressed as a function of profit (π) and patient health outcome (h) in Eq. (10):

$$u = u(\pi, h) \quad (10)$$

If we linearize the utility function of hospital in Eq. (10), the utility for the hospital for treating the representative patient is:

$$u = (1 - \lambda)\pi + \lambda h \quad (11)$$

u represents the hospital's utility and π reflects the profit that the hospital makes from treating the representative patient. h refers to the health outcome of the patient, with health outcome a function of LoS as specified previously in Eq. (6).

λ measures the hospital's degree of altruism. A higher λ indicates that the hospital cares more about the patients' health outcome and less about profit. For a purely altruist hospital, λ should equal to 1 since it only cares about the patient's health outcome. For a hospital that only cares about profit maximization, λ equals 0. Of course, the hospital may balance both arguments, caring about both profit and the health outcome of the patient, such that $0<\lambda<1$ (Makris and Siciliani, 2013).

In Eq. (12), under FFS, the profit of treating the representative patient equals the bill of treating the patient minus the actual cost c for treating the patient.

$$\pi = b - c \quad (12)$$

Eq. (13) and Eq. (14) outlines the properties of the profit function under FFS.

In Eq. (13), the actual cost is positively related with LoS, and related to the other treatments received by the patient (e.g., the medicine on the prescription and the surgical treatments), since we use LoS to measure the treatment intensity in the empirical section, we mainly focus on the relation between the cost and Los.

$$c = c(LoS, z) \quad (13)$$

Moreover, there is the mark-up other than the marginal cost reimbursement, as the payment for the physician's own input and effort in China, and the mark-up increases with LoS (Yip and Hsiao, 2009, Yip et al., 2010). Thus, in Eq. (14), the first order condition of the difference between bill and cost with respect to LoS should be positive if there is a profit margin for treating a patient.

$$(b - c)' > 0 \quad (14)$$

2.2 The comparison between FFS and GB

In China, when the payment method is shifted from FFS to GB, the change only impacts the relationship between the payer and the hospital, with patients paying the same OOP amount under both FFS and GB. Thus, the utility of the patient under GB should be in the same form as the utility under FFS, which has been introduced in Eq. (3) (4) (5) and (6). Consequently, we focus on the relationship between the payer and hospitals in this section.

Under FFS, the objective function of the hospital for treating a representative patient would be:

$$u_{FFS} = (1-\lambda)(b-c) + \lambda h \quad (15)$$

Taking the first order condition on u_{FFS} with respect to LoS yields:

$$u'_{FFS} = (1-\lambda)(b-c)' + \lambda h' \quad (16)$$

In Eq. (16), $(1-\lambda)(b-c)'$ represents the income effect: under FFS, there exists a profit margin on LoS for treating the patient, thus, the hospital has an incentive to increase LoS to maximise profit. Alongside the profit motive, $\lambda h'$ represents the hospital's altruism effect: the altruistic hospital might increase LoS due to its own concern about the patient's health outcome to guarantee the recovery, but the altruistic hospital would not keep increasing LoS beyond the optimal LoS* after which the patient gains no further health benefit from remaining in hospital.

Under GB the payer reimburses the hospital according to a total fixed amount of budget determined prospectively at the start of a fiscal year rather than according to the total actual costs of providing treatment to patients. In contrast, payments are made retrospectively under FFS. For the payer, the utility function is the same as Eq. (8), where $W = -(1-\eta)P_a + \eta H_a$.

Under GB, the reimbursed payment from the payer to the hospital for the representative patient (i.e., the implicit price under GB) can be expressed as:

$$p_{GB} = \frac{Q_0 P_0}{Q} \quad (17)$$

The global budget for the hospital equals $Q_0 P_0$, a per case GB payment for the representative patient which is the average of bills for patients in the past year for all the hospitals in an area (P_0) multiplied by the number of patients treated by the hospital in the past year (Q_o).

The objective function of the hospital to treat the representative patient is:

$$u_{GB} = (1-\lambda)(\frac{Q_0 p_{GB}}{Q} - c) + \lambda h \quad (18)$$

According to Eq. (18), the hospital can change the profit for treating a patient by either changing LoS or changing the number of patients in that year (Q), which means the hospital can reduce the cost for treating a patient by reducing the treatment intensity, or the hospital can reduce the number of patients in that year (decreasing Q) to ensure the higher per case budget ($\frac{Q_0 p_{GB}}{Q}$).

We take first order condition on u_{GB} with respect to LoS:

$$u'_{GB} = -(1-\lambda)c' + \lambda h' \quad (19)$$

Compared with u'_{FFS}, u'_{GB} takes a smaller value, which indicates that hospitals have a lower motivation to increase LoS under GB compared to the hospitals under FFS. This result is consistent with the aim of GB adoption to control health care costs. Moreover, if the effect of cost control ($-(1-\lambda)c'$) is larger than the effect of altruism ($\lambda h'$), u'_{GB} could even be negative. In this case, the hospital will reduce LoS and might even provide insufficient care to the patients, subsequently leading to possible adverse health outcome for the patient.

Under GB the objective function of the hospital to treat its patients becomes:

$$U_{GB} = (1-\lambda)(\frac{Q_0 p_{GB}}{Q} - c)Q + \lambda H \quad (20)$$

Total profit amounts to the profit for the average patient multiplied by total patients (Q) in that year. H is the health outcomes for all the patients from the hospital. It is noteworthy that the average health outcome for the population in the area (H_a) should be the aggregation of patient health outcome in the hospital (H), and patient health outcome in the hospital should be the aggregation of the patient health outcome (h): the altruistic hospital should increase LoS to the optimal value for each patient to guarantee his recovery and to improve the patient health outcome in the hospital.

Thus, at hospital level, if the hospital reduces the LoS for each patient, the total health income for the hospital will decrease. And if the hospital provides insufficient care to each patient and leads to adverse health outcomes for him, the health outcome of that hospital will become worse (i.e., the health service quality of the hospital will decrease) as a result.

Taking the first order condition on the utility of the hospital to treat all its patient under GB with respect to Q yields:

$$U'_{GB} = -(1-\lambda)c \quad (21)$$

In Eq. (21), U'_{GB} has a negative value, which indicates that under GB, hospital utility decreases the more patients are patients treated. We can regard $\frac{Q_0 p_{GB}}{Q}$ as the actual payment to the hospital for treating a patient, given the fixed GB payment ($Q_0 p_{GB}$), the decrease in the number of patients (Q) can increase the value of p. In other words, under GB, if the hospital decreases the number of patients, given the fixed total number of GB payment, it will have a higher budget for treating each patient.

Then we have several hypotheses relating to the behaviour of hospitals when payment is shifted from FFS to GB.

Hypothesis 1: The bills will decrease after GB adoption, since the hospital has the motivation to decrease the bill for each patient and ensure the total health expenditures in the hospital do not exceed the global budget. In the empirical section, the average bills for patients in a hospital in a year is calculated by the quotient of the total income from the patients divided by the total number of the patients in that hospital.

Hypothesis 2: The GB policy will result in a decrease in average LoS for the hospital because the hospital tends to reduce treatment costs by reducing treatment intensity. In the empirical section, the average LoS for patients in a hospital in a year is calculated by the quotient of the total number of occupied bed days in a year divided by the number of discharged patients in a year in that hospital.

Hypothesis 3: The GB policy will lead to a decrease in the number of patients treated by the hospital, since other than reducing the per-patient health payment, the hospital might also decrease the number of patients to ensure the total costs do not exceed the global budget.

In this section, our framework refers the model of Hodgkin and McGuire (1994) which illustrate PPS effect on cost control, and we specify the payment function to adopt the framework to GB reform. The theoretical results extend current literature of GB studies in China by predicting the changes in bills, OOP and service volume and illustrating hospital reaction under GB in China more comprehensively. Most empirical findings (Dou et al., 2019, Huang et al., 2016, Guan et al., 2020, Xie et al., 2015, Yuan et al., 2019, Liang et al., 2012, Yang et al., 2022, Wang, 2015) in China support our predictions on bills and OOP.

2.3 The comparison between FFS and PPS

Recall Eq. (12), under FFS, the profit of treating the representative patient equals the bill of treating the patient minus the actual cost c for treating the patient.

$$\pi = b - c \quad (12)$$

Under PPS, the income for treating a patient becomes the sum of the patient OOP (f) and the PPS payment made by the payer (m), which are both in the fixed amount:

$$\pi = f + m - c \quad (22)$$

The transition from FFS to PPS decreases *e* (i.e., OOP: bill ratio under FFS) and *r* (i.e., the reimbursement: bill ratio under FFS) from some positive value to closer to 0, and the profit under PPS is not related solely to the bill. PPS can reduce the hospital´s revenue in two ways. First, as the reimbursement is a fixed tariff (that is, so called average cost reimbursement), the marginal income has been eliminated under PPS. Second, the PPS tariff (*f*+*m*) can be reduced, until it is only slightly higher than or equal to the hospital´s actual cost.

Recall Eq. (15) and Eq. (16), under FFS, the objective function of the hospital for treating a representative patient would be:

$$u_{FFS} = (1 - \lambda)(b - c) + \lambda h \quad (15)$$

Taking the first order condition on u_{FFS} with respect to LoS yields:

$$u'_{FFS} = (1 - \lambda)(b - c)' + \lambda h' \quad (16)$$

Under PPS, b is substituted for the PPS tariff *(f+m)* so the objective function of the hospital for treating a representative patient becomes:

$$u_{PPS} = (1 - \lambda)(f + m - c) + \lambda h \quad (23)$$

Then we take first order condition on the utility under PPS, with respect to LoS:

$$u'_{PPS} = -(1 - \lambda)c' + \lambda h' \quad (24)$$

By comparing Eq. (16) with Eq. (24), we find that the parameter $\lambda h'$ is unchanged in both equations, which indicates that irrespective of payment model (i.e., FFS or PPS), the hospital is motivated to increase LoS to the optimal value LoS* and provide sufficient care to the patient to guarantee recovery due to its own altruism concern for the patient. However, the parameter $(1 - \lambda)(b - c)'$ under FFS changes to $-(1 - \lambda)c'$ under PPS and turns negative. This change represents the policy effect of payment reform of switching from FFS to PPS. Since under FFS hospitals are reimbursed the marginal cost plus a profit margin, the hospital is motivated to increase LoS to increase revenue and maximize profits. When the payment is shifted to PPS, reimbursement at marginal cost and the profit margin are both eliminated. Thus, the hospital is incentivised to decrease LoS to reduce the cost of treating the patient so that is less than the fixed PPS payment.

u'_{PPS} has two properties. First, u'_{PPS} is smaller than u'_{FFS}, which indicates that the hospital has a lesser motivation to increase LoS under PPS compared to when paid under FFS. This result is consistent with the aim of PPS adoption to control for the increasing trend in health care expenditure. Second, if the effect of cost control ($-(1 - \lambda)c'$) is larger than the effect of altruism ($\lambda h'$), U'_{PPS} could be negative. In this case, the hospital will reduce LoS and might have insufficient incentive to treat patients, which would have adverse consequences on their health outcomes.

Based on this theoretical framework, we have several hypotheses relating to the behaviour of hospitals when payment is shifted from FFS to PPS:

Hypothesis 1: The bills for patients will decrease after the policy because the hospital has the motivation to ensure bills do not exceed the PPS tariffs.

Hypothesis 2: The OOP for patients will decrease after the policy. The OOP becomes a fixed amount as a portion of PPS tariff under PPS while it is related with the quantity of health service under FFS, thus the hospital has the motivation to control costs and, hence, the OOP for the patients.

Hypothesis 3: The payment policy will result in a decrease in LoS because the hospital tends to reduce the bill for patients by reducing the treatment intensity.

Hypothesis 4: The payment policy will lead to an increase in unplanned 30-day readmissions, suggesting a decrease in health outcomes because the decrease in treatment intensity might cause insufficient care for the patients.

In this section, we extend the model of Hodgkin and McGuire (1994) to three agents-patient, payer and hospital, where the payer as the government plays the role of reimbursing to hospitals as well as monitoring the performance of hospital, to adopt the framework in China. Then we expand the predictions to the changes in bills, OOP, LoS and health outcomes for patients. The theoretical results extend current literature of PPS studies in China by predicting the changes in bills, OOP, LoS and health outcomes and illustrating hospital reaction after PPS in China more comprehensively. Also, the theoretical framework explains the mechanism through which PPS affects bills, OOP and health outcomes.

As most empirical studies in China found that PPS incentivized cost control (Jian et al., 2015a, Peng et al., 2021, Wu et al., 2022, Peng et al., 2018a, Yuan et al., 2019, Li et al., 2018, Li et al., 2019, Meng et al., 2022, Hu et al., 2023, Huang et al., 2022, Jian et al., 2015b), and cost control was associated with decrease in LoS (Jian et al., 2015b, Jian et al., 2015a, Li et al., 2018, Wu et al., 2022, Peng et al., 2018a, Meng et al., 2022, Huang et al., 2022). Thus, our predictions on bills, OOP and LoS are plausible.

Although several studies found PPS had no effect on health outcomes (Jian et al., 2019, Yuan et al., 2019, Wu et al., 2022), these studies all focused on AMI patients and may not apply more widely; and the other studies (Jian et al., 2015b, Hu et al., 2014, Zhang and Hu, 2015) among patients from various diseases found worsened health outcomes after PPS. Thus, our predictions that health outcomes will be worsened after PPS should be plausible.

Conclusion

This chapter developed a three-agent theoretical framework, being based on the model of Hodgkin and McGuire (1994), and predicts that LoS and the number of patients for hospitals will decrease after the GB adoption and the bills, OOPs and LoS for patients will decrease after PPS adoption, while the 30-day readmissions for patients will increase after PPS adoption. The next chapter will provide two literature reviews of GB studies in China and PPS studies in China separately. Corresponding to the predictions on outcome variables presented in the theoretical framework, the literature review will focus on LoS and the number of patients for hospitals for GB studies. It will also focus on outcomes as bills, OOPs, LoS, and 30-day readmissions for PPS studies.

3. Literature review

Summary

Chapter 3 identifies and summarizes the studies on GB reforms in China and PPS reforms in China in two reviews. Flow diagrams are presented to illustrate the literature selection process, while Summary Tables and forest plots summarize the findings of previous studies.

3.1 Methods

Keywords

I, as the sole reviewer, searched databases including PubMed, Springer, Science Direct, and the LSE Library, conducting two independent literature reviews for GB studies and PPS studies in China, separately. I used the keywords "global budget payment China" and "*global budget*" (to ensure 'global budget' appears as a fixed phase) for the GB reforms literature, and "DRG payment China" and "DRGs payment China" for the PPS-DRG reforms literature. The literature search was augmented by referencing the bibliographies of prior reviews to identify additional relevant reviews and potentially eligible studies. No restrictions were applied regarding publication status or date.

Selection strategy

Literature search records were managed using NoteExpress. I screened titles and abstracts of all articles included in the search results after the exclusion of duplicates. Studies related to PPS reforms in China underwent a full-text review.

I selected studies for review based on the following criteria: (1) The intervention of the study should be based on GB reforms in China for GB study reviews, and on PPS-DRG reforms in China for PPS study reviews. For studies with a background of mixed interventions, the primary policy evaluation should focus on one of the aforementioned interventions, which can be applied at the institutional, regional, or individual level. (2) Studies must provide a comparison between the policy-targeted patients and other untargeted patients, or a before-and-after comparison. (3) I focused on empirical studies using either hospital-level data or individual-level data by applying econometric methods. (4) The outcome variables of the study may include payments (the payment or OOP for the patient) or treatment intensity (LoS).

Study Selection

After excluding duplicates, I screened titles and abstracts of all identified articles. Based on the selection criteria, I initially excluded studies unrelated to the payment reforms, studies without an abstract or that were unavailable, and studies not in English or Chinese. I then excluded studies whose payment reforms were not in China, and those in China but not focused on GB or PPS-DRG reforms. After abstract screening, studies eligible for inclusion were subjected to full-text evaluation. During this phase, I excluded studies based on their methods, research objectives, and outcome variables, ultimately selecting those that met the inclusion criteria for this review.

Data Extraction

I independently retrieved the following characteristics of each study: first author name, year of publication, year of policy intervention, sample size, estimation methods, and the effects of the policy on outcome variables. This step facilitated the assessment of the combinability of the studies included in this review.

Statistical Analyses

For continuous outcome variables, which include changes in patient payments, length of stay (LoS), and the number of patients, the Standard Mean Difference (SMD) with a 95% Confidence Interval (CI) was used to indicate effect size. The threshold for statistical significance was set at $P \leqslant 0.05$. It is noteworthy that the studies investigated either bills or OOP; therefore, the effects on both were consolidated into a single forest plot, and 'the effect on payment' was used to denote the impact on either bills or OOPs.

For the dichotomous outcome variable (i.e., the change in readmissions estimated by Logit regression), Odds ratios (ORs) with 95% CIs were employed to denote effect size. The threshold for statistical significance was maintained at $P \leq 0.05$.

Among the selected studies, effect sizes were reported as relative changes, expressed as percentages, and absolute changes, expressed as difference values. I converted absolute changes into relative changes by dividing the difference by the baseline value. Following the methodology of previous literature reviews on payment reforms (Chen et al., 2023, Meng et al., 2020), I listed all selected GB and PPS studies separately in two forest plots, created sub-groups according to the various methods used in these studies, and estimated the general effect size for all studies within each forest plot.

Heterogeneity was assessed using the Higgins I^2 test; an I^2 value greater than 50% indicated significant heterogeneity among the studies. Consequently, a random-effects model was employed to address this heterogeneity.

RevMan 5.4.1 was utilized for all statistical analyses.

3.2 Results

3.2.1 The results of GB studies in China

Figure 2 illustrates the process of selecting literature on GB reforms in China. A total of 399 studies were identified, with 368 studies retained for title and abstract screening after removing duplicates.

After excluding studies unrelated to the payment reforms, those without an abstract, and those not in English or Chinese, 150 studies remained. Further exclusions of studies whose payment reforms were not in China, or were in China but not focused on GB or PPS-DRG reforms, resulted in 30 studies being preserved for full-text screening.

Following the full-text screening, 15 studies were excluded due to the methods, research objectives, outcome variables, or the unavailability of the full text. This resulted in 15 studies being finally selected for inclusion in the review.

There are 15 studies selected for the review, comprising three Treatment-Control (TC) studies, five Before-After (BA) studies, two Interrupted Time Series (ITS) studies, and five Controlled Before-After (CBA) studies.

Among the 13 studies investigating the GB effect on bills or OOP, 10 identified a reduction in patient payments among targeted patients, 2 noted a decrease in bills but an increase in OOP among targeted patients, and 1 found an increase in bills but a decrease in the OOP ratio. 4 studies explored the GB effect on LoS, all of which reported a reduction in LoS for GB-targeted patients. Thus, the evidence suggests that the GB adoption in China has strong effects on cost containment and LoS reduction, however, hospitals might transfer the financial burden to patient OOPs. Three studies examined the GB effect on service volumes in hospitals: two reported an increase in service volume (1 in the absolute number of patients and 1 in the proportion of inpatients), and 1 observed a decrease in service volume through the reduction in the number of discharged patients. The summaries of these studies are listed in Table 2 below.

To quantify the overall effect sizes of GB from the selected studies, I employ forest plots, following the methodology of previous literature reviews on payment reforms in China. These forest plots synthesize the GB effects from the selected studies on patient payments, LoS, and readmissions separately. Due to some studies not reporting standard errors or standard deviations, the CI cannot be calculated for them, and consequently, they are excluded from the forest plots. Additionally, when a study employs different analytical samples, each analytical sample's result is treated as an individual study result in the forest plots.

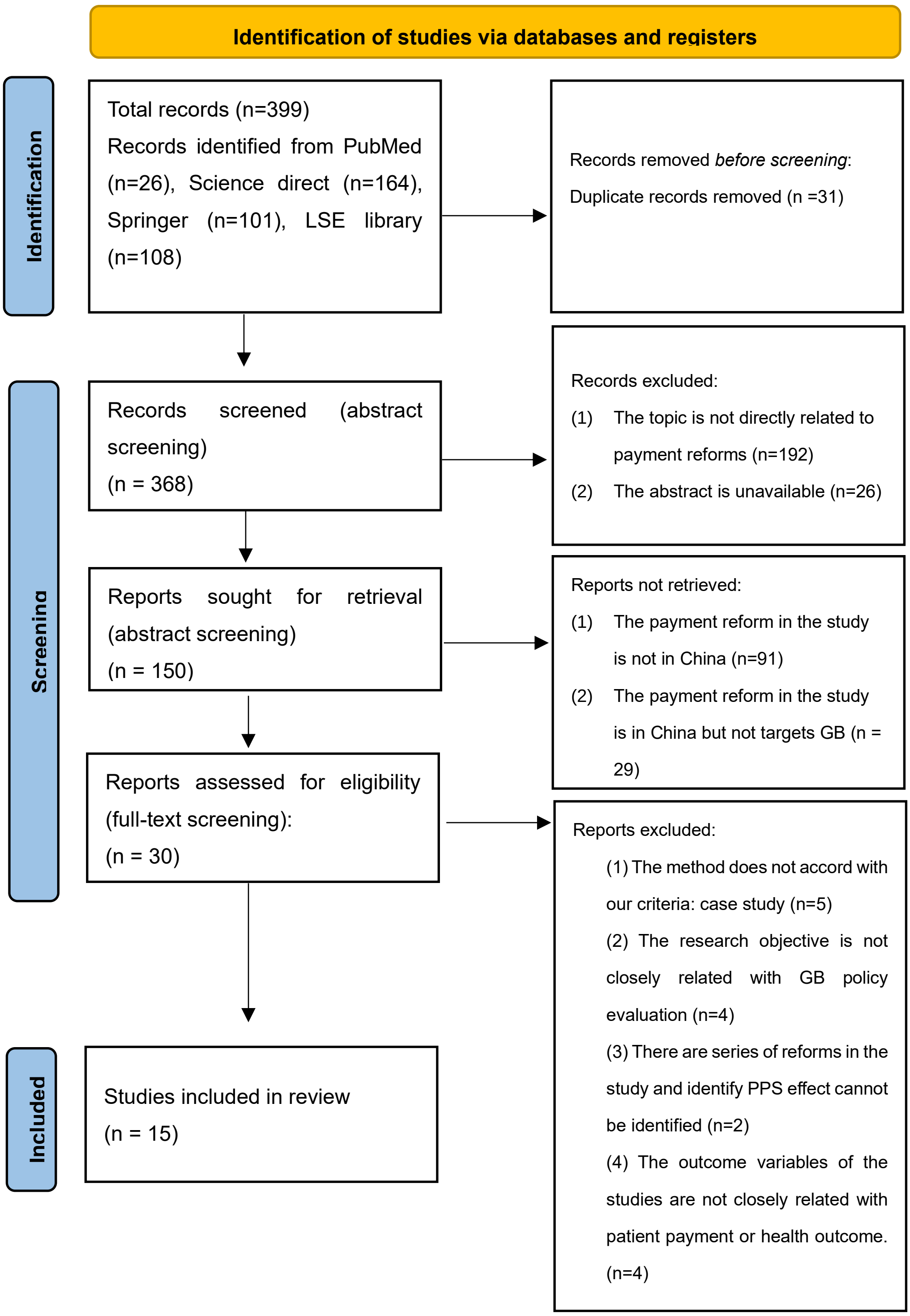


Figure 2. PRISMA flow diagram of studies on GB reforms

Table 2. The summary of studies on GB reforms in China

Reform	year of intervention	Studies	GB targeted samples in the study	Method	Policy effects
YC county in Hubei province	2013	He et al. (2017)	33,175 NRCMS inpatients before and 36,883 NRCMS inpatients after the reform	CBA(DiD)	A decrease in growth of bills (CNY263.35) and reimbursed expenses (CNY447.46), but with a increase (CNY 188.6) in OOPs. The increase (11.4%) in readmissions.
Shanghai	2011	Dou et al. (2019)	The inpatients of cardiovascular diseases from tier 1, tier 2 and tier 3 hospitals in Shanghai.	ITS	An instant drop in bills (CNY 55.71 million), the decrease in the monthly increasing trend (CNY 4.23 million) and the instant decrease in discharge number (10.4%) .
Tianjin	2011	Huang et al. (2016)	102,492 UREBMI outpatients of hypertensive from Tianjin's primary hospitals	BA	Under GB, the bills (CNY 640.28 vs. CNY 700.64), reimbursed payments (CNY 491.87 vs. CNY 532.37) and OOPs (CNY 148.42 vs. CNY 168.27) are all lower compared to FFS.
Weiyuan county in Gansu province	2015	Li et al. (2018)	127,491 NRCMS inpatient observations	BA	The increase in bills and decrease in OOP ratio and LoS.
Yunnan province	2015	Guan et al. (2020)	The patients from a secondary hospital	BA	The reduction in bills (1.7%), LoS (0.5 day), and the reduction the median value of monthly admission and readmission.
Guizhou province	2016	Zhou et al. (2021)	Inpatients of pneumonia, chronic asthma, acute myocardial infarction and stroke, from 16 county-level hospitals	CBA(DiD+PSM)	Increase in rate of sputum culture in patients with pneumonia, rate of aspirin at discharge, rate of discharge with β-blocker and rate of smoking cessation advice in AMI patients; decrease in rate of oxygenation index assessment in patient with chronic asthma; no changes in other indicators of process quality.
Shunyi district, Beijing	2011	Xie et al. (2015)	The 91 hospitals in Shunyi district from 2009 to 2014	BA	The increasing rate of total reimbursed payment of the hospital decreased after the policy (from 134.18% in 2011 to 52.09% in 2012, 28.46 in 2013 and 27.94 in2014.)
Zhongshan, Guangdong province	2010	Yuan et al. (2019)	2895 patients diagnosed with acute myocardial infarction (AMI) from two tier 3hospitals in Zhongshan from 2008 to 2014, with the local insurers as treatment group and migrant insurers as control group.	CBA(DiD)	Decrease in bills (15%) for targeted patients and no significant changes in targeted patients' hospital mortalities or 30-day readmissions.
Shunde district, Foshan	NA	Liang et al. (2012)	The UEBMI funding annual statistics in two districts in Foshan city from 2007 to 2010	TC	The bills in targeted district decreased during 2009 and 2010 (from CNY6393 to CNY5712) while they increased in untargeted district (from CNY7402 to CNY7740).
City A	2009 and 2010	Zhang et al. (2013)	The UEBMI patients collected from the survey on the health service utilization of the health insurees in China in city A (treatment) and city B (control) from 2009 to 2011	TC	The growth rate of bills for city A is lower than city B (8.07% v.s. 9.80%).
Chengdu	2013	Li and Zhang (2022)	UEBMI patients from 2012 to 2014 provided by the Chengdu Healthcare Security Administration, 2876 observations for patients with acute suppurative appendicitis under GB payment (treatment), and 1556 observations for patients with acute appendicitis under DRG payment (control)	CBA(DiD)	Death reduced by 7.1%, transfer reduced by 5.0%, and rehabilitation increased by 12.1% for GB targeted patients after the policy compared with DRG targeted patients.
Nationwide study	NA	Yang et al. (2022)	400672 stroke patients of urban health insurance claims randomly collected from the China Health Insurance Research Association (CHIRA) database from 2013 to 2016	TC	The bills for GB were lower than FFS (CNY 13674.5 v.s. CNY13860.7), The OOPs for GB were higher than FFS (CNY3869.6 v.s. CNY3665.9), the LoS for GB was lower than FFS (14.1 days v.s. 14.7 days)
Beijing	2013	Wu and Guo (2016)	The patients from a dental hospital in Beijing	BA	After the policy, the number of outpatients increased (from 166784 in 2011 and 206239 in 2012 before the policy to 231525 in 2013 and 243823 in 2014 after the policy), the number of inpatients increased (from 644 in 2011 and 800 in 2012 before the policy to 921 in 2013 and 1011 in 2014 after the policy). The bills for outpatients increased (from CNY237 in 2011 and CNY252 in 2012 to CNY276 in 2013 and CNY 293 in 2014), The bills for inpatients increased (from CNY11366 in 2011 and CNY10284 in 2012 to CNY12017 in 2013 and CNY12863 in 2014),
Gaolan county, Gansu province	2011 and 2013	Wang (2015)	NRCMS claims from 2010 to 2013, with 118607 claims in total, 16741 fromGaolan County and 101866 from the control county	CBA (DiD)	The bills for targeted patients were unchanged, OOPs and LoS for targeted patients decreased by CNY157.79 and 0.082 days after the policy for Shidong THC. The bills, OOPs and LoS for targeted patients decreased by 5.4%, CNY 92.486 and 0.141days for other THCs.
City A	2017	Xiang et al. (2022)	7 targeted hospitals in city A from 2016 to 2018	ITS	The trend of the portion of inpatients for hospitals which measured the service volume (assuming that the hospitals would decompress hospitalization to increase the volume) increased by 0.8 for tier 3 hospitals after the policy.

In Figure 3, the studies investigating the effect of GB on patient payments comprise of BA, ITS, TC, and CBA studies. As the I^2 statistic exceeded 50%, indicating heterogeneity among the studies, a random-effects model was employed to address this heterogeneity. A significant decrease in patient payments by 20% (with a CI from -31% to -8%) was observed. Yang et al. (2022) investigated both bills and OOPs for a general analytical sample. Wang (2015) and Li et al. (2019) investigated both bills and OOPs for different analytical samples separately. Dou et al. (2019) examined the policy effects of two waves of GB reforms separately. Therefore, each outcome variable or analytical sample result was individually included, akin to the result of a separate study.

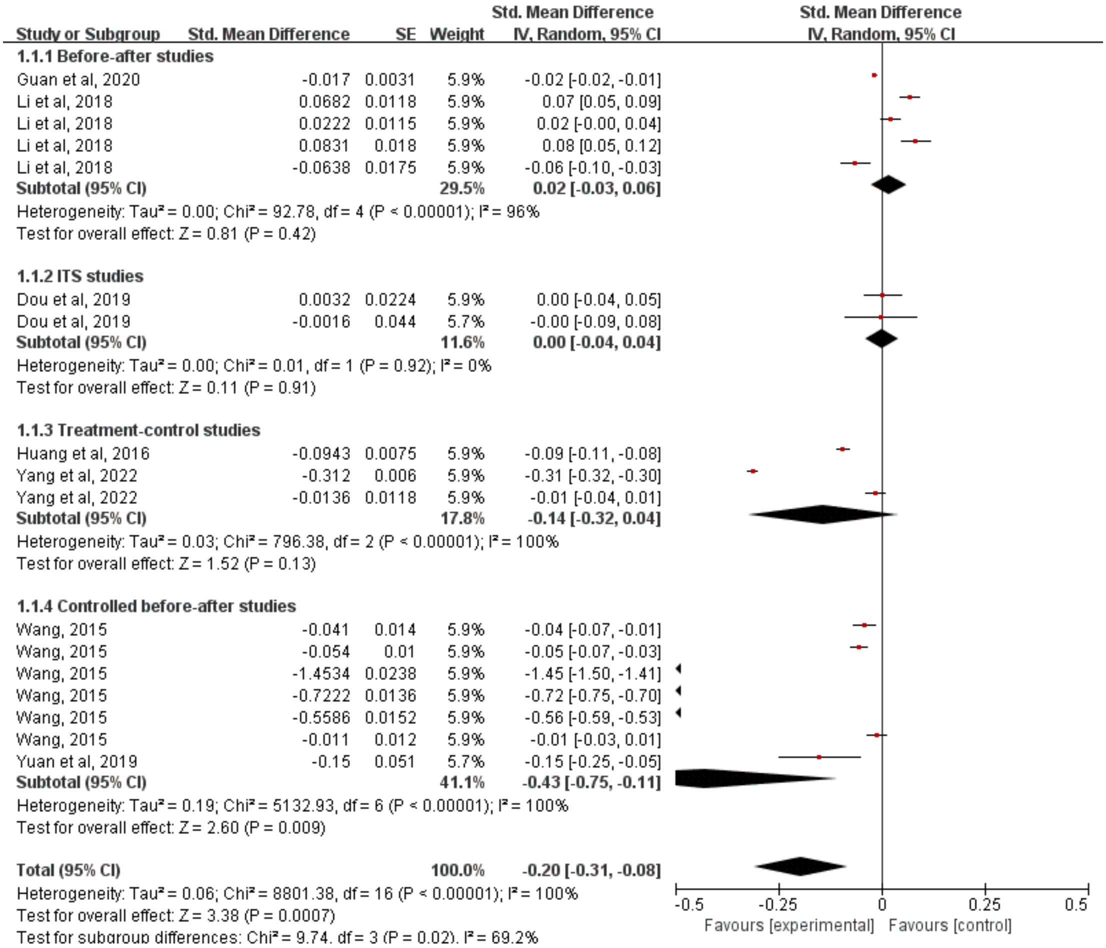


Figure 3. Forest plot for GB effect on payments

In Figure 4, studies investigating the effect of GB on LoS include BA and CBA studies. Since the I^2 statistic exceeded 50%, indicating heterogeneity among the studies, a random-effects model was employed to address this heterogeneity, resulting in no significant change observed. Wang (2015) and Li et al. (2019) investigated LoS for different analytical samples separately. Consequently, the result of each analytical sample was individually included, akin to the result of a separate study.

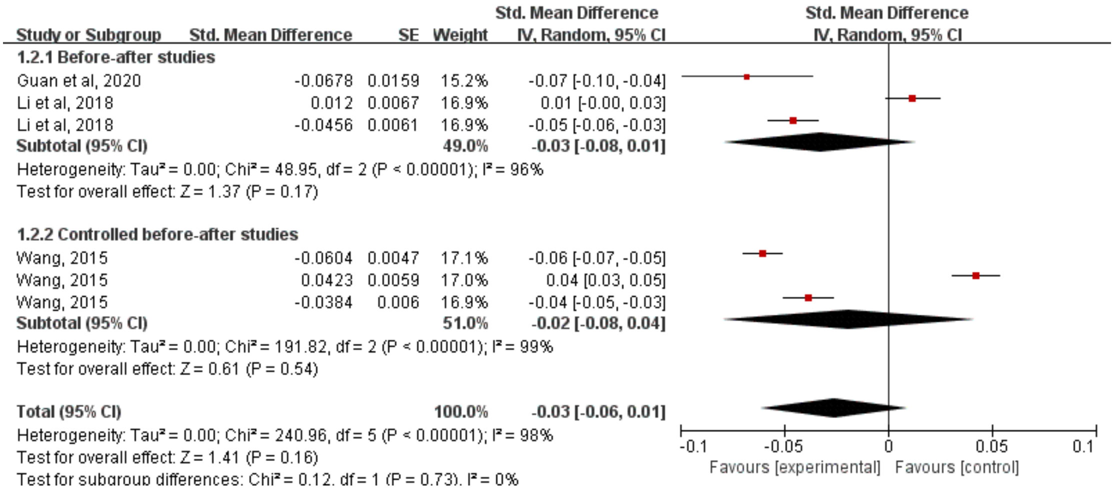


Figure 4. Forest plot for GB effect on LoS

3.2.2 Results of PPS studies in China

Figure 5 illustrates the literature selection process for PPS reforms in China. Initially, a total of 1045 studies were identified, of which 998 studies were retained for screening titles and abstracts after eliminating duplications.

Following the exclusion of studies irrelevant to payment reforms, those lacking abstracts or availability, and those not in English or Chinese, 432 studies remained. Subsequently, studies not focused on payment reforms in China or not specifically on GB or PPS-DRG reforms were excluded, leaving 104 studies for full-text screening.

After full-text assessment, 88 studies were excluded based on methodology, research objectives, outcome variables, and availability of full text, resulting in the final selection of 16 studies for review.

In total, 16 studies were selected for review, comprising 2 Treatment-Control (TC) studies, 6 Before-After (BA) studies, 1 Interrupted Time Series (ITS) study, and 7 Controlled Before-After (CBA) studies.

Among the 13 studies investigating the PPS effect on patient payments, 11 identified a reduction in patient payments (either bills or OOP) among targeted patients, while 2 found no significant change. However, among these 2 studies, Meng et al. observed a decrease in Out-of-Pocket payments despite no change in bills. Nine studies explored the PPS effect on LoS, with all identifying a reduction in LoS for PPS-targeted patients. Consequently, PPS adoption in China demonstrated robust effects on cost containment and LoS reduction. 6 studies investigated the PPS effect on readmissions, wherein 3 observed no significant change for PPS-targeted patients, 1 noted higher readmissions, 1 found lower readmissions, and 1 observed both higher and lower readmissions for PPS-targeted patients, respectively, in 2012 and 2013. Thus, the PPS effect on readmissions in China remains undetermined. The summaries of these studies are listed in Table 3 below.

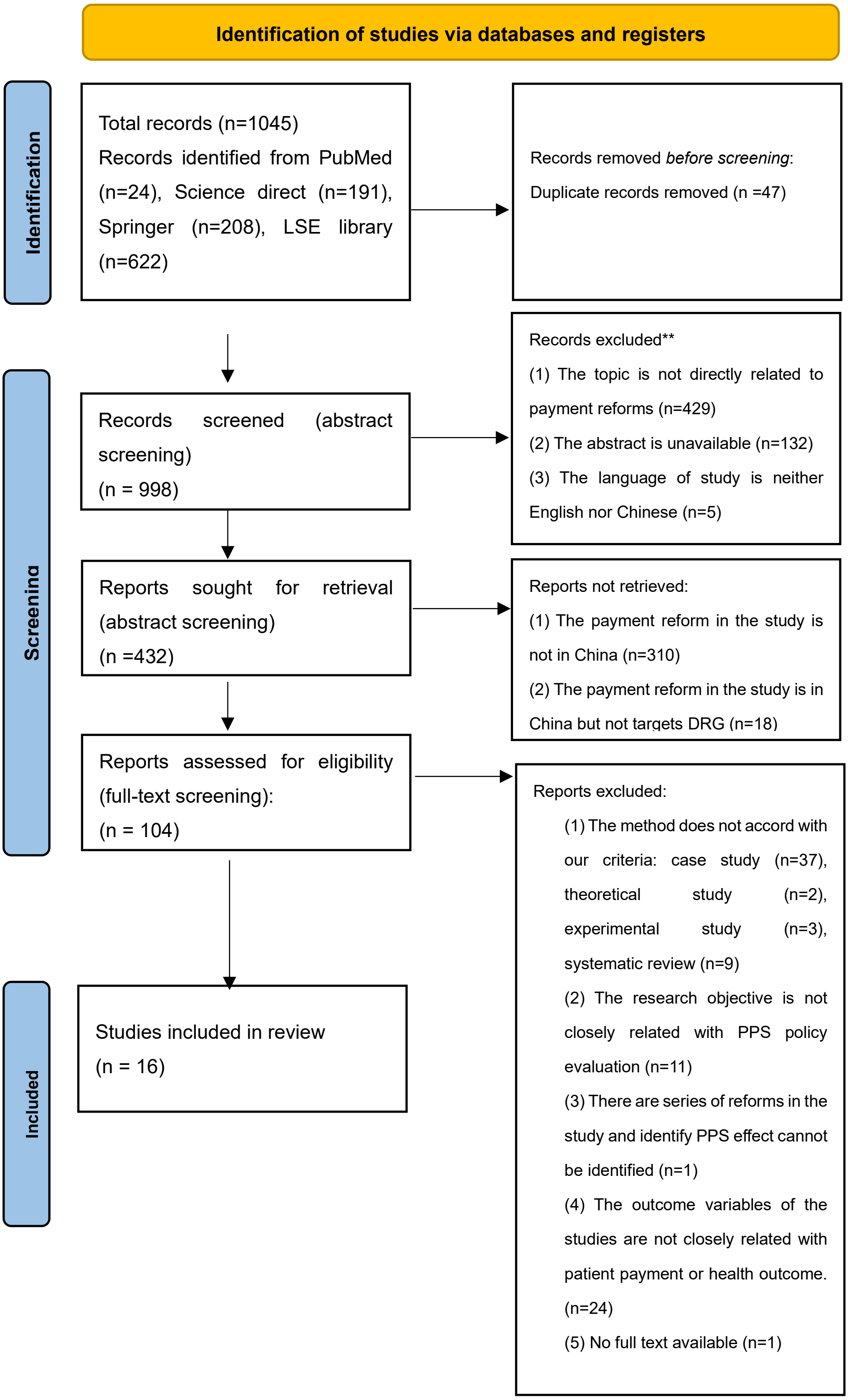


Figure 5. PRISMA flow diagram of studies on PPS reform

Table 3. The summary of studies on PPS reforms in China

Reform	years	Evaluations	Study sample	Method	Policy effects
Shanghai	2004	Zhang (2010)	A 35% random sample from 2004 and 2005 containing about 14 000 inpatient patients collected from a tier 3 hospital.	CBA (DiD and DDD)	No change in bills; decrease (13.1%) in LoS for targeted patients after the policy.
Zhongshan, Guangdong province	2010	Yuan et al. (2019)	2895 patients diagnosed with acute myocardial infarction (AMI) from two tier 3hospitals in Zhongshan from 2008 to 2014, with the local insurers as treatment group and migrant insurers as control group.	CBA (DiD)	Decrease in bills (15%) for targeted patients and no significant changes in targeted patients' hospital mortality or 30-day readmissions.
Beijing	2011	Hu et al. (2014)	UEBMI patients from 6 targeted hospitals and other 8 similar untargeted hospitals from 2008 to 2012.	CBA	After the policy, in 2012, LoS and 14-day readmissions for targeted hospitals are 0.47 days lower and 0.68% higher than other hospitals.
		Jian et al. (2015b)	584825 discharged insured patients from tier all the 3 hospitals in Beijing during 2008 and 2012, collected from the health insurance department in Beijing	CBA	The reduction in OOPs for patient with targeted diseases in the 6 targeted hospitals were higher than the reduction in OOPs for all the tier 3 hospitals (CNY135.8 v.s. CNY71.3). Even in targeted hospitals, not all the patients with targeted diseases still chose FFS payment rather than DRG payment. Among the patients with targeted diseases and from targeted hospitals, the 14-day readmissions for FFS group is higher than DRG group (6.84% v.s. 2.85%), and the LoS for FFS group is longer than DRG group (10.0 days v.s. 8.9 days).
		Jian et al. (2015a)	294989 discharge patients collected from the Beijing Health Insurance Bureau during 2010 and 2012 from 14 tier 3 general hospitals (6 hospitals piloted DRG payments and the other 8 hospitals used only FFS-based payments).	CBA (DiD)	After the policy, compared to control hospitals, there were reductions in bills both for eligible DRG cases at pilot hospitals (4.8%) and for DRG cases that actually received DRG payment (6.2%), and reductions in LoS for eligible DRG cases (3.6%), cases that actually received DRG payment (3.0%).
		Zhang and Hu (2015)	The 14 tier 3 hospitals (6 intervention DRG hospitals and 8 FFS hospitals) in Beijing during 2012 and 2013.	TC	In 2012 and 2013, the LoS for DRG hospitals were lower than FFS hospitals (by 0.47 days and 0.36 days). The 14-day readmissions for DRG hospitals were higher than FFS hospitals (by 0.68%) in 2012 and lower (by 0.94%) in 2013.
		Peng et al. (2021)	513 elderly patients from the Peking Union Medical College Hospital diagnosed with unilateral displaced femoral neck fractures from 2006 to 2017.	BA	The overall treatment costs increased by CNY21,028 for the targeted patie nts.
		Jian et al. (2019)	7119 AMI patients under UEBMI from 14 tier 3 hospitals from 2010 to 2012.	CBA (DiD)	No changes in in-hospital death and 30-day readmissions for AMI patients; no improvements on the indicators of process quality (Aspirin at arrival, Aspirin prescribed at discharge, β-blocker prescriptions on arrival and β- blocker prescriptions at discharge).
Weiyuan, Gansu province	2015	Li et al. (2018)	127491 inpatient records during 2014 and 2016 NRCMS database in Weiyuan County	BA	The decrease in bills (by CNY 213.16), the decrease in LoS (by 0.28 days) and the decrease in OOP ratio (by 2.29%).
	2016	Li et al. (2019)	9 targeted hospitals in Yuxi from 2016 to 2017	BA	The growth rate of bills for targeted hospitals decreased by 6% after the policy.
L city, Guizhou Province	2017	Wei and Feng (2019)	141987 cases from 4 hospitals (with tier 3 hospitals A and B, tier 2 hospitals C and D) in L city	ITS	The growth rate of bills declined by CNY43.98 per month in hospital A, while that increased by CNY5.91 per month in hospital B, the increase of bills in hospital C by CNY68.33 per month.
Tianjin	2017	Wu et al. (2022)	All admissions of angina and AMI during 2015 to 2018 in Tianjin, 43698 in the intervention group (20 088 pre- and 23610 post-program), and 13134 in control group (6768 pre- and 6366 post-program)	CBA (DiD)	The decrease LoS, bills, and OOPs by 20.8%, 14.2%, and 95.5%, respectively, while no change in readmissions.
Anhui province	2013	Peng et al. (2018a)	1432 uterine leiomyoma inpatients from 11 hospitals in three regions of Anhui province from 2010 to 2015 that were randomly collected from New Rural Cooperative Medical Office of Anhui Province.	BA	The per-case hospitalization bill dropped from 919.08 ± 274.92 USD to 834.91 ± 225.29 USD and length of hospital stay reduced from 9.96 ± 2.39 days to 8.83 ± 1.95 days($P < 0.01$).
Sanming, Fujian province	2018	Meng et al. (2022)	The patients with hip fracture aged above 60 years from all the public hospitals enrolled in the Sanming Basic Health Insurance Scheme from 2016 to 2018.	BA	No significant change in bills. OOPs decreased from CNY9587 to CNY7618, LOS decreased from18.89 days to 17.48 days
Nationwide study	NA	Hu et al. (2023)	28200 insured urban resident lung cancer inpatients from the China Medical Insurance Research Association (CHIRA) from 23 provinces and 3 municipalities during 2010 and 2016	TC	The payment for lung caner patient under PPS was 8.4% lower than the payment under FFS.
Guangxi province	2021	Huang et al. (2022)	3571 Tuberculosis patients from a tier 3 hospital (1741 for control and 1830 for treatment) in Guangxi province from Jan 2021 to Dec 2021.	BA	Compared with the patients before the policy, the patients after policy had lower per-admission payment (CNY12619.5 v.s. CNY15187.8) and shorter LoS (12.3 days v.s. 14.3 days).

To quantify the overall PPS effect sizes of the selected studies, forest plots are utilized to synthesize the PPS effects on patient payments, LoS, and readmissions separately. Due to some studies not reporting standard errors or standard deviations, CI cannot be calculated for them, rendering their exclusion from the forest plots.

In Figure 6, studies examining the PPS effect on patient payments comprise BA and CBA studies. As the I^2 statistic exceeded 50%, indicating heterogeneity among the studies, a random-effects model was employed to address this heterogeneity. A significant decrease in patient payments, by 13 % (with a CI from -26% to 0), was observed. Meng et al. (2022) investigated policy effects on both bills and OOPs, while Zhang (2010) studied two waves of reforms separately. Thus, the result of each outcome variable or each analytical sample was individually included, akin to the result of a separate study.

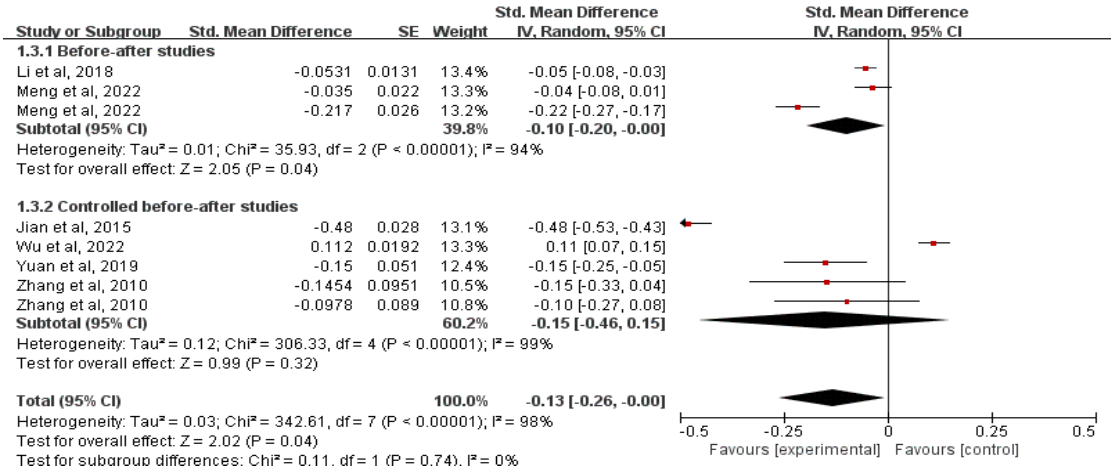


Figure 6. Forest plot for PPS effect on payments

In Figure 7, studies investigating the PPS effect on LoS included BA and CBA studies. Since the I^2 statistic surpassed 50%, indicating heterogeneity among the studies, a random-effects model was employed to address this heterogeneity. A significant reduction in LoS, by 12% (with a CI from -18% to -5%), was observed. Zhang (2010) investigated two waves of reforms separately. Thus, the result of each analytical sample was included individually, akin to the result of a separate study.

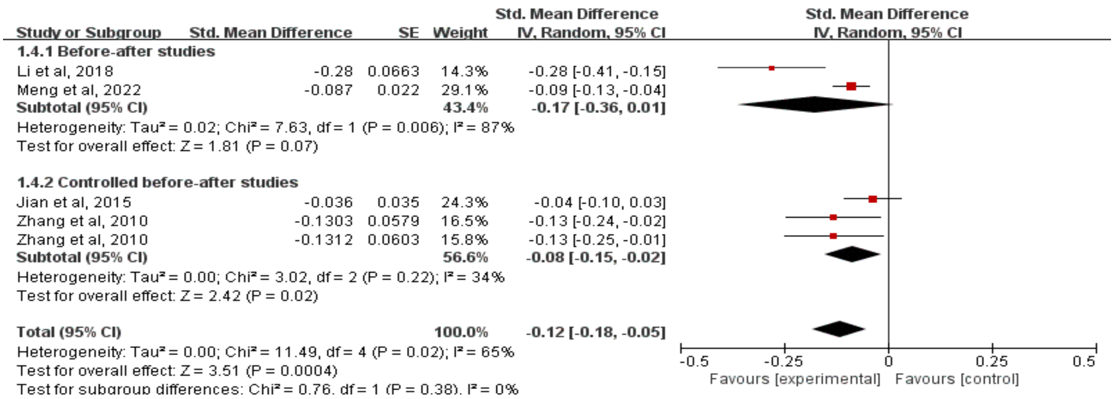


Figure 7. Forest plot for PPS effect on LoS

In Figure 8, studies investigating the PPS effect on readmissions solely comprise CBA studies. As the I^2 statistic was less than 50%, indicating no heterogeneity among the studies, a fixed-effects model was utilized for estimation. Consequently, no significant policy effect was observed. Jian et al. (2019) examined policy effects for different samples separately. Thus, the result of each analytical sample was individually included, akin to the result of a separate study.

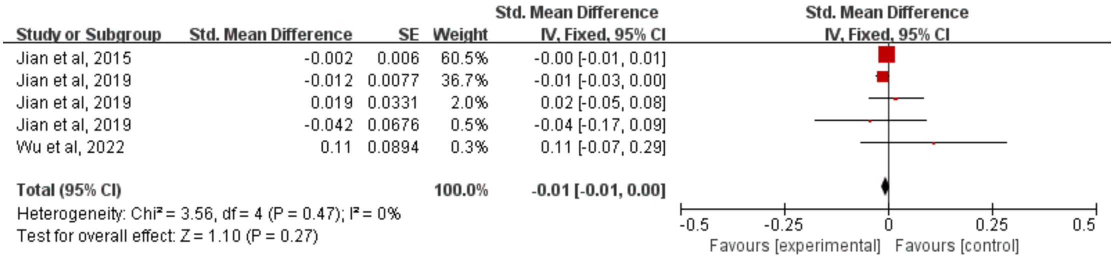


Figure 8. Forest plot for PPS effect on readmissions

3.3 Discussion

When reviewing economic studies, a challenge arises from the application of various econometric methods to observe effects, leading to different interpretations of the results. For instance, the coefficient of multivariate regression indicates the marginal effect, the coefficients of ITS indicate the immediate before-after policy effect and its effect on trends, and the coefficient of DiD indicates the effect on the treatment group compared with the control group after the policy. Therefore, when incorporating findings from all these economic studies into one plot, it is essential to create sub-groups based on econometric methods and subsequently calculate the overall effect size for all the sub-groups. However, a limitation of this approach exists: since the effects of different sub-groups are not entirely comparable, it is not entirely valid to calculate the general effect of all the sub-groups, despite previous studies (Chen et al., 2023, Meng et al., 2020) doing so.

Conclusion

This chapter provided two separate reviews of the existing literature examining the implications of GB and PPS reforms in China separately. The following chapter will introduce the context and health system in Chengdu, where the payment reforms evaluated by this dissertation were implemented, and introduce the data sources for the empirical studies on the payment reforms in Chengdu.

4. Context and data sources

Summary

The empirical element of this dissertation focuses on evaluating the payment reform in Chengdu, which is the capital city of Sichuan province and located in the southwestern part of China. Chapter 4 introduces the context and health system in Chengdu and introduces the data sources for the empirical studies on the payment reforms in Chengdu.

4.1 Chengdu context

In 2016, Chengdu had a population of 15.9 million (Chengdu Bureau Of Statistics, 2017), which makes it the 5th most populous agglomeration in China. Due to the technology and innovation boom, Chengdu attracts an increasing number of internal migrants from other areas of China and the population size exceeded 20 million in 2020 (Chengdu Bureau Of Statistics, 2021). According to the 2019 Annual Report of Health Service Development in Chengdu, the life expectancy of the residents is 81.01 years, 78.42 for males and 83.84 for females. The average annual disposable income per capita has reached CNY 39,503 (USD 6107) in 2019, CNY 45,878 (USD 7093) for urban residents and CNY 33,202 (USD 5133) for rural residents. Thus, the empirical evidence of PPS reform in Chengdu as one of largest cities in China will provide strong policy implications for implementing a payment reform and may guide future payment reforms across other parts of China.

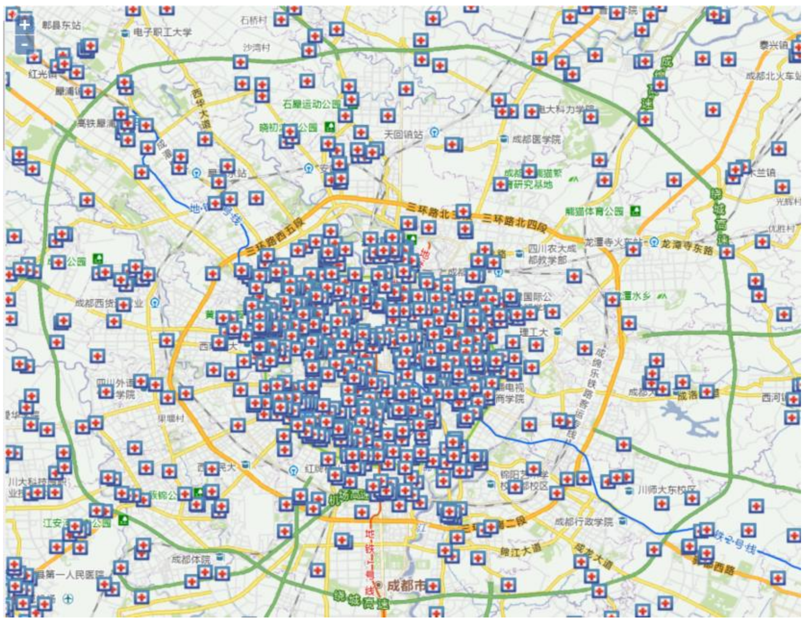


Figure 9. The distribution of health institutions in Chengdu

According to the 2019 Annual Report of Health Service Development in Chengdu, there are 12,121 health institutions in total, including 629 hospitals, 11355 primary health care institutions and 77 public health institutions. Figure 9 shows the distribution of hospitals in Chengdu on Baidu map. According to the type of ownership, 172 hospitals are public hospitals while 482 hospitals are private hospitals. According to the classification, there are 74 tier 3[4] hospitals, 130 tier 2[5] hospitals and 450 other hospitals (including tier 1 hospitals and hospitals without tier). According to bed capacity scale, there are 377 hospitals with less than 100 beds, 107 hospitals with 100 ~ 199 beds, 78 hospitals with 200 ~ 499 beds, 35 hospitals with 500 ~ 799 beds, and 32 hospitals with 800 beds and above.

In 2019, the number of health staff in Chengdu reached 237,600. Among the health staff, 184,900 are health technicians, 8,400 of other technicians, 13,200 administrative staff, and 27,400 logistics staff. Among the health technicians, there are 68,400 medical practitioners, 4.13 medical practitioners per thousand population, 87,800 registered nurses, 5.30 registered nurses per thousand population. 66.07% of health technicians work in hospitals, 27.56% work in primary healthcare institutions, 4.96% work in public health institutions and 1.40% work in other institutions.

In 2019, the total number of visits to health institutions in Chengdu reached 164.51 million. Among all the visits, hospital visits account for 82.32 million, 73.82 million visits for primary healthcare institutions, and 8.36 million visits to other medical institutions. Among the hospital visits, there are 65.96 million visits (80.12% of total hospital visits) to public hospitals and 16.37 million visits (19.88%) to private hospitals.

In 2019, a doctor (including the doctor of hospitals and primary health care institutions) treated 9.30 patients per day on average among health institutions (excluding village clinics). The utilization rate of hospital beds among health institutions is 87.35%. The average LoS of a hospitalized patient is 9.40 days.

Then Figure 10 shows the stream of the hospitals' income in Chengdu, as reported in the statistics of Chengdu Healthcare Security Administration. According to the stream of the hospitals' income in Chengdu, the income from UEBMI local patients made up 20% of the total income. The income from URBMI local patients made up 50% and the income from intercity patients and uninsured patients made up 20%.

---

[4] [3] In China, according to the government document that The Measures for The Administration of The Hospital Grade (1989), health institution are classified in 3 tiers: Tier-1hospitals are township clinics that contains less than 100 beds. They are responsible for providing preventive care, minimal health care and rehabilitation services. Tier-2 hospitals tend to be affiliated with a medium size city, county or district, the number of beds that they contain is more than 100 but less than 500. They are responsible for providing comprehensive health services, as well as medical education and conducting research on a regional basis. Tier-3 hospitals round up the list as comprehensive or general hospitals with the bed capacity over 500. They are responsible for providing specialist health services, perform a bigger role regarding medical education and scientific research and they serve as medical centers providing care to multiple regions.

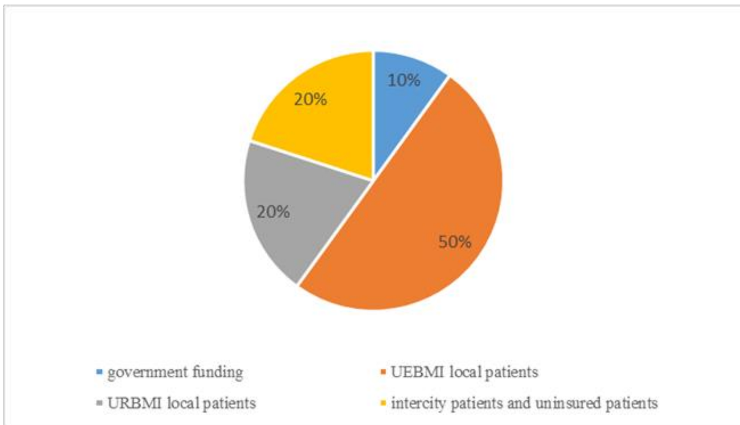


Figure 10. The stream of hospitals' incomes in Chengdu

Thus, even though the PPS reform Chengdu in 2011 only targeted UEBMI and URBMI patients, due to the income from UEBMI and URBMI patients made up 70% of the total income, the PPS effect was still strong enough to motivate hospitals to decrease bills.

4.2 Data sources

In chapter 5, for evaluating the GB effects on patients and hospitals, we use the hospital level data collected from Chengdu Healthcare Security Administration, which is annual data in the form of annual report from the hospital submitted to Chengdu Healthcare Security Administration.

In chapter 6, to evaluate the causal impact of first PPS reform in Chengdu in 2011, we use discharge patient data that was collected from Chengdu Healthcare Security Administration. The data is a random extract, covering 10% of all enrolees of UEBMI and URBMI in Chengdu city.

In chapter 7, to evaluate the causal impact of the second PPS reform in Chengdu in 2018, we use the frail elderly patient data collected from Chengdu Healthcare Security Administration, which were consisted of the frail elderly patients who were the enrolees of the long-term care insurance in Chengdu.

Conclusion

This chapter introduced the context and health system in Chengdu. Moreover, it introduced the data resources for the empirical studies employed in this dissertation. We next turn to the evaluation of the GB reform in Chengdu.

5. The effects of GB on patient payments and health outcomes in Chengdu

Summary

Background: In June 2013, Chengdu, the capital city of Sichuan province, China, introduced a GB reform targeting enrollees of UEBMI, URBMI, and NRCMS from both tier 2 (<500 beds) and tier 3 (>500 beds) hospitals.

Objectives: As we drew on the theoretical framework set out in Chapter 2, subsequently, we evaluated the GB effect on bills, average LoS, and hospital the number of patients.

Data: Hospital-level data collected from the Chengdu Healthcare Security Administration spanning from 2009 to 2016, comprising 1003 observations, was utilized.

Methods: Interrupted Time Series (ITS) estimation was employed to investigate initial changes and changes in trends for bills, LoS, and the number of patients following GB adoption. The Cumby–Huizinga test was utilized for residual autocorrelation, and the model was re-estimated, specifying lag(2) for bills and the number of patients, and lag(1) for LoS, to account for autocorrelation accurately.

Results: In the first year of GB adoption in 2013, medical bills increased by 39.8%, with no significant change in trend thereafter. LoS remained unaffected in the first year of GB adoption, and its trend remained unchanged after GB. The number of patients increased annually by 3.3%, with a 12.7% decrease in the first year of GB adoption, showing no change in the post-policy trend. After adjusting for residual autocorrelation, GB adoption decreased the trend by 19.8%, with no impact on average LoS or the number of patients.

Conclusion: In Chengdu, after GB adoption, bills increased, and the number of patients decreased in the first year of the reform. After adjusting for autocorrelation, GB adoption led to a 19.8% decrease in trend but had no effect on average LoS or the number of patients. Thus, the results of number of patients were consistent with theoretical predictions, while the results of bills and LoS were not.

5.1 The GB reform in Chengdu

In June 2013, the GB reform was instituted in Chengdu, targeting tier 2 and tier 3 hospitals and beneficiaries of UEBMI, URBMI, and NRCMS. This reform nearly encompassed all hospital discharges in Chengdu (tier 1 hospitals being clinics in China), aiming to significantly influence hospital behavior by incentivizing reductions in payments and unnecessary treatments for patients

Similar to other regions in China, Chengdu implemented the GB reform using an expenditure cap and 'soft target' approach. Under Chengdu's GB system, a per-case GB payment was calculated by averaging bills from the previous year across all tier 2 and tier 3 hospitals. The GB payment to a hospital equaled the per-case GB payment multiplied by the quantity of cases from the past year for that hospital. Therefore, unlike a regional GB payment, Chengdu's adoption of GB in 2013 initially followed a basic GB payment scheme. Additionally, hospitals in Chengdu could receive partial reimbursement for expenditures exceeding the benchmark under special conditions. Thus, Chengdu's GB system encountered similar potential shortcomings as other regions in China, whereby the basic GB payment with a 'soft target' might not have been sufficiently efficient for cost containment.

To assess the efficiency of GB adoption in Chengdu for cost control and its impact on hospital performance, we compare the trends in bills, LoS, and the number of discharged patients from every hospital in Chengdu before and after GB adoption, estimating the GB effect in the empirical sections.

As detailed in Chapter 3, three agents are involved in the payment model: the patient, the hospital, and the payer. The patient pays out-of-pocket expenses as a portion of the bill to the hospital, while the payer, Chengdu Healthcare Security Administration in China, covers the remaining portion of the bill and monitors the quality of health services provided by hospitals. When investigating the GB effects, we focus on the relationship between the hospital and the payer, predicting that the GB reform will reduce bills and LoS for hospitals, concurrently leading to a decrease in the number of treated patients.

5.2 Empirical strategy

5.2.1 Interrupted time series (ITS) analysis

We employ ITS analysis to assess the GB effect on hospital bills, LoS, and patient numbers. Utilizing hospital-level data, as detailed in the subsequent section, we calculate the average bill per patient in a hospital by dividing the hospital's income by the number of patients, and the average LoS per patient in a hospital by dividing the occupied bed days by the number of discharged patients. We estimate the following equation:

$$Y_{jt} = \alpha + \beta Year_t + \gamma After_{jt} + \delta After * Year_{jt} + \rho \boldsymbol{X_{jt}} + \epsilon_{jt} \quad (25)$$

The outcome variable Y_{jt} represents bills, LoS, or the number of patients for hospital j in year t. $After_{jt}$ denotes the policy indicator, taking the value 1 for post-GB adoption and 0 for pre-GB adoption, reflecting the immediate GB effect. $After * Year_{jt}$ is the interaction of year and policy, indicating the sustained GB effect in each year following the policy. $\boldsymbol{X_{jt}}$ is a vector of characteristic variables of hospital j in year t, which includes tier, number of medical staff, number of beds and amount of valuable equipment (exceeding CNY10,000) in the hospital.

5.2.2 Correcting for autocorrelations

An essential assumption of the linear regression model is the independence of error terms. When error terms exhibit correlation over time, known as autocorrelation, the true standard deviation of regression coefficients may be significantly underestimated. To assess residual autocorrelation, we employ the Cumby-Huizinga test (Cumby and Huizinga, 1992). Upon confirming residual autocorrelation, we refine the model by specifying lag(2) for bills and the number of patients, and lag(1) for LoS, ensuring proper adjustment for autocorrelation.

One potential issue with ITS estimation is the presence of residual autocorrelation. We simplify the ITS equation as $y_t = \alpha + \beta X_t + \mu_t$, and $\mu_t = \rho\mu_{t-1} + v_t$, where v_t follows an independent identically distribution, and the lag (1) of the equation is $y_{t-1} = \alpha + \beta X_{t-1} + \mu_{t-1}$. And we take the difference of y_t and ρy_{t-1}, and get $y_t - \rho y_{t-1} = \alpha(1-\rho) + \beta X_t - \rho\beta X_{t-1} + \mu_t - \rho\mu_{t-1}$. Since $\mu_t - \rho\mu_{t-1} = v_t$, the auto-correlation does not exist in the residuals of $y_t - \rho y_{t-1}$. Thus, incorporating lagged terms into the ITS equation resolves the autocorrelation issue in residuals.

Furthermore, we conduct the Cumby-Huizinga test to assess residual autocorrelation. According to this test, residuals for bills and the number of patients follow an AR(2) process, while residuals for LoS adhere to an AR(1) process. Consequently, we re-estimate the model by specifying lag(2) for bills and the number of patients, and lag(2) for LoS, ensuring accurate adjustment for autocorrelation.

5.2.3 Heterogeneity analysis

Given the disparity in budgets between tier 2 and tier 3 hospitals, their responses to the GB policy may vary. Tier 3 hospitals typically operate on a larger scale and employ medical staff with higher qualifications and professional ranks compared to tier 2 hospitals. Consequently, tier 3 hospitals face higher performance standards. For instance, tier 2 hospitals tend to treat less severe patients and may plausibly reduce treatment intensity to cut costs, a strategy less feasible for tier 3 hospitals due to their higher proportion of severe cases. Moreover, following GB adoption, tier 2 hospitals may refer severe cases to tier 3 hospitals to capitalize on financial incentives, potentially resulting in a decrease in patient numbers.

To assess whether the effects of GB differ across hospital tiers, we conduct a heterogeneity analysis, examining the impact of GB separately for each tier.

5.3 Data

We utilize hospital-level data obtained from the Chengdu Healthcare Security Administration, comprising annual reports submitted by hospitals from 2009 to 2016. The dataset encompasses 1003 hospital-level observations, including 276 from tier 3 hospitals and 727 from tier 2 hospitals. Information captured includes total patient income, the number of patients, number of discharges, and bed utilization.

5.4 Results

5.4.1 Descriptive analysis

Figure 11[6] depicts the trends of outcome variables, Figure 11-1 to Figure 11-3 are the trends of bills (adjusted for inflation), LoS, and the number of patient across the study period, respectively. Bills showed a consistent increase since 2009, with a decline in 2011, though this downward trend ceased post-2013. LoS experienced a reduction in the year preceding GB adoption but subsequently increased after the first year of implementation. The number of patients exhibited continuous growth until the year following GB adoption, after which the increasing trend appeared to stabilize from 2014 onward.

Table 4 provides the mean and standard deviation for both outcome variables and key

[6] In Figure 11, the blue line represents the trend of the outcome variables for the hospitals in the study period. The X-axis is the time point for each year. And the Y-axis is bills, LoS and the total number of patients respectively.

explanatory variables, categorized by tier 2 and tier 3 hospitals.

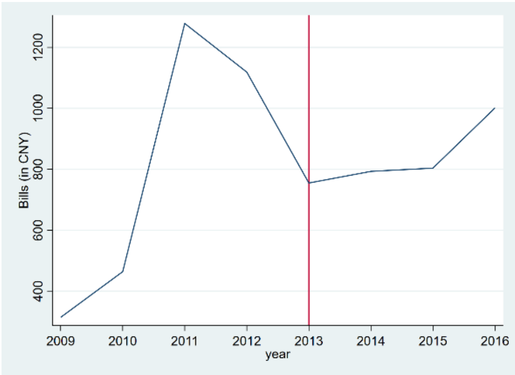


Figure 11-1. The trend of bills

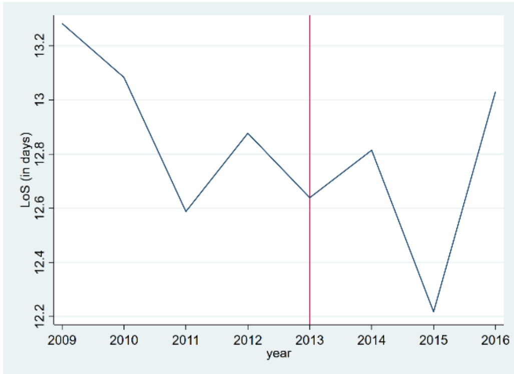


Figure 11-2. The trend of LoS

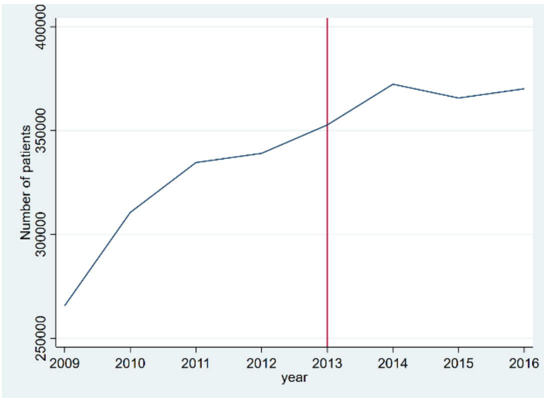


Figure 11-3. The trend of number of patients

Figure 11. The trends of outcome variables

Table 4. Descriptive statistics for outcome variables and key explanatory variables

	Tier 2 hospitals		Tier 3 hospitals	
Variable	Mean	Std. Dev	Mean	Std. Dev
Bills (in CNY)	837.48	3990.51	832.52	1040.88
Average LoS	13.01	13.75	12.21	8.16
The number of patients (in thousands)	174.39	151.72	794.84	904.36
The portion of post-policy observations to the full sample (After=1)	54.61%		69.20%	

5.4.2 Empirical results

Table 5. The results of ITS estimation

VARIABLES	(1) Bill	(2) LoS	(3) The number of patients
year	0.0117	0.0331	0.0333**
	(0.0147)	(0.231)	(0.0169)
after	0.398***	-0.958	-0.127*
	(0.0607)	(1.046)	(0.0730)
After*year	0.0167	-0.182	-0.0180
	(0.0257)	(0.448)	(0.0311)
Observations	982	981	982
Number of hospital ID	401	400	401

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As illustrated in Table 5, the first year of GB adoption in 2013 witnessed a significant 39.8% increase in bills, with no subsequent change in the trend during the post-policy period. Average LoS remained unchanged in the initial year of GB adoption, with no alteration in trend observed post-implementation. Notably, the number of patient displayed a substantial 3.3% yearly increase before experiencing a 12.7% decrease in the year following GB adoption. However, there was no discernible change in the trend during the post-policy period.

5.4.3 Correcting for autocorrelation

Table 6 presents the results of ITS estimation after accounting for residual correlation. For the estimation concerning bills, 75 hospitals were included in the sample, with 192 hospitals omitted for unevenly spaced observations (e.g. the hospitals had observations from 2009 to 2011, and from 2013 to 2016, while the observation in 2012 was missing), and 134 hospitals excluded for having fewer than 2 observations. In the LoS estimation, 75 hospitals were part of the sample, with 192 omitted for unevenly spaced observations, and 133 excluded for fewer than 2 observations. Regarding the number of patients, the sample comprised 40 hospitals, with 192 omitted due to unevenly spaced observations, and 169 excluded for fewer than 3 observations.

Bills increased by 11.9% annually, experiencing no change post the first year of GB adoption. Additionally, the post-policy trend of bills decreased by 17.8% compared to the pre-policy trend. However, the GB effects on average LoS or the number of patients, whether for initial changes or changes in trends, became insignificant.

Table 6. The results of ITS estimation after adjusting residual correlation

VARIABLES	(1) Bills lag(2)	(2) LoS lag(1)	(3) The number of patients lag(2)
year	0.119*	0.545	0.00138
	(0.0684)	(0.692)	(0.0708)
after	-0.0250	-0.653	0.0919
	(0.130)	(1.747)	(0.174)
After*year	-0.178**	-0.986	0.0829
	(0.0755)	(0.854)	(0.0865)
Observations	228	228	155
Number of hospital ID	75	75	40

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Upon adjusting for residual autocorrelation, GB adoption appeared to decrease the increasing trend in bills but had no discernible effect on average LoS or the number of patients.

5.4.4 Heterogeneity analysis

To examine the impact of the GB initiative on hospitals across different tiers, ITS analysis was conducted separately for tier 3 and tier 2 hospitals.

As indicated in Table 7, bills for tier 3 hospitals escalated annually by 8.1%, surged by 27.7% in the first year following GB implementation, and the post-policy trend declined by 10.3% relative to the pre-policy trend. Additionally, the introduction of GB did not influence the LoS or patient numbers in tier 3 hospitals.

Table 8 reveals that for tier 2 hospitals, bills rose by 38.0% in the initial year of GB implementation, and the post-policy trend in bills increased by 7.2% in comparison to the pre-policy trend. The LoS in tier 2 hospitals diminished by three days in the first year of GB implementation, with no subsequent change in the average LoS trend. Although the annual increase in patient numbers was 2.5%, there was no initial change following the implementation of GB; however, the post-policy trend showed a 5.1% reduction compared to the pre-policy trend.

Table 7. The results of ITS estimation for tier 3 hospitals

VARIABLES	(1) Bills	(2) LoS	(3) The number of patients
year	0.0808***	0.318	-0.0353
	(0.0223)	(0.327)	(0.0222)
after	0.277***	-1.418	-0.100
	(0.0746)	(1.252)	(0.0691)
After*year	-0.103***	-0.790	0.0444
	(0.0314)	(0.514)	(0.0299)
Observations	269	269	269
Number of hospital ID	167	167	167

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8. The results of ITS estimation for tier 2 hospitals

VARIABLES	(1) Bills	(2) LoS	(3) The number of patients
year	-0.0119	0.0753	0.0245*
	(0.0165)	(0.259)	(0.0148)
after	0.380***	-3.029***	0.0150
	(0.0704)	(1.160)	(0.0640)
After*year	0.0722**	0.0506	-0.0511*
	(0.0304)	(0.507)	(0.0277)
Observations	713	712	713
Number of hospital ID	372	371	372

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Post-GB implementation, tier 2 hospitals recorded an increase in bills during the first year, contrary to predictions of reduced costs. Conversely, the LoS reduction in the first year and the downward trend in patient numbers during the post-policy period were consistent with the forecasts that GB would reduce both the LoS and patient admissions in hospitals.

According to the results, bills for both tier 3 and tier 2 hospitals increase in the first year following the implementation of GB. Subsequently, the post-policy trend in bills decreases for tier 3 hospitals and increases for tier 2 hospitals. For LoS, there is no change observed in tier 3 hospitals, while a decrease is noted in the first year of GB for tier 2 hospitals. Additionally, the number of patients shows a declining trend in tier 2 hospitals post-GB. It is speculated that tier 2 hospitals may refer more severe cases to tier 3 hospitals for financial incentives, leading to a noted decrease in patient numbers at tier 2 facilities.

It is noteworthy that the combined number of hospital IDs in Table 7 (tier 3 hospital analysis) and Table 8 (tier 2 hospital analysis) exceeds the total in Table 5 (overall analysis). This discrepancy arises because the tiers of some hospitals changed; for example, a hospital classified as tier 2 until 2011 could be upgraded to tier 3 in 2012. Consequently, when performing subgroup analysis by tier, a hospital ID could be included in either the tier 3 or tier 2 group.

Due to the lack of sufficient data groups—with some groups omitted because they were either equidistantly spaced or had fewer than two observations—we cannot conduct the ITS estimation after adjusting for residual correlation.

5.5 Discussion

To evaluate the impact of GB adoption, our study assessed changes in bills, LoS, and patient numbers at hospitals in Chengdu following the implementation of GB.

In our main analyses, we utilized ITS estimation to assess both immediate and long-term impacts on bills, LoS, and patient numbers. ITS results indicated immediate effects on bills and patient counts, with a 39.8% increase in bills and a 12.7% decrease in patient numbers in the first year of GB implementation. However, GB had no immediate or lasting effect on average LoS.

After adjusting for residual correlation, the post-policy trend in bills showed a 17.8% decrease, rendering the immediate effects on bills statistically insignificant. Moreover, there were no immediate or sustained effects on average LoS or patient numbers.

The findings that GB instantly increased bills and had no impact on average LoS contradicted predictions that GB would reduce both metrics, suggesting that GB adoption in Chengdu did not effectively control healthcare expenditures or encourage a reduction in treatment intensity. These outcomes may be attributed to the hospitals shifting financial burdens to patient OOP (unfortunately, our data only covered OOP starting in 2013, preventing observation of changes prior to GB adoption). Additionally, the implementation of GB in Chengdu featured a 'soft' target, allowing hospitals to receive partial reimbursement for expenditures exceeding benchmarks under certain conditions. This 'soft' target likely diminished the incentive for hospitals to reduce healthcare spending, as it was common for hospital expenditures to surpass the benchmarks in Chengdu.

Figures 12 and 13 present forest plots from GB studies, updated to include results from this study labeled as Zhao (2023. Our main findings reveal that neither bills, with inflation control, nor average LoS decreased immediately following GB implementation, contradicting the majority of prior research. This discrepancy may be attributed to two factors. Firstly, unlike other studies that utilized patient-level data, our research employed hospital-level data to compute average bills and LoS, potentially reducing accuracy. Secondly, GB in Chengdu employs a 'soft' target allowing partial reimbursement for expenditures exceeding benchmarks under specific conditions, which has hindered cost containment (Chen and Fan, 2016). Conversely, GB reforms in other regions are implemented more stringently, offering hospitals stronger incentives to reduce costs.

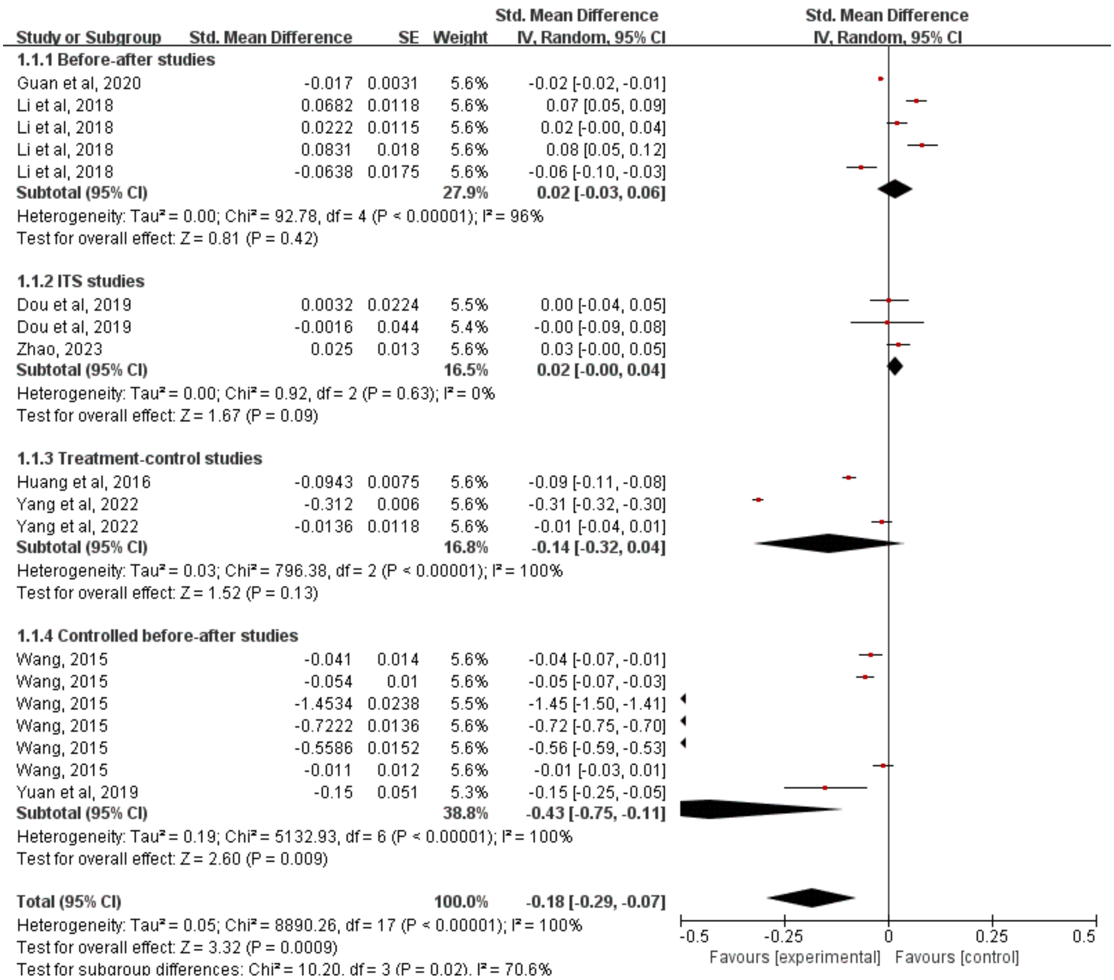


Figure 12. Forest plot of GB effect on payments including results of this study

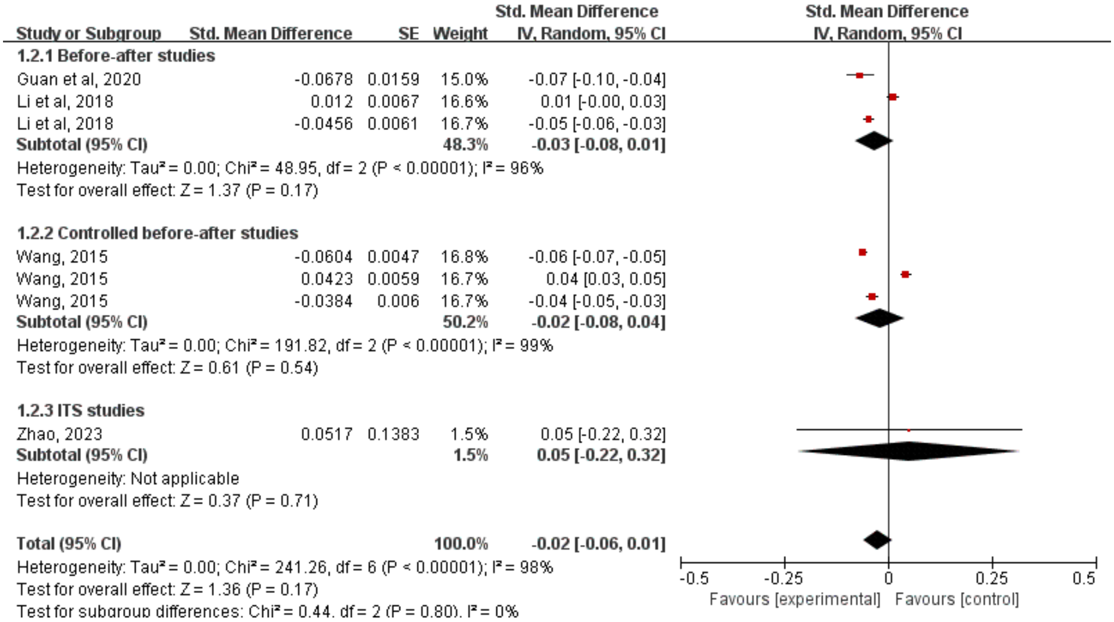


Figure 13. Forest plot of GB effect on LoS including the results of this study

Additionally, our study identifies heterogenous responses among tier 2 and tier 3 hospitals under GB. Tier 3 hospitals initially saw a bill increase of 27.7% and a subsequent trend decrease of 10.3%. Neither the values nor the trends of LoS and patient numbers showed immediate changes. In contrast, for tier 2 hospitals, GB implementation initially affected bills (increasing by 38.0%) and LoS (decreasing by 3.0 days) in the first year, with sustained effects on the trends of bills (increasing by 7.2%) and patient numbers (decreasing by 5.1%).

However, our study has several limitations. Primarily, the dataset is unbalanced, leading to the exclusion of many groups due to irregular spacing or having fewer than two observations. Consequently, after adjusting for residual autocorrelation, the reduced number of hospitals in the sample limits the validity of the ITS estimation when adjusted for autocorrelation.

Missing values may arise from two sources. Initially, Chengdu implemented an electronic information system post-2010, leading hospitals to transition from manual to electronic records. This shift resulted in some data being lost during the transfer. Secondly, only hospitals that entered into agreements with the payer were required to report data to the government; therefore, any hospital that signed a contract post-2010 would lack prior data.

A second limitation involves the absence of information on OOPs, precluding analysis of whether hospitals shifted financial burdens from reimbursed payments to OOPs. A third limitation is the unavailability of additional quality indicators, such as self-reported health status of discharged patients or readmission rates, which restricts testing of health service quality assumptions in the empirical analysis.

The empirical findings suggest several policy implications. Firstly, it appears that hospitals in Chengdu are unable to reduce bills under GB. Given the 'soft' target GB payment scheme, where hospitals may receive partial reimbursement for expenditures exceeding benchmarks under specific conditions—variations in reimbursement may be substantial enough to nearly offset the excess costs. Hospitals may exaggerate certain details (e.g., the severity of patient diagnostics to increase the count of severe cases) in annual reports to secure higher reimbursements. This practice shifts the hospital's focus from controlling costs to maximizing reimbursement claims, undermining the objectives of GB. To counteract this, payers could enforce stricter verification of hospital reports and provide stronger incentives for cost reduction. Secondly, while no change in patient numbers is noted in tier 3 hospitals post-GB, a decrease is observed in tier 2 hospitals. This suggests a need for more thorough oversight of hospital performance under GB, particularly to ensure that hospitals do not reject patients as part of strategies to manipulate service metrics.

5.6 Conclusion

In Chengdu, following the adoption of GB, there was an increase in bills and a decrease in the number of patients during the first year of reform. After adjusting for autocorrelation, GB adoption resulted in a 19.8% decrease in the trend of bills but had no impact on average LoS or patient numbers. According to the heterogeneity analysis, tier 3 hospitals saw an increase in bills during the first year of GB, followed by a decrease in the post-policy trend in bills. For tier 2 hospitals, there were increases in bills in the first year of GB and in the post-policy trend in bills, a decrease in average LoS during the first year of GB, and a decrease in the post-policy trend of patient numbers.

6. The effects of PPS on patient payments and health outcomes in Chengdu in 2011

Summary

Background: In May 2011, Chengdu, the capital city of Sichuan province in China introduced a PPS. The policy targeted enrollees of the URBMI and UEBMI and was implemented across ten diseases and among 147 tier 2 (<500 beds) and 42 tier 3 (>500 beds) hospitals.
Objectives: In this study, as a theoretical framework set out in Chapter 2 predicted changes associated with the payment reform, including on expenditure and quality of care, we then evaluated the PPS policy effect on bills, OOPs, LoS and 30-day emergency readmission rates.
Data: We used patient-level data collected from Chengdu Healthcare Security Administration from quarter 1 in 2010 to quarter 4 of 2013. The dataset comprised 19489 observations in 167 hospitals.
Methods: Following the allocation of patients into PPS targeted (treatment) group and control group based on diagnoses codes, we performed a difference-in-differences (DiD) estimation. As a robustness check groups were formed using Propensity Score Matching (PSM) methods.
Results: We found that the adoption of PPS resulted in a 10.0% reduction in total health payments, a 10.0% reduction in OOPs, 0.935 (0.892) times lower rate for LoS, and a 2.4% rise in the probability of 30-day emergency readmission compared to the counterfactual. Our analysis of heterogenous effects revealed that the decrease in 30-day readmissions was concentrated among tier 2 hospitals, with no change observed in tier 3 hospitals.
Conclusion: In Chengdu, after PPS adoption, total health payments, OOPs and LoS for PPS targeted patients decreased, while 30-day readmissions increased. The increase in re-admissions was observed in tier 2 but not in tier 3 hospitals.

6.1 The PPS reforms in Chengdu

Chengdu implemented PPS for the enrollees of URBMI and UEBMI who were hospitalized for treatment in the tier 2 and tier 3 hospitals in 2011.

This PPS project in Chengdu targeted the following 10 diseases with a relatively large volume of patients and corresponding operations: acute appendicitis treated by appendicectomy, acute mastitis treated by abscess incision drainage, benign prostatic hyperplasia treated by transurethral resection of prostate, ureteral calculus treated by ureteroscopy taken under rubble nephrolithotomy, idiopathic thrombocytopenic purpura, spontaneous pneumothorax treated by internal medicine conservative treatment, spontaneous pneumothorax treated by pulmonary bullectomy, benign ovarian tumor treated by ovarian cyst excision, external hemorrhoid treated by thrombotic external hemorrhoidectomy and uterine fibroid treated by myomectomy. Thus, in this wave of reform, simple case-mix PPS was applied in Chengdu, and the PPS reimbursement standard of targeted diseases are listed in Table 9.

Table 9. The PPS reimbursement standard in Chengdu

Diagnosis Name	Operation Name	Tier	Total price	Insured price	OOP price
Acute appendicitis (AA)	Appendectomy	Tier 3	5410	3730	1680
		Tier 2	3400	2600	800
Acute mastitis (AM)	Incision and Drainage of Abscess	Tier 3	4310	3060	1250
		Tier 2	4250	3480	770
Benign Prostatic Hyperplasia (BPH)	Transurethral Resection of the Prostate	Tier 3	11330	8710	2620
		Tier 2	9480	7690	1790
Ureteral Calculi (UC)	Transurethral Ureteroscope Lithotripsy	Tier 3	8200	6270	1930
		Tier 2	5510	4430	1080
Idiopathic Thrombocytopenic Purpura (ITP)		Tier 3	14120	10550	3570
		Tier 2	10650	8010	2640
Spontaneous pneumothorax (conservative medical treatment) (SPC)		Tier 3	7160	5950	1210
		Tier 2	4880	4290	590
Spontaneous pneumothorax (surgical operation) (SPS)	Pulmonary Resection	Tier 3	24290	17000	7290
		Tier 2	15400	11550	3850
Benign Ovarian Tumor (BOT)	Ovarian Cyst Removal	Tier 3	7050	5360	1690
		Tier 2	4660	3790	870
Thrombotic external hemorrhoids (TEH)	Thrombotic External Hemorrhoidectomy	Tier 3	3400	2380	1020
		Tier 2	2890	2180	710
Uterine fibroids (UF)	Uterine fibroids removal	Tier 3	7990	5590	2400
		Tier 2	5100	3980	1120

The PPS tariffs contained the insured price and OOP price, both with the fixed amount. It was noteworthy that the tariffs for tier 3 and tier 2 hospitals were separated: Tier 3 hospitals round up the list as comprehensive or general hospitals, they tended to receive the severe patients and employ more advanced equipment and technology compared to tier 2 hospitals, which consequently caused the higher costs in tier 3 hospitals.

Recalling the set-up in theoretical framework, patients paid the hospital the amount of OOPs when they discharged, and this payment should be fixed; then the hospitals claimed the reimbursement of insured payments from the Chengdu Healthcare Security Administration, any profit or loss of insured payments should be afforded by the hospital. For example, if a patient with an acute appendicitis received an appendectomy in a tier 3 hospital, the total payment for his/her case was CNY 5410, he/she needed to pay the amount of CNY 1680 to the hospital, and the hospital claimed the reimbursement of CNY 3730 from Chengdu Healthcare Security Administration.

Hence, the reimbursement process design in Chengdu made hospitals strictly observant of the PPS budget, and hospitals needed to reduce costs under such a strong constraint[7].

In Chengdu, since only 10 diagnoses were targeted by PPS in the 2011 wave, the hospitals did not control Case Mix Index, and the patients with complications would be removed from DRG clinical path. The Notice on Implementing DRGs in Chengdu (Chengdu Healthcare Security Administration, 2011) stipulates that if the treatment deviated the clinical path due to complications, the hospitals should inform the patients, retreat them from PPS and return to FFS. Thus, due to the asymmetric information in the treatment, the physicians might induce the patients to implement FFS for financial incentives.

To further control for the unreasonable increase in health expenditures, the State Council of China issued the Guidelines on Further Deepening the Payment Reform under Basic Medical Insurance Schemes in 2017, which comprehensively promoted the multiple medical insurance payment methods mainly relying on DRGs under PPS. In September 2017, Chengdu issued the Notice on Deepening the Medical System Reform and it pointed out that the relevant departments (including Municipal Development and Reform Commission, Health and Family Planning Commission, Chengdu Healthcare Security Administration and Human Resources and Social Security Bureau) should comprehensively promote the PPS, and the number of DRGs should be no less than 100 in this wave. On December 20 in 2017, Chengdu Municipal Development and Reform Commission issued the reimbursing Standard for 101 Diseases in Chengdu Municipal Public Hospitals, which clarified the tariffs for 101 PPS target diagnoses.

According to the PPS reform background in chapter 1, since this reform was in the form of single disease PPS, the local government in Chengdu set up a PPS tariff for each diagnosis separately for tier 2 and tier 3 hospitals, and this PPS tariff was directly announced in the official document (Notice on Implementing PPS among Several Diagnoses) that was issued in April 2010. Being based on the experience of PPS reform from other cities (e.g., Xiangtan, Guangzhou, and Shanghai), with the PPS tariffs being calculated by averaging the costs for patients in the last 3 years, we assume the PPS tariffs in Chengdu to be set up in the similar way.

Due to the lack of official statistics on patients' average costs for the PPS targeted

[7] In China, after 1980s, with the end of centralized economy for health service sectors, the funding from government for public hospitals only accounts for a limited portion of their total revenue. According to the Health Policy Report for China, by the end of 2009, the government funding only makes up 10% of the revenue of public hospitals, while the 90% of revenue comes from their health service. Hence, the deficit of the public hospitals should be afforded by themselves, the bankruptcy hospitals can exit the health service market, or they can be acquired/merged by private hospitals.

diagnoses, we calculate the average costs for these diagnoses by using the samples from the data, and we calculate the FFS bill for each targeted diagnosis by averaging the treatment costs of the samples who had the same diagnosis but in the pre-policy stage. The treatment cost for a patient is the sum of costs for all the services, including prescription costs.

By comparing the PPS tariffs and FFS bills, we can have a preliminary judgement on the performance of the hospitals. In other words, do PPS tariffs approximate FFS bills? If not, by what margin do PPS tariffs deviate from FFS bills? Do PPS tariffs set up a constraint for hospitals and motivate them to reduce bills for patients?

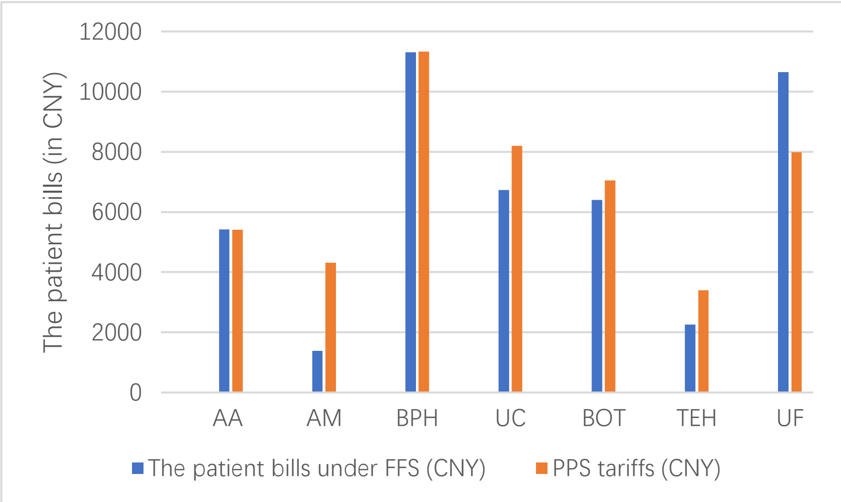


Figure 14. The comparison of FFS bills and PPS tariffs for tier 3 hospitals

Figure 14 shows the comparison of PPS tariffs and FFS bills[8] among tier 3 hospitals, which covers the 9 PPS targeted diagnoses from our data. For most of the diagnoses, the PPS tariffs are similar to FFS bills, while the FFS bills for 2 diagnoses are distinct with the PPS tariffs for them.

Figure 15 shows the comparison of PPS tariffs and FFS bills among tier 2 hospitals. The PPS tariffs for 4 diagnoses (acute mastitis, benign prostatic hyperplasia, ureteral calculi and spontaneous pneumothorax with surgical operation) are larger than the corresponding FFS bill, with the PPS tariff for the other targeted diagnoses aligning closely with the FFS bill.

[8] We have controlled the inflation of health service costs when calculating the average FFS bills before PPS policy.

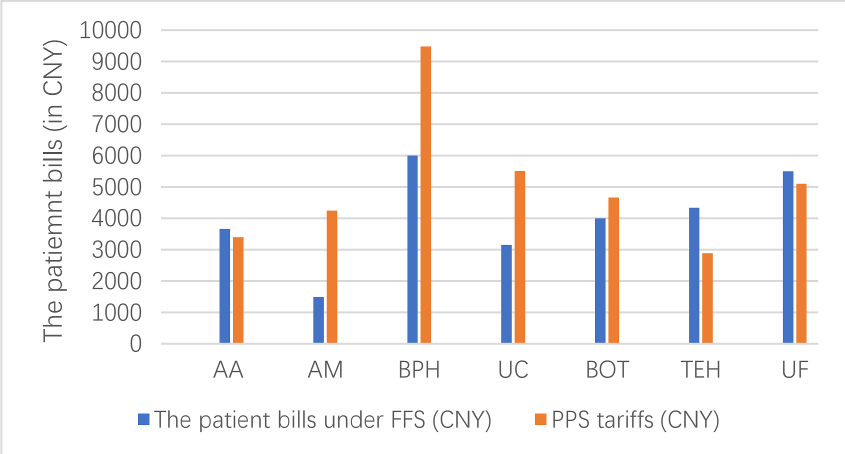


Figure 15. The comparison of FFS bills and PPS tariffs for tier 2 hospitals

According to the comparison of PPS tariffs and FFS bills, it remains unclear to what extend cost difference have the potential to impact behavior given that they align quite closely across most clinical groups. To evaluate the general policy effect on bills for patients, we will employ a DiD estimation; and to test whether tier 2 and tier 3 hospitals have different reactions to PPS policy, we will evaluate the policy effect for them separately in the section of heterogeneity analysis.

Being based on the predictions from theoretical framework, bills, OOPs and LoS for patients will decrease after the policy because the hospital tends to reduce the bills for patients and reduce the treatment intensity; and the unplanned 30-day readmissions will increase after the policy, suggesting a decrease in health outcomes because the decrease in treatment intensity might cause insufficient care for the patients.

6.2 Empirical strategy

6.2.1 Difference-in-difference estimation

We use a DiD approach to estimate the causal impact of the PPS policy implemented in May 2011. We estimate the impact on bills, OOPs and unplanned 30-day readmissions by applying the following equations, estimated using ordinary least squares estimators:

$$y_{ijt} = \beta_0 + \delta PPS * Post_{ijt} + \beta_1 PPS_{ijt} + \beta_2 Post_{ij} + \alpha los_{ijt} + \gamma X_{ijt} + \theta_j + \varphi_t + \varepsilon_{ijt} \quad (26)$$

In Eq. (26), the outcome y_{ijt} can be total amount of health service price on the bill, OOP and unplanned 30-day readmission for patient i discharged from hospital j at time t.

For the unplanned 30-day readmission, if the patient experiences an unplanned readmission within 30 days at period t, y_{ijt} equals 1, otherwise y_{ijt} equals 0. We restrict our analysis to unplanned 30-day readmissions, which requires some definitional criteria. First, in China, if a patient was discharged from a higher tier hospital and readmitted to a lower tier hospital (e.g., a patient discharges from a tier 3 hospital and is readmitted to tier 2 hospital), the readmission is likely a planned recovery treatment. Therefore, this would not constitute an unplanned 30-day readmission. Second, aligned with previous studies that have used readmission rates as a quality indicator (Maurer, 2004, Demir et al., 2008, Rumball-Smith et al., 2009), we define unplanned readmissions as those where the patient is readmitted to a hospital that has the same or higher tier compared to the previous hospital within 30 days. As such, 30-day unplanned readmission is used as our indicator for a poor health outcome.

$PPS * Post_{ijt}$ is the interaction term of Post and PPS, with the coefficient yielding the DiD estimator. If patient i from hospital j is under PPS after implementing the policy, $PPS * Post_{ijt}$ equals 1, otherwise, $PPS * Post_{ijt}$ equals 0. PPS_{ijt} is the dummy, which equals 1 if the patient's diagnosis is under PPS (treatment group) and equals 0 if not (control group). By taking the first derivative, $\delta PPS_{ijt} + \beta_1$ represents the underlying difference in bills between PPS targeted patients and other patients, and β_1 reflects the difference between control group and treatment group before the policy. $Post_{ij}$ is a dummy to measure the time, which equals 0 if patient i is at pre-policy stage and equals 1 if patient i is at post-policy stage. According to the theoretical framework, the bills and unplanned 30-day readmissions are related to LoS, in Eq (18), LoS is included as a control variable. Controlling for patient-level characteristics, we include X_{ijt} as a vector of patient characteristics. It contains information on patient i's gender, age, type of health insurance and the 3-digit ICD10 diagnostic group of the patient. There may be unobservable factors other than the patients' characteristics that can influence the patients' bills and health outcomes. To diminish the impact of unobservable factors that are time invariant (*e.g.* the type and the location of the hospital), we employ the fixed effect model. θ_j controls for the fixed effect at hospital level; φ_t controls for time dummies (*i.e.*, quarter).

The equation for LoS has the same structure as Eq. (26) and is written as follows:

$$E(los_{ijt}) = \exp\left(\beta_0 + \delta PPS * Post_{ijt} + \beta_1 PPS_{ijt} + \beta_2 Post_{ij} + \gamma X_{ijt} + \theta_j + \varphi_t + \varepsilon_{ijt}\right) \qquad (27)$$

In Eq. (27), LoS is the outcome variable. Due to the assumed Poisson distribution of the dependent variable LoS, Eq. (27) is estimated using Poisson regression, which is not a linear model. By taking the exponent of the coefficient, we obtain the incidence rate ratio (IRR). In the empirical results, we report the IRR for LoS estimation.

6.2.2 Robustness checks

We conduct robustness checks by constructing alternative treatment and control groups using PSM, and by testing for the anticipation effect.

PSM+DiD
An important assumption for the DiD estimation is that treatment group and control group must be comparable and share a common trend before policy adoption. To the robustness of our results, we combine the DiD estimation with PSM to make more comparable treatment and control groups, and we apply the 1:1 nearest non-replacement matching.

Anticipation effect test
It is likely that hospitals receive information of implementing PPS policy in advance, so there might be an anticipation effect impacting on the estimation of our main model specification. To address this concern, we choose the time point which is one quarter earlier than the PPS adoption time as the time of policy intervention and estimate the policy effect by taking out anticipation period (Friebel et al., 2017).

6.2.3 Heterogeneity analysis

Since the PPS tariffs for tier 2 hospitals and tier 3 hospitals differs, tier 2 and tier 3 hospitals may not be faced with the same incentives, hence, they may not have the same motivation to reduce bills or reacting to PPS policy in line with our hypotheses. To test whether PPS policy effects differ by hospital tier, in the heterogeneity analysis section we evaluate the impact of PPS separately by hospital tier.

6.3 Data

To evaluate the causal impact of PPS in Chengdu, we use discharge patient data that was collected from Chengdu Healthcare Security Administration. All patients covered in this dataset are health insurance enrolees. The data is a random extract, covering 10% of all enrolees of Urban Employee Basic Medical Insurance (UEBMI) and Urban Residents Basic Medical Insurance (URBMI) in Chengdu city. The dataset contains a total of 927,107 observations and covers a period from January 2010 to December 2013, with the policy taking effect in May 2011(i.e., the second quarter of that year). We use quarterly patient-level data, thus, the period covered by the analysis starting from the first quarter of 2010 and ending at the fourth quarter of 2013.

Patient information includes the anonymized patient ID, gender, age, insurance status (i.e. enrolled in UEBMI or URBMI), International Classification of Diseases 10th Revision (ICD-10) code, date of admission, date of discharge, total payments, OOP, the date of admission and the date of discharge. Information about the patient´s hospital is also recorded, including the hospital ID, tier and ownership form.

Since the PPS policy was only implemented in tier-2 and tier-3 hospitals and targeted to 10 diseases, we restrict our study population to patients discharged from those hospitals. We identify patients with PPS targeted diseases (treatment group) and construct control groups of patients with similar diseases: if the first three digits of ICD-10 code are the same as the one for the diagnosis under PPS, the patient is selected into the control group.

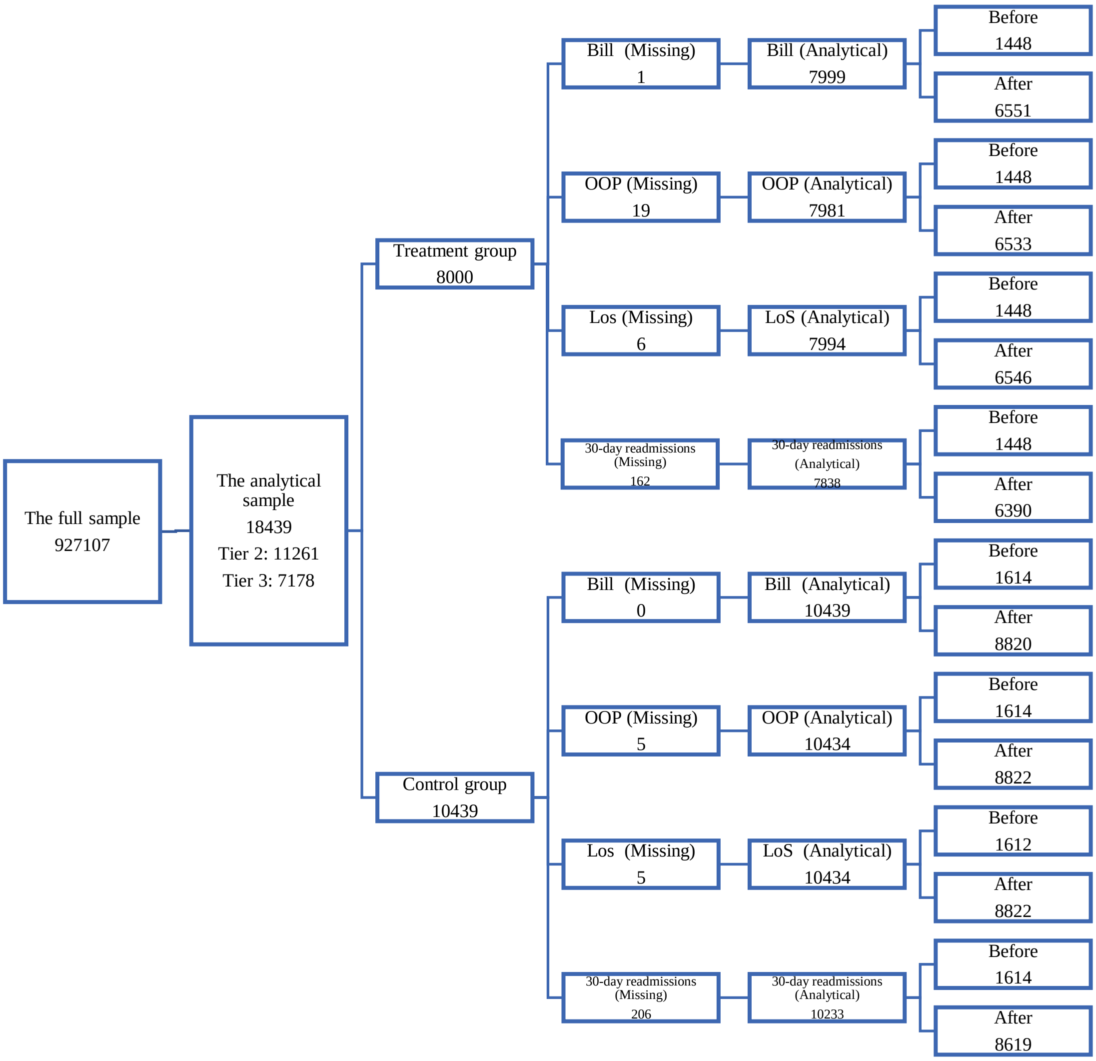


Figure 16. The number of observations at each stage of DiD

There are 19489 observations in our analytical samples, with 11765 observations from tier 2 hospitals and 7724 observations from tier 3 hospitals. As figure 16 shows, some were dropped from the analysis due to missing data. Information was available for almost everyone for Bills, OOP and LoS, but there were 226 patients for whom information about 30-day readmission was missing.

6.4 Results

6.4.1 Descriptive statistics

Table 10 and 11 depict the number of observations and hospitals in treatment and control group before and after PPS.

Table 10. The sample size of treatment and control group before and after PPS

	Before	After
Treatment	1449	6551
Control	1614	8825

Table 11. The sample size of treatment and control group before and after PPS

	Before	After
Treatment	107	151
Control	105	149

Figure 17 [9] show the trends of outcome variables, Figure 17-1 to Figure 17-4 respectively show the trends of bills, OOPs, LoS, and unplanned 30-day readmission rates of PPS targeted patients (treatment group) and the patients that are unaffected by PPS (control group) across the study period. The trends of treatment groups were similar as the trend of control groups before the policy, and the bifurcation of treatment and control group appeared after the policy.

And the visual inspection of trends appears to indicate possible policy effects appearing with the begin of the first quarter of 2011. This finding highlight anticipation effects given that we assume that hospitals received the information of implementing PPS policy in advance, which led to behavior change even before PPS policy was officially initiated. Moreover, this anticipated change only occurs to bills, OOPs and LoS, which appears plausible: the hospitals decreased the health payments and decreased the supply of health service once they received the massage before the policy was formally initiated, and the adverse outcomes for patients appeared subsequently. To investigate

[9] In Figure 17, the blue line represents the trend of treatment group and the red line represents the trend of control group. The X-axis is the time point for each quarter. And the Y-axis is the average bills (in CNY) in, the average OOP (in CNY), the average LoS and the average 30-day readmission rates respectively.

the impact of anticipation effects, in the robustness checks section, we assume the first quarter of 2011 to be the timepoint at which any policy effects may become visible, in line with the visual assessment.

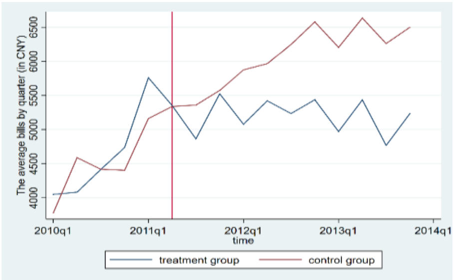


Figure 17-1. The trend of bills

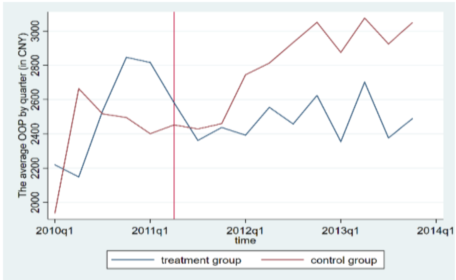


Figure 17-2. The trend of OOP

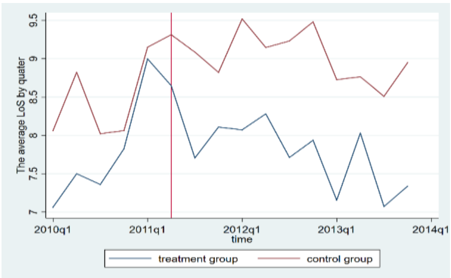


Figure 17-3. The trend of LoS

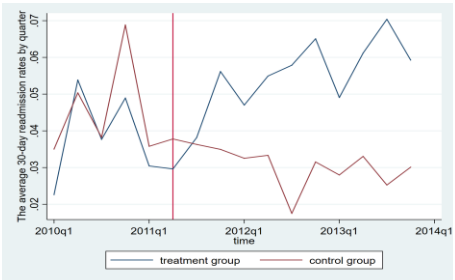


Figure 17-4. The trend of 30-day readmissions

Figure 17. The trends of outcome variables

Table 12 describes the mean value and standard deviation for the outcome variables and key explanatory variables, separately for patients treated at tier 2 hospitals and tier 3 hospitals.

Table 12. Descriptive statistics for outcome variables and key explanatory variables

	Tier 2 hospitals		Tier 3 hospitals	
Variable	Mean	Std. Dev	Mean	Std. Dev
Bill (in CNY)	4364.3	2709.1	7254.9	3720.5
OOP (in CNY)	1730.7	1367.8	3969.7	3226.0
LoS	8.0	5.3	9.0	6.6
Unplanned 30-day readmission rate	4.2%	0.2	4.5%	0.2
PPS targeted and post-policy group (PPS*Post=1)	35.1%	0.5	36.2%	0.5
PPS targeted group (PPS=1)	45.5%	0.5	40.1%	0.5
Post-policy group (Post=1)	78.3%	0.4	91.4%	0.3

6.4.2 DiD results

Table 13 shows the empirical results of DiD estimation, column (1), column (2), column (3) and column (4) show the results of regressions on bills, OOP, LoS and 30-day readmissions.

Table 13. The results of DiD estimation

VARIABLES	(1) Bills	(2) OOP	(3) LoS	(4) 30-day readmissions
PPS*post	-0.0993***	-0.100***	0.935***	0.0243***
	(0.0171)	(0.0188)	(0.0133)	(0.00814)
PPS	-0.0905***	-0.0945***	0.892***	-0.0109
	(0.0162)	(0.0178)	(0.0121)	(0.00771)
post	0.303***	0.242***	0.940***	7.86e-05
	(0.0253)	(0.0278)	(0.0202)	(0.0124)
los	0.0533***	0.0522***		0.000381
	(0.000592)	(0.000652)		(0.000283)
Constant	7.784***	6.807***		0.0130
	(0.0353)	(0.0389)		(0.0169)
Observations	18,438	18,415	18,428	18,071
R-squared	0.467	0.416	0.122[10]	0.026
Number of hospitals	167	167	157	166

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The first row shows the PPS policy effect on treatment group: for PPS targeted patients, bills decreased by 10.0%[11], OOPs decreased by 10.0%, they had the rate for LoS 0.935 times lower for LoS, and the probability of being readmitted in 30 days increased by 2.4%, compared to other patients after PPS adoption, which were consistent with the predictions from theoretical framework.

Moreover, as we stated in the section of empirical strategy, $\delta PPS_{ijt} + \beta_1$ represents the underlying difference in bills between PPS targeted patients and other patients, and β_1 reflects the difference between control group and treatment group before the policy. Thus, compared to other patients, PPS targeted patients had 9.1% lower bills, 9.5%

[10]The fixed effect of Poisson regression does not output the intercept; thus, it is important to calculate the R-square. In the pattern of Least Square Dummy Variable (LSDV) estimation, the Poisson model can output a pseudo-R-square. Hence, we report the pseudo-R-square that is calculated in LSDV estimation to reflect the goodness to fit of the Poisson regression.

[11] We have taken the logarithm on bills; the coefficient represents the percentile number of change.

lower OOP, 0.892 times lower rate for LoS and same probability of being readmitted before PPS. After PPS adoption, the bills increased by 30.3%, the OOP increased by 24.2%, the rate for LoS was 0.940 times lower and 30-day readmissions had no significant change. There was a positive relationship between LoS and bills and a positive relationship between LoS and OOP, which aligned with the prediction in theoretical framework. And according to column (4), there was no significant relationship between LoS and 30-day readmissions.

It was noteworthy that the number of hospitals varied in the regression on the bills, LoS and 30-day readmissions. The number was 167 in the regression on the bills and OOP in column (1) and column (2) while it reduced to 166 in the regression on 30-day readmissions in column (4). This was because there was only 1 observation in a hospital (ID is 160) in the sample, and the information of 30-day readmission was missing for this patient, and the hospital is subsequently omitted in the regression analysis of 30-day readmissions. For the regression on LoS in column (3), since it was Poisson regression, 10 hospitals were dropped in the regression since there is only 1 observation per group.

6.4.3 The results of robustness checks

DiD+PSM

Table 14. The results of DiD+PSM estimation

VARIABLES	(1) Bills	(2) OOP	(3) LoS	(4) 30-day readmissions
PPS*post	-0.101***	-0.0985***	0.892***	0.0266**
	(0.0220)	(0.0245)	(0.0251)	(0.0106)
PPS	-0.0416**	-0.0396*	0.954*	-0.0164
	(0.0208)	(0.0232)	(0.0255)	(0.01000)
post	0.293***	0.231***	1.040	0.00765
	(0.0334)	(0.0372)	(0.0460)	(0.0165)
LoS	0.0529***	0.0517***		0.000253
	(0.000781)	(0.000872)		(0.000377)
Constant	7.634***	6.655***		0.0305
	(0.0461)	(0.0514)		(0.0222)
Observations	11,220	11,201	5,738	11,018
R-squared	0.483	0.423	0.113	0.023
Number of hospitals	162	162	146	162

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 14 shows the results of DiD+PSM estimation[12]. The effects of *PPS*after* on the bills, OOP, LoS and 30-day readmissions were still significant after matching process, and the coefficients of *PPS* and *post* in DiD+PSM estimations were nearly consistent with the results of DiD estimation, these results enhanced the robustness of DiD estimation.

Anticipation effect test

The results of anticipation effect test were in Table 15. After the first quarter of 2011, bills and OOPs for PPS targeted patients decreased, 30-day readmissions for PPS targeted patients increased, while LoS for PPS targeted patients had no significant change. Anticipation effects happened to bills and OOP as predicted, happened to 30-ady readmissions even though we did not observe any anticipation effect for the trend of 30-day readmissions in Figure 17, but did not happen to LoS which was not aligned with the prediction. Thus, we could explain that when the hospital anticipated PPS adoption, they decreased the patients' health payments through the channel other than decreasing LoS (e.g. decreasing drug payments).

Table 15. DiD estimation after taking out anticipation time

VARIABLES	(1) Bills	(2) OOP	(3) LoS	(4) 30-day readmissions
PPS*post	-0.0627***	-0.0648***	0.988	0.0205**
	(0.0200)	(0.0221)	(0.0168)	(0.00954)
PPS	-0.117***	-0.120***	0.854***	-0.00888
	(0.0192)	(0.0212)	(0.0141)	(0.00915)
post	0.286***	0.226***	0.918***	0.00185
	(0.0257)	(0.0284)	(0.0200)	(0.0126)
LoS	0.0533***	0.0523***		0.000368
	(0.000592)	(0.000652)		(0.000283)
Constant	7.797***	6.821***		0.0139
	(0.0356)	(0.0397)		(0.0171)
Observations	18,438	18,415	18,428	18,071
R-squared	0.467	0.416	0.122	0.026
Number of hospitals	167	167	157	166

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[12] In the PSM estimation of column (1) column (2) and column (4), the matching of the sample included LoS as one of the covariates; while in the PSM of column (3), the matching of the sample did not obtain LoS in the covariates, since LoS was the explained variable in this estimation.

6.4.4 Heterogeneity analysis

Since the PPS tariffs were not identical for tier 2 and tier 3 hospitals, we then investigated whether the PPS reform had the same effect on hospitals of different tiers.

Table 16 and 17 showed the effect of PPS on tier 3 hospitals and the effect of PPS on tier 2 hospitals. According to these results, under PPS reform, both tier 2 and tier 3 hospitals had reduced bills, OOPs and LoS. However, the impact of PPS on readmissions was not significant in tier 3 hospitals, but the probability of readmission increased in tier 2 hospitals.

The different health service qualities of tier 2 and tier 3 hospitals under PPS could be explained by the different resource allocation of these two types of hospital. In China, since tier 3 hospitals are all general hospitals at city level with the bed capacity over 500, while tier 2 hospitals are general or specialized hospitals at district/county level with the bed capacity less than 500, the resources of tier 3 hospitals are more abundant than tier 2 hospitals. Compared to tier 2 hospitals, the diploma and professional rank of the medical staff are higher on average in tier 3 hospitals. And ratio of doctor to nurse is also higher in tier 3 hospitals, which means there are more nurses allocated to a doctor for taking care of the doctor’s patient the discharged patients in tier 3 hospitals can receive more adequate care (Zheng et al., 2013, Song and Chen, 2018). Thus, even if tier 2 and tier 3 hospitals both faced with the budget constraint, the higher quality of human resource of tier 3 hospitals could maintain their health services quality.

Table 16. DiD estimation for tier 3 hospitals

VARIABLES	(1) Bills	(2) OOP	(3) LoS	(4) 30-day readmissions
PPS*post	-0.103***	-0.124***	0.922***	0.0278
	(0.0388)	(0.0415)	(0.0258)	(0.0175)
PPS	0.226***	0.258***	0.879***	0.00847
	(0.0693)	(0.0742)	(0.0241)	(0.0316)
post	-0.110***	-0.122***	0.862***	-0.0216
	(0.0379)	(0.0406)	(0.0446)	(0.0171)
LoS	0.0445***	0.0462***		0.000810**
	(0.000904)	(0.000968)		(0.000409)
Constant	8.142***	7.389***		0.0130
	(0.0812)	(0.0870)		(0.0367)
Observations	7,177	7,166	7,175	7,023
R-squared	0.418	0.384	0.133	0.036
Number of hospitals	42	42	40	42

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 17. DiD estimation for tier 2 hospitals

VARIABLES	(1) Bills	(2) OOP	(3) LoS	(4) 30-day readmissions
PPS*post	-0.0962***	-0.0750***	0.937***	0.0287***
	(0.0185)	(0.0208)	(0.0158)	(0.00938)
PPS	0.290***	0.178***	0.893***	-0.00938
	(0.0273)	(0.0306)	(0.0142)	(0.0144)
post	-0.0854***	-0.0937***	0.979	-0.00882
	(0.0172)	(0.0193)	(0.0247)	(0.00870)
LoS	0.0610***	0.0567***		-5.76e-05
	(0.000783)	(0.000878)		(0.000397)
constant	7.659***	6.760***		0.00784
	(0.0403)	(0.0452)		(0.0204)
Observations	11,261	11,249	11,252	11,048
R-squared	0.519	0.449	0.118	0.023
Number of hospitals	146	146	137	145

This table only report the results for key explanatory variables
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

6.5 Discussion

To evaluate PPS adoption, our study investigates the change in bills, patients' OOP health payments, LoS and health outcomes following the replacement of FFS with PPS for selected treatments in Chengdu.

In our main analyses, we employed DiD estimation to investigate the change in bills, OOP, LoS and 30-day readmissions, and performed a robustness check that combined the DiD estimation with PSM to assure that treatment group and control group have a similar trend before the PPS policy. The main results of DiD estimation were generally consistent with the results of DiD+PSM estimation, and both results were aligned with the predictions of the theoretical framework. According to the DiD (DiD+PSM) estimation, bills decreased by 10.0% (10.1%), OOP decreased by 10.0% (9.9%), the rate for LoS was 0.935 (0.892) times lower, while the patients' probability of being readmitted in 30 days increased by 2.4% (2.7%). Moreover, our study found the heterogenous performance of tier 2 and tier 3 hospitals under PPS. Both tier 3 and tier 2 hospitals decreased bills, OOP and LoS for patients, while tier 3 hospitals maintained the health outcomes but in tier 2 hospitals readmissions increased by 2.8% following the payment reform.

Figure 18, Figure 19, and Figure 20 are the forest plots of PPS studies reproduced from chapter 3, by adding the results of this study labelled as Zhao (2023). In the forest plots, our finds that PPS reform can reduce bills, OOPs, and LoS are aligned with most of the other studies. And for the PPS effect on readmissions, the results of previous studies are quite controversial, while our study finds an increase in 30-ady readmissions.

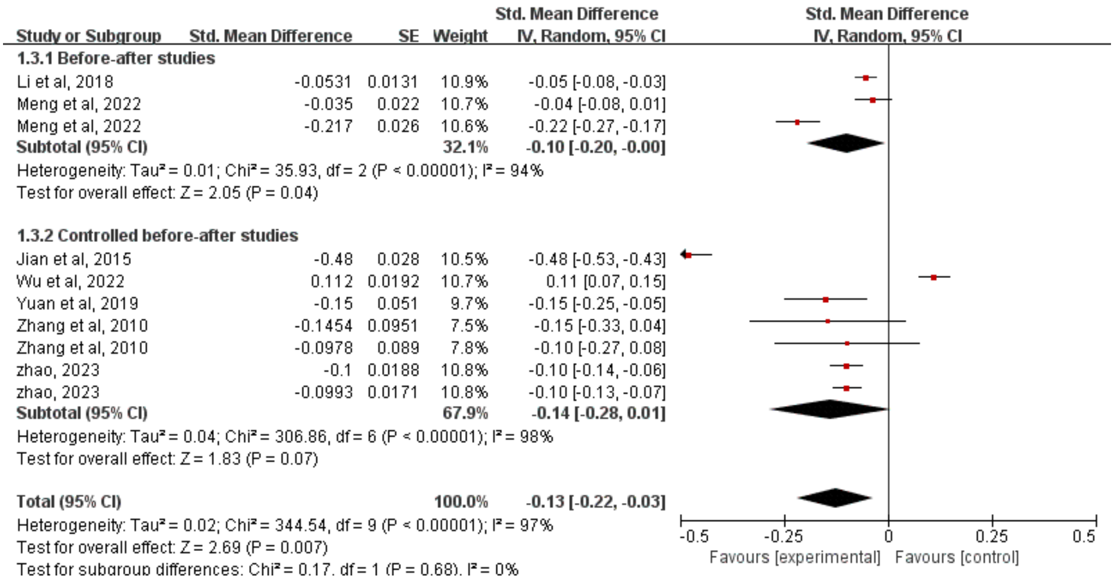


Figure 18. Forest plot of PPS effect on payments including results from this study

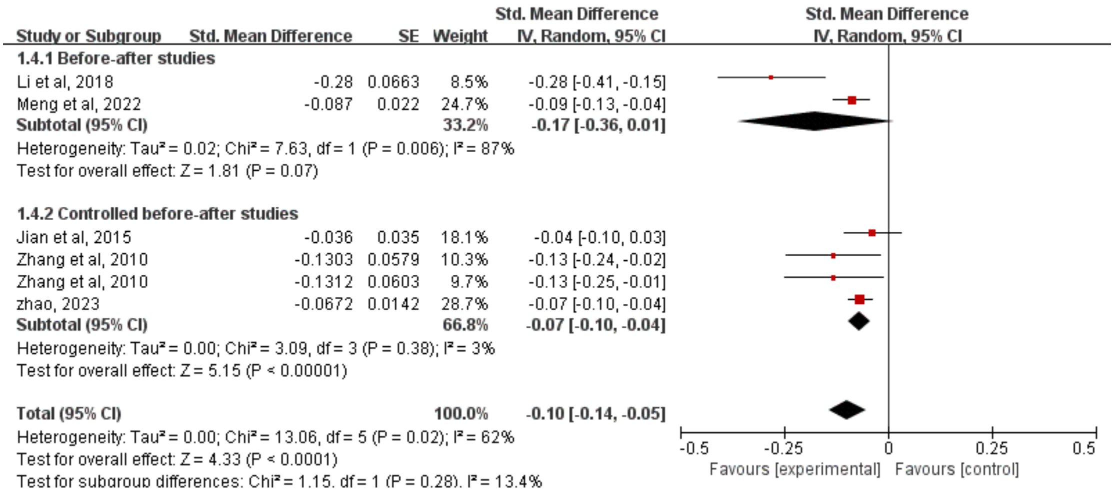


Figure 19. Forest plot of PPS effect on LoS including results of this study

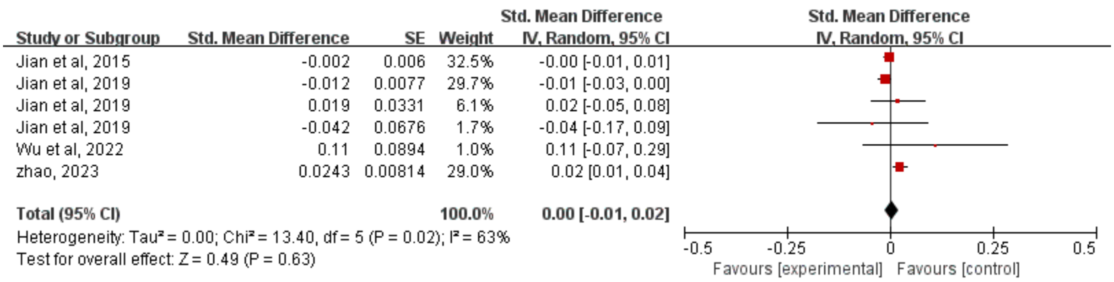


Figure 20. Forest plot of PPS effect on readmissions including results of this study

It is noteworthy that the studies of Jian et al. (2015a) and Zhang (2010) use a similar empirical strategy to evaluate the effects of PPS on patients' health payments and health outcomes in Shanghai and Beijing in China.

The study of Jian et al. (2015a) focuses on the PPS pilot policy which was implemented among six tier 3 hospital in Beijing in 2011, by comparing the health payments, OOP, LoS and readmissions of the patients from the six targeted tier 3 hospitals with the patients from another eight tier 3 hospitals. They find that the total health payments decrease by 6.5%, OOP decreases by 10.5% after PPS adoption, while they do not observe any reduction in LoS or increase in readmissions.

The study of Zhang (2010) finds that PPS did not reduce patients' health payment significantly. However, since PPS was implemented among health insurance enrollees in Shanghai, this study drew comparisons between the insured and uninsured patients, which may lead to biased estimates due to the heterogeneity of the insured and uninsured (e.g., the difference in socio-economic status between these two groups).

Compared to the previous studies on PPS in China which employ DiD estimations, our study has three main strengths. First, we develop the theoretical framework, illustrating the goal of the payer, the patient and the hospital and relationship between them, describing the changes of these three agents under PPS reform, and making predictions about bill, OOP, LoS and 30-day readmissions. Second, we combine the DiD estimation with PSM, assuring the treatment group and control group have a similar trend before PPS policy, with results being consistent. Third, we investigate the policy effect on tier 3 and tier 2 hospitals separately and observe heterogenous effects.

However, our study has several limitations. First, though LoS can be the indicator of treatment intensity (Charunwatthana and Supakankunti, 2014), there are still other studies claiming that LoS can be the indicator of hospital performance (Boes and Napierala, 2021). Hence, LoS is not a perfect measurement for treatment intensity since its ambiguity in either relating to treatment intensity or health outcomes. However, due to the limitation of data, we have no other information about treatment. Second, since our data only contains the patients who are the enrolled in UEBMI and URBMI and discharged from tier 2 and tier 3 hospitals, which does not allow our study to select the patients with PPS targeted diagnoses but unaffected by PPS policy (e.g., the patients with PPS targeted diagnoses but enrolled in NRCMS) as control group. This restricts the pool of patients from which to draw our control group. This might undermine the comparability of our treatment and control groups, though we applied PSM to reduce heterogeneity. Third, since the data do not contain the patients with PPS targeted diagnoses but unaffected by PPS policy, we cannot investigate whether the hospitals shifted cost from the patients affected by the policy to the patients with same targeted diagnoses but unaffected by the policy.

6.6 Conclusion

In this study, we find that PPS adoption in Chengdu led to decreases in bills, OOP and LoS decreased but quality, as measured by 30-day readmission, worsened. We find that hospitals in Chengdu reduced the bills and OOPs one quarter before the official document of PPS was published, possibly reacting in advance before the reform was officially initiated. According to heterogeneity analysis, the increase in re-admissions was observed in tier 2 but not in tier 3 hospitals. Thus, in Chengdu, following the introduction of PPS tier 3 hospitals reduced health payments and LoS for their patients, maintaining quality; while tier 2 hospitals reduced health payments and LoS but there were adverse health outcomes for their patients.

7. Effects of PPS on patient payments and health outcomes among the frail elderly in Chengdu in 2018

Summary

Background: In December 2017, Chengdu implemented the second phase of the PPS reform, targeting enrollees of the URBMI, UEBMI, and NRCMS across 44 diseases in tier 3 grade 1 hospitals.

Objectives: This study assessed the impact of the PPS policy on patient bills and LoS for frail elderly patients.

Data: Data on frail elderly patients was sourced from the Chengdu Healthcare Security Administration, spanning January 2014 to December 2019, encompassing 503,643 observations across 1055 hospitals.

Methods: We applied a triple difference (DDD) estimation to analyze the PPS policy variations by disease and hospital. Initially, we performed the first difference-in-differences (DiD) analysis among patients from targeted hospitals by segregating them into groups with targeted and untargeted diseases. Subsequently, a second DiD analysis was conducted among patients from untargeted hospitals, using the same grouping strategy. The final step involved calculating the differences between the outcomes of the two DiD analyses.

Results: The introduction of PPS led to a 15.7% reduction in bills and an 8.9% increase in LoS for patients with targeted diseases in targeted hospitals, relative to the counterfactual. In the sub-group analysis, for cataract patients, both bills and LoS remained unchanged post-reform. For patients with herniated intervertebral disc (HIVD), there was a 19.8% reduction in bills, while LoS remained stable. The lack of significant change in the bills for cataract patients may be attributable to a spillover effect.

Conclusion: The 2018 PPS reform in Chengdu resulted in reduced bills and increased LoS among targeted frail elderly patients, the results of bills were consistent with predictions while the results for LoS were not. However, these findings are likely biased due to the small sample size and the violation of the parallel trend assumption in the DDD analysis, suggesting that the results may not accurately reflect the true impact of the PPS reform.

7.1. Introduction

After the initial PPS reform in May 2011, Chengdu launched a second phase starting on December 28, 2017. Both waves were structured around single-disease PPS, but differed significantly in scope and target. Firstly, the initial wave focused on only 10 diseases, whereas the second expanded to include **44**. Secondly, the first wave applied to both UEBMI and URBMI across all tier 2 and tier 3 hospitals in Chengdu, while the second was limited to UEBMI and URBMI patients in grade 1 tier 3 hospitals. As of 2019, there were 69 tier 3 hospitals in Chengdu, comprising 40 grade 1 and 29 grade 2 facilities.

In contrast to the initial wave, the second reform targeted specific diseases and diagnostic groups. The 2011 reform aimed to contain costs and reduce unnecessary treatments for the 10 targeted diseases. With a shift to DRGs, the second wave sought not only to control costs but also to enhance hospital service quality. Under PPS, the pricing for each diagnostic group was standardized across similar tier hospitals, enabling the government to directly evaluate and compare the quality of health services provided. Consequently, DRGs served as a mechanism to both monitor and incentivize hospitals to enhance their service quality.

However, the ambitious goals of the second PPS wave demanded advanced capabilities from hospitals. The complexity of diagnostic groups required experienced coding staff to ensure accurate patient classification; additionally, hospitals needed to maintain superior levels of staffing and equipment to uphold or improve service quality under financial constraints. Thus, the second wave targeted only grade 1 tier 3 hospitals, which possessed the largest scale and most advanced human resources compared to lower-tier facilities. Similarly, this phase focused on diseases with a high the number of patient.

The second wave of PPS reform targeted a broader range of diseases and affected a larger patient group. Since the 2018 PPS reform aimed to contain costs and improve hospital service quality, this study investigates changes in both bills and service quality among frail elderly individuals to assess the effects of PPS. By utilizing data on frail elderly people in Chengdu, we plan to examine the impact of PPS on bills and LoS for this demographic group.

Unfortunately, the data available for analysis have several limitations. First, the policy document distinguishes between 'cataract' and 'senile cataract' as separate conditions, yet the data often recorded both under the generic term 'cataract.' Consequently, we can only identify patients under the broad category of 'cataract.' In subsequent sections, all patients whose condition is recorded as 'cataract' are considered targeted patients.

Second, the dataset comprises only elderly individuals and does not encompass all 44 diseases, with the majority of targeted patients suffering from cataracts and herniated intervertebral disc (HIVD), common conditions in this population. This limits the scope of the analysis due to the small sample size and specific patient population. In the empirical sections, we utilize a robust methodological approach to maximize the use of the available data. Lastly, the absence of detailed patient surgical information means we can only identify those with targeted diseases, rather than individuals from every targeted diagnostic group.

The subsequent sections of this chapter will outline the diseases targeted by the 2018 PPS reform, present the empirical strategy, describe the data, and discuss the empirical results.

7.2 Targeted diseases

Chengdu launched the second wave of PPS reform on December 28, 2017, but for analytical purposes, the start date is considered to be January 1, 2018. The 2018 PPS reform targeted 44 diseases and 101 diagnostic groups (detailed in Appendix 9.3.3). The reform established distinct clinical pathways for each disease based on the type of operation, anesthesia method, and treatment quantity, setting specific PPS tariffs for each operation.

In our dataset, the majority of PPS-targeted patients were diagnosed with cataracts and HIVD. Table 18 illustrates the PPS tariffs for each diagnostic group under these conditions.

Table 18. Several targeted diagnostic groups and their PPS tariffs

Disease	Operation	Anesthesia method	Quantity	PPS tariff
Cataracts	Cataract extracapsular extraction + intraocular lens (IOL) implantation	Local anesthesia	Single Eye	4710
Cataracts	Cataract extracapsular extraction + IOL implantation	Local anesthesia	Both eyes	7610
Cataracts	Cataract ultrasonic phacoemulsification extraction + IOL implantation	Local anesthesia	Single Eye	7380
Cataracts	Cataract ultrasonic phacoemulsification extraction + IOL implantation	Local anesthesia	Both eyes	10120
herniated intervertebral disc (HIVD)	Percutaneous Laser Lumbar Disc Removal	General Anesthesia	1 intervertebral disc	12520

For instance, for cataracts, there are four sub-groups, each with a unique PPS tariff. A patient undergoing extracapsular cataract extraction and IOL implantation with local anesthesia for one eye is charged CNY 4,710, while the procedure for both eyes costs CNY7610. Similarly, for cataract ultrasound emulsification and IOL implantation with local anesthesia for one eye, the tariff is CNY 7380, and for both eyes, CNY10120. For HIVD, there is a single sub-group where the procedure of percutaneous laser lumbar disc removal with general anesthesia has a PPS tariff of CNY 12,520. Thus, the 101 PPS tariffs derive from the 44 targeted diseases.

Based on the theoretical framework outlined in Chapter 2.3, we anticipate that bills and LoS for targeted patients will decrease post-policy, as hospitals may reduce treatment intensity to lower bills.

7.3 Empirical strategy

We examine the effects of PPS on two outcomes: bills (b_{ijt}) and LoS (LoS_{ijt}), where i denotes the individual patient, j represents the hospital and t represents the time. According to our theoretical framework, we anticipate reductions in both bills and LoS following the implementation of PPS. We also adjust for inflation in our analysis of bills.

The 2018 PPS reform was targeted at several tier 3 hospitals and encompassed 44 diseases, allowing for variation in the reform either by disease or by hospital.

However, using either disease or hospital variation to identify the policy's effects presents challenges in our study. Since the 2018 reform was limited to certain tier 3 hospitals, while all tier 2 hospitals were excluded, this creates heterogeneity between the targeted and unaffected hospitals. Selecting patients with PPS-targeted diseases from tier 3 hospitals as the treatment group and those from unaffected hospitals as the control group compromises comparability, violating the assumption of a pre-policy common trend. Conversely, selecting patients with PPS diseases from tier 3 hospitals as the treatment group and patients with similar diseases from unaffected hospitals as the control group yields comparability but results in a limited number of observations.

To address the heterogeneity between targeted and untargeted hospitals and to increase the sample size, we employ a Triple Differences (DDD) estimation to assess the effects of the 2018 PPS policy on bills and LoS for the frail elderly.

The DDD estimation operates as follows: 1) conduct the first Difference-in-Differences (DiD) estimation among patients with PPS-targeted diseases from targeted hospitals and patients with similar diseases from the same hospitals; 2) conduct the second DiD among patients with PPS diseases from untargeted hospitals and similar patients from these hospitals; 3) calculate the difference between the results of the first and second DiD estimations to derive the DDD effects.

To analyze both the overall and specific effects for cataract and HIVD patients, we apply the DDD estimation comprehensively and separately for each disease group. The model for estimating bills is structured in the following general form:

$$b_{ijt} = \alpha + \beta_0 disease * hospital * post_{ijt} + \beta_1 disease * post_{ijt} + \beta_2 hosital * post_{ijt} + \beta_3 disease * hospital_{ijt} + \beta_4 disease_{ijt} + \beta_5 hospital_{ijt} + \beta_6 post_{ij} + \delta los_{ijt} + \gamma X_{ijt} + \theta_j + \varphi_t + \varepsilon_{ijt} \quad (28)$$

In Eq. (28), the outcome b_{ijt} represents the total health service charges on the bill for patient I discharged from hospital j at time t, and this equation is estimated using OLS.

$disease * hospital * post_{ijt}$ is the interaction term of *disease*, *hospital* and *post*, with the β_0 coefficient representing the DDD estimator. If patient I from hospital j is with PPS disease, from the PPS targeted hospital and after implementing the policy, $disease * hospital * post_{ijt}$ equals 1, otherwise, it equals 0.Thus, if β_0 is positive and significant, it represents that PPS adoption increases bills for patients with targeted diseases from targeted hospitals (intervention group) compared to all other patients (i.e., those with untargeted diseases from targeted hospitals, those with targeted diseases from untargeted hospitals, and those with untargeted diseases from untargeted hospitals) post-policy.

$disease * post_{ijt}$is the interaction between *disease* and *post*, representing the policy effect on patients with targeted diseases. $hosital * post_{ijt}$ is the interaction between *hospital* and *post*, reflecting the policy effect on patients from targeted hospitals. $disease * hospital_{ijt}$ is the interaction between *disease* and *hospital*, illustrating the differential between patients with targeted diseases from targeted hospitals (intervention group) and all other groups (i.e., patients with untargeted diseases from targeted hospitals, patients with targeted diseases from untargeted hospitals, and patients with untargeted diseases from untargeted hospitals).

$disease_{ijt}$ is a dummy variable set to 1 if the patient is diagnosed with a PPS-targeted disease and 0 otherwise. $hospital_{ijt}$ is a dummy variable set to 1 if the patient is from a PPS-targeted hospital and 0 otherwise. $post_{ijt}$ is a dummy variable used to indicate the timing of the policy effect, set to 0 pre-policy and 1 post-policy.

Thus, $\beta_3 + \beta_4 + \beta_5$ indicates the difference between patients with targeted diseases and from targeted hospitals and all the other patients before PPS. Among the patients from targeted hospitals, $\beta_3 + \beta_4$ represents the difference in being with targeted diseases or not before PPS. Among the patients from untargeted hospitals, β_4 represents the difference in being with targeted diseases or not before PPS. Among the patients with targeted diseases, $\beta_3 + \beta_5$ represents the difference in being from targeted hospitals or not before PPS. Among the patients with untargeted diseases, β_5 represents the difference in being from targeted hospitals or not before PPS.

According to the theoretical framework, bills are associated with the LoS, in Eq (1), LoS is included as a control variable.

Controlling for patient-level characteristics, we include X_{ijt} as a vector of patient attributes. This vector comprises data on patient i's gender, age, type of health insurance, and degree of frailty. There may be unobservable factors, aside from patient characteristics, that influence bills. To mitigate the impact of time-invariant unobservable factors, such as the type and location of the hospital, we utilize a fixed effect model. θ_j controls for the fixed effects at hospital level; φ_t controls for time dummies (i.e., half year).

To assess the policy's impact on LoS, the following model is estimated:

$$E(\,LoS_{ijt}) = \exp\,(\alpha + \beta_0 disease * hospital * post_{ijt} + \beta_1 disease * post_{ijt} + \beta_2 hosital * post_{ijt} + \beta_3 disease * hospital_{ijt} + \beta_4 disease_{ijt} + \beta_5 hospital_{ijt} + \beta_6 post_{ij} + \gamma X_{ijt} + \theta_j + \varphi_t + \varepsilon_{ijt}) \quad (29)$$

In Eq. (29), the independent variable is LoS for all patients, specifically for those with cataracts and HIVD. Given the assumed Poisson distribution of the dependent variable LoS, Eq. (29) is estimated using Poisson regression which is a non-linear model. As the previous chapter, we report IRR for LoS estimation.

We classify patients with PPS-targeted diseases as those in group A and those with similar but untargeted diseases as in group B. Similarly, patients from targeted hospitals are categorized as in group C, and those from untargeted hospitals in group D.

The expected outcomes for patients **with targeted diseases from targeted hospitals post-policy** (time=1) and pre-policy (time=0) are represented in Eq. (30):

$$E(\overline{y_{AC1}}) = \alpha + \beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 \quad (30)$$

We have identified the expected outcomes for various patient groups as follows: patients with targeted diseases from targeted hospitals before the policy (AC_0), patients with untargeted diseases from targeted hospitals after the policy (BC_1), patients with untargeted diseases from targeted hospitals before the policy (BC_0), patients with targeted diseases from untargeted hospitals after the policy (AD_1), patients with targeted diseases from untargeted hospitals before the policy (AD_0), patients with untargeted diseases from untargeted hospitals after the policy (BD_1), and patients with untargeted diseases from untargeted hospitals before the policy (BD_0):

$$E(\overline{y_{AC0}}) = \alpha + \beta_3 + \beta_4 + \beta_5 \quad (31)$$

$$E(\overline{y_{BC1}}) = \alpha + \beta_2 + \beta_5 + \beta_6 \quad (32)$$

$$E(\overline{y_{BC0}}) = \alpha + \beta_5 \quad (33)$$

$$E(\overline{y_{AD1}}) = \alpha + \beta_1 + \beta_4 + \beta_6 \quad (34)$$

$$E(\overline{y_{AD0}}) = \alpha + \beta_4 \quad (35)$$

$$E(\overline{y_{BD1}}) = \alpha + \beta_6 \quad (36)$$

$$E(\overline{y_{BD0}}) = \alpha \quad (37)$$

Thus, the difference in expected outcome for patients with targeted diseases from targeted hospitals before and after the policy should be $E(\overline{y_{AC1}}) - E(\overline{y_{AC0}})$ and equal $(\beta_0 + \beta_1 + \beta_2 + \beta_6)$, which includes the policy effect on intervention group (β_0), the policy effect on patients with targeted diseases (β_1), the policy effect on patients from targeted hospitals (β_2) and the time effect of the policy (β_6).

The difference in expected outcomes for patients with untargeted diseases from targeted hospitals before and after the policy should be $E(\overline{y_{BC1}}) - E(\overline{y_{BC0}})$ and equal $(\beta_2 + \beta_6)$, which includes the policy effect on patients from targeted hospitals and the time effect of the policy. Then the estimated effects of the first DiD should be$[E(\overline{y_{AC1}}) - E(\overline{y_{AC0}})] - [E(\overline{y_{BC1}}) - E(\overline{y_{BC0}})] = \beta_0 + \beta_1$, which include the policy effect on intervention group and the policy effect on patients with targeted diseases.

The difference in expected outcomes for patients with targeted diseases from untargeted hospitals before and after the policy should be $(\overline{y_{AD1}}) - E(\overline{y_{AD0}})$ and equal $(\beta_1 + \beta_6)$, which includes the policy effect on patients with targeted diseases and the time effect of the policy.

The difference in expected outcomes for patients with untargeted diseases from untargeted hospitals before and after the policy should be $E(\overline{y_{BD1}}) - E(\overline{y_{BD0}})$ and equal β_6, representing solely the time effect of the policy. Thus, the estimated effects of the second DiD should be $[E(\overline{y_{AD1}}) - E(\overline{y_{AD0}})] - [E(\overline{y_{BD1}}) - E(\overline{y_{BD0}})] = \beta_1$, which includes the policy effect on patients with targeted diseases.

By taking difference between the first DiD and the second DiD, we estimate the DDD effect, which should be β_0. And in the process of estimating DD effects, the difference between targeted hospitals and untargeted hospitals (β_5) has been eliminated.

7.4 Data

Our study dataset comprises hospitalization records of frail elderly individuals enrolled in Chengdu's long-term care insurance, with 563043 observations from 1055 hospitals spanning January 2014 to December 2019. Given that this insurance is subsidized by UEBMI and URBMI in Chengdu, all patients in the dataset are concurrently enrolled in both UEBMI and URBMI.

The dataset includes records of discharged patients; each record details the patient's gender, age, insurance ID, degree of frailty (i.e., mild, moderate, or severe), dates of admission and discharge, diagnoses, hospital identifier, bill, detailed fees (e.g., drug and hospitalization fees), reimbursement, and OOP.

From January 2014 to December 2019, the dataset contains 503,643 observations from 1,055 hospitals. Although the PPS reform targeted 44 diseases, we identified only two—cataract and HIVD—with sufficient observations (over 1,000 each) among the frail elderly. For cataract analysis, we selected patients with lens, vitreous, cornea, and retina diseases as controls. For HIVD, we selected patients with all other non-targeted **vertebral diseases** as controls.

We calculated an average PPS tariff for all diagnostic groups associated with senile cataract, amounting to CNY 5,202 per single-eye treatment. This figure was defined as the average PPS tariff for cataract single-eye treatments. Subsequently, we compared this tariff with the average bill for similar treatments in our dataset before the implementation of the policy.

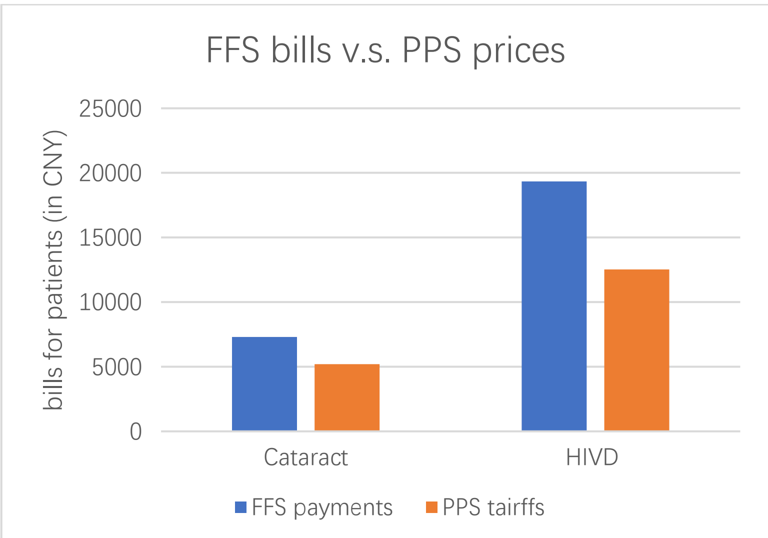


Figure 21. The comparison between FFS bills and PPS tariffs

According to the data, after adjusting for inflation, the average bill for cataract single-eye treatment in tier 3 grade 1 hospitals (i.e., the targeted hospitals in the 2018 PPS) is

CNY 7,312, and for HIVD treatment in the same hospital category, it is CNY 19,346. Figure 21 provides a comparison of FFS bills and PPS tariffs for these two diseases, focusing on the targeted hospitals.

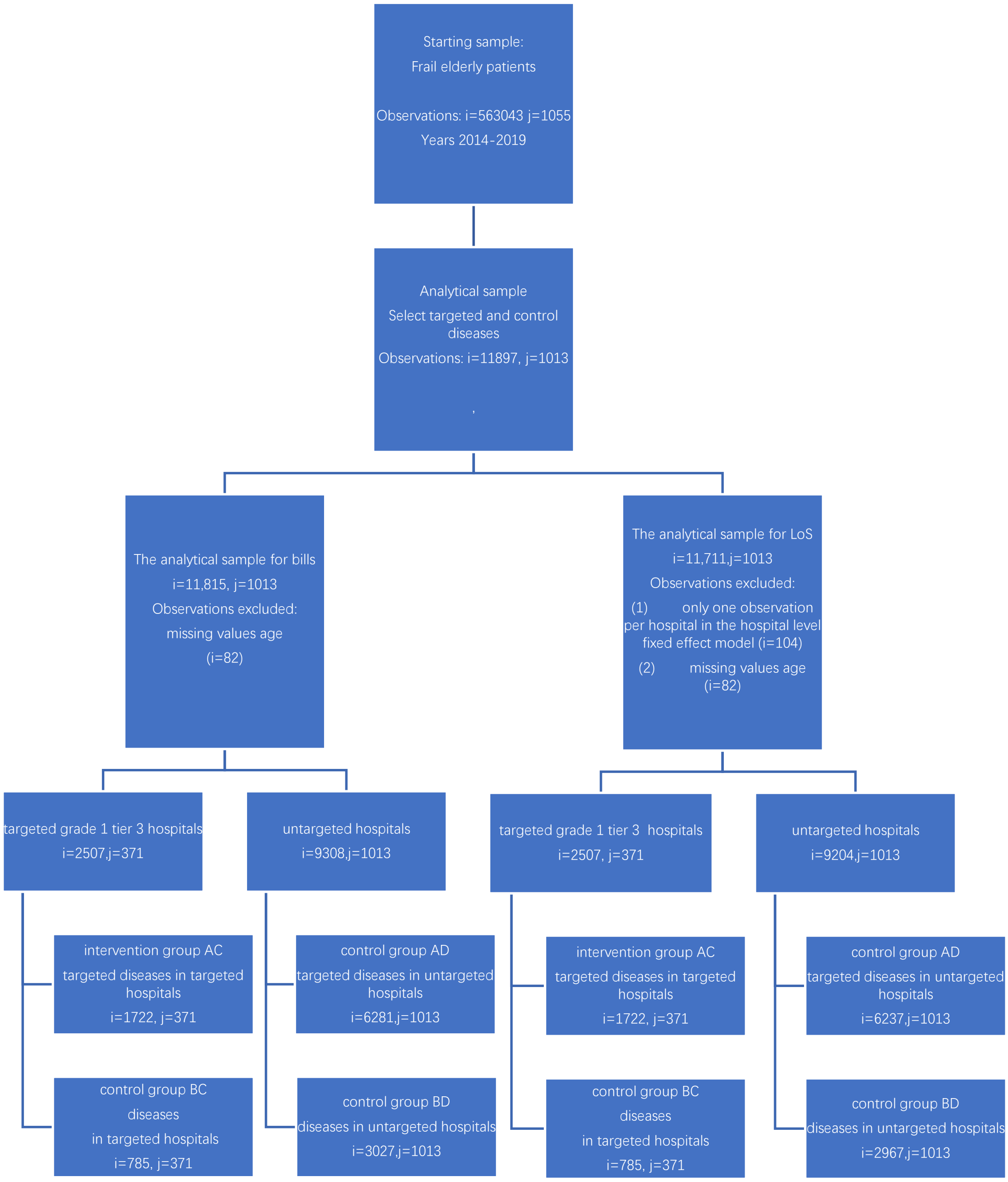
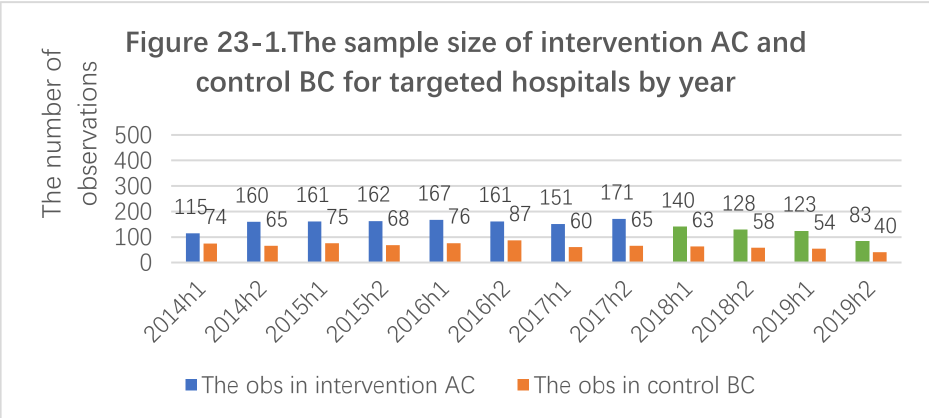


Figure 22. Flow diagram for sample selection

In the DDD estimation, there is one intervention group and three control groups. The intervention group, AC, and control group, BC, both originate from targeted hospitals but differ in that they treat targeted and untargeted diseases, respectively. The control groups, AD and BD, derive from untargeted hospitals and similarly manage targeted and untargeted diseases, respectively. Figure 22 illustrates the selection process of analytical samples for bills and LoS, detailing the number of observations and hospitals in each analytical sample for the intervention and control groups.

The first DiD analysis involves comparing the intervention group AC with the control group BC, which reflects the distinction between targeted and untargeted diseases within targeted hospitals. The second DiD compares the control group AD with control group BD, assessing the difference between targeted and untargeted diseases in untargeted hospitals, where no significant effects are expected. We then calculate the difference between the results of these two DiDs to estimate the DDD effects.

Next, we present the distribution of samples for the overall analysis. Based on the DDD estimation, we focus on the first DiD among patients from targeted hospitals. In our analysis, the sample sizes for the intervention group AC and control group BC are detailed at the top of Figure 23. Figure 23-1 and Figure 23-2 respectively show the sample distribution for targeted hospitals and untargeted hospitals. Patients from **targeted hospitals** diagnosed with targeted diseases and in the post-policy stage (indicated by green bars) represent the group directly affected by the reform. This group, alongside those who received treatment pre-reform, serves as the controls, illustrating their relevance to the empirical strategy employed in the DDD analysis.



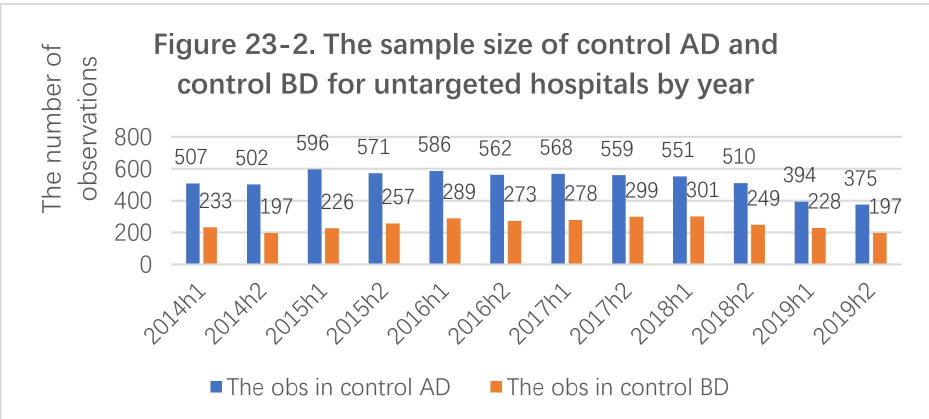
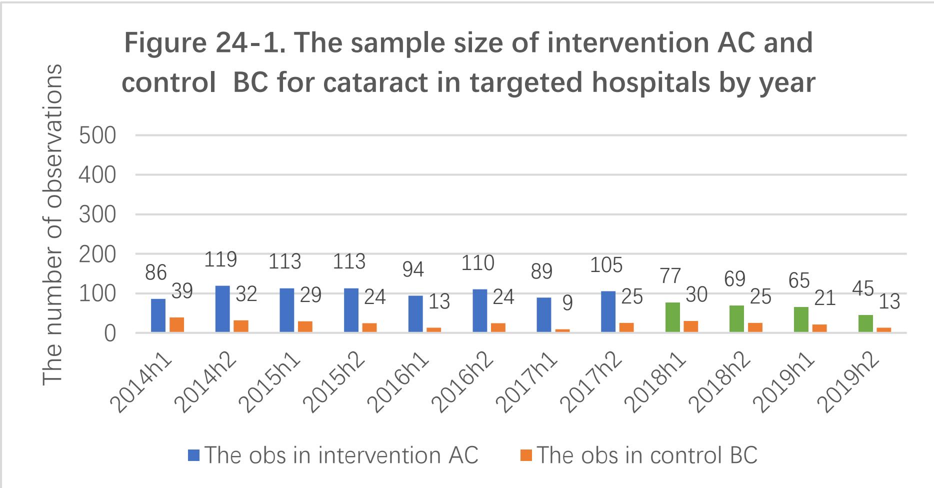


Figure 23. The sample size for overall analysis by half year

Subsequently, we conduct the second DiD analysis among patients from untargeted hospitals. For **these hospitals**, the sample sizes of control groups AD and BD are displayed at the bottom of Figure 23. In this second DiD estimation, we hypothetically assign "treatment" and "control" statuses to these groups under the assumption that the policy is implemented in untargeted hospitals, although it is not actually in effect. Therefore, all observations within control groups AD and BD are expected to be unaffected by the policy and are classified according to disease variation.

We then detail the distribution of samples for cataract and HIVD patients separately.

In the analysis of **cataract** within **targeted hospitals**, the sample sizes for the intervention group AC and control group BC are presented at the top of Figure 24. For **untargeted hospitals**, the sample sizes of control groups AD and BD are indicated at the bottom of Figure 24. Figure 24-1 and Figure 24-2 respectively show the sample distribution for targeted hospitals and untargeted hospitals.



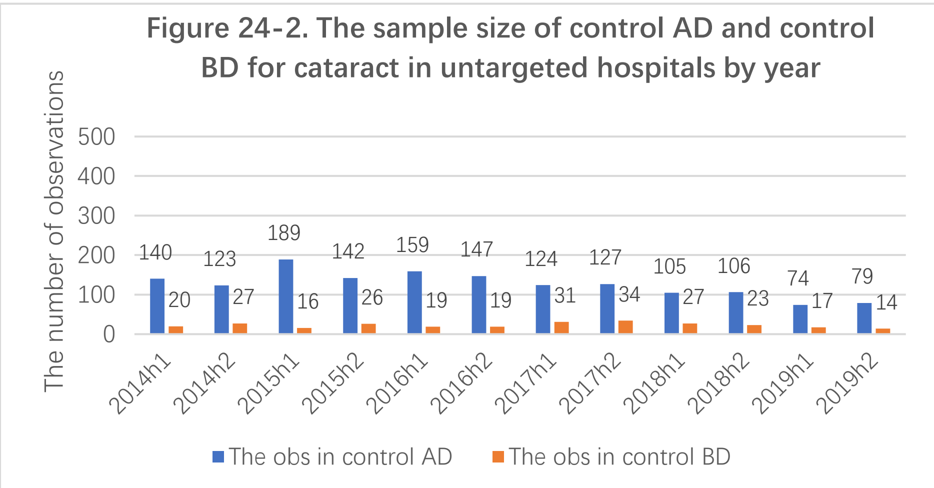
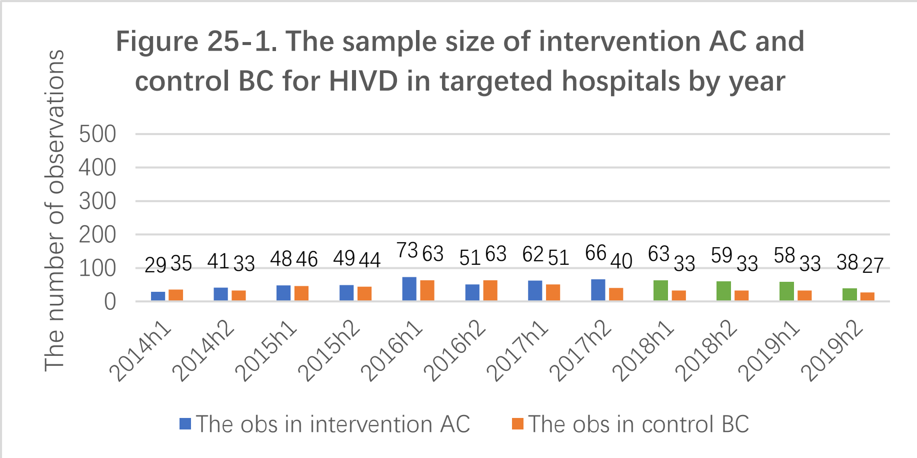


Figure 24. The sample size for cataract analysis by half year

For **HIVD** analysis, the sample sizes for the intervention group AC and control group BC at targeted hospitals are depicted at the top of Figure 25. Conversely, the sample sizes for the control groups AD and BD at **untargeted hospitals** are presented at the bottom of the same figure. Figure 25-1 and Figure 25-2 respectively show the sample distribution for targeted hospitals and untargeted hospitals.



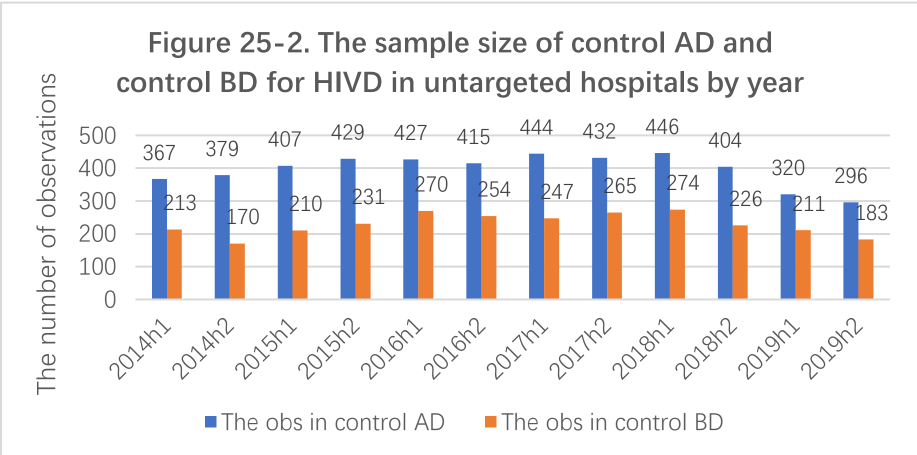


Figure 25. The sample size for HIVD analysis by half year

7.5 Empirical results

7.5.1 Descriptive results

Initially, Figure 26 displays the trends of outcome variables, Figure 26-1 to Figure 26-3 display bills for each group (i.e., groups AC, AD, BC, BD) across the overall, cataract, and HIVD analyses, respectively.

Subsequently, Figure 27 displays the trends of outcome variables, Figure 27-1 to Figure 27-3 display the trends of LoS for each group in the overall, cataract, and HIVD analyses, respectively.

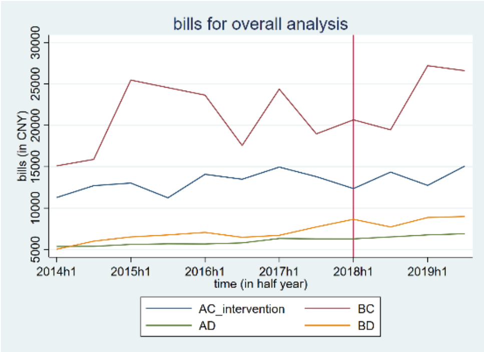


Figure 26-1. The trends of bills for overall sample

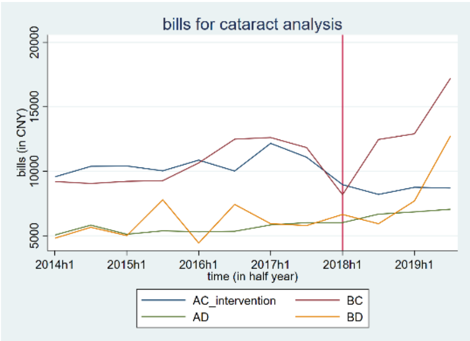


Figure 26-2. The trends of bills for cataract

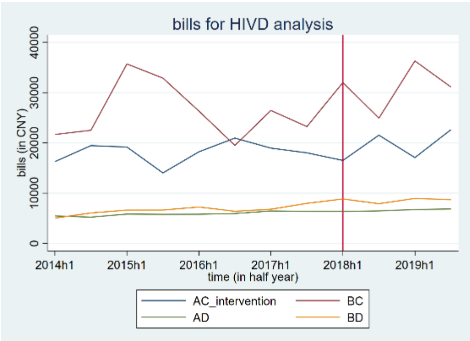


Figure 26-3. The trends of bills for HIVD

Figure 26. The trends of bills for overall analysis and sub-group analysis

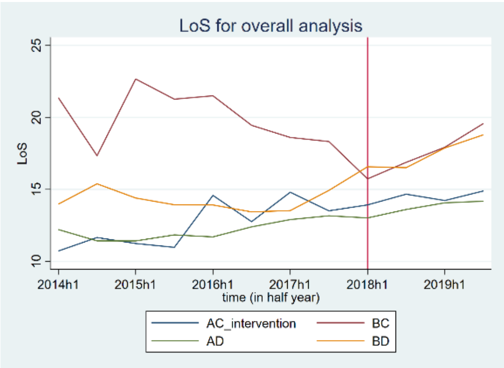


Figure 27-1. The trends of LoS for overall sample

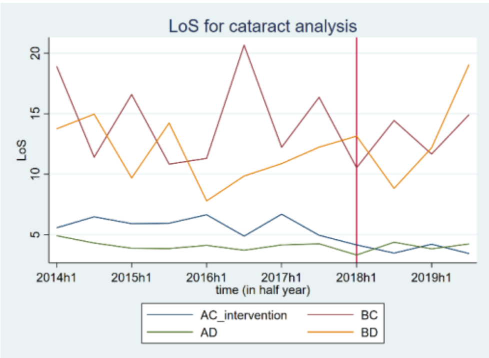


Figure 27-2. The trends of LoS for cataract

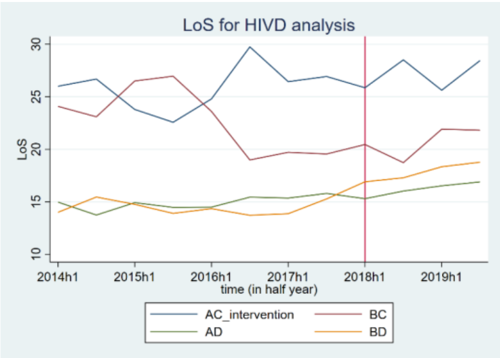


Figure 27-3. The trends of LoS for HIVD

Figure 27. The trend of LoS for overall analysis and sub-group analysis

According to the DiD assumption under the Triple Difference (DDD) framework, the first DiD analysis presumes that both the intervention and control group 1 share a common trend in bills and LoS prior to the policy implementation. In the second DiD analysis, control groups 2 and 3 are expected to maintain a common trend both before and after the policy implementation, as they are not subject to the policy. The figures indicate violations of this common trend assumption across the intervention and control groups, for both bills and LoS, thereby compromising the validity of the DDD approach and potentially resulting in biased estimates. Therefore, we will briefly summarize the estimation results and subsequently critique the reliability of these findings.

7.5.2 DDD results

Initially, we presented the DDD results for the combined analysis of all patients (*i.e.*, cataract and HIVD). Subsequently, we conducted sub-group analyses for cataract and HIVD separately.

Table 19 displays the DDD results for the overall analysis, where column 1 and column 2 detail the results for bills and LoS, respectively. According to the coefficients of the DDD estimator (β_0), bills for patients at targeted hospitals with targeted diseases were reduced by 15.7% after the policy implementation, compared to other patients. Conversely, patients at targeted hospitals with targeted diseases had the rate 1.093 times higher for LoS, relative to patients in the control groups. Therefore, while PPS adoption reduced bills, it also led to adverse health outcomes in patients from targeted hospitals with specified diseases. The unexpected results regarding LoS can be attributed to two factors. First, hospitals may have lowered bills by reducing drug utilization or other services rather than by decreasing LoS. Second, hospitals aimed to diminish treatment intensity, but this reduction led to worse health outcomes, necessitating a longer recovery period for these patients, thereby increasing the LoS for targeted patients following the policy.

In addition to elucidating the results, we must address the significant issue that neither the bills nor the LoS of the intervention group exhibited a parallel trend when compared with any control group during the pre-policy stage in the descriptive analysis. Moreover, due to the limited sample size, we observed that the trend of outcome variables often fluctuated widely, particularly for LoS. Indeed, it was impossible to visually identify the policy effects on bills or LoS in the descriptive results; however, the DDD method estimated notable policy effects for both bills and LoS. These findings, as estimated by DDD, may be attributable to the small sample size and substantial standard deviation of the data, thus rendering them less reliable.

Table 19. DDD results for overall analysis

	(1)	(2)
VARIABLES	bills	LoS
$\beta_0$hospital*disease*post	-0.157***	1.093***
	(0.0525)	(0.0303)
$\beta_1$hospital*post	0.0297	0.973
	(0.0440)	(0.0218)
$\beta_2$disease*post	-0.0351	1.075***
	(0.0250)	(0.0146)
$\beta_3$disease*hospital	0.0425	0.751***
	(0.0293)	(0.0116)
$\beta_4$hospital	-0.0189	1.266***
	(0.0677)	(0.0443)
$\beta_5$disease	-0.0279*	0.924***
	(0.0154)	(0.00797)
$\beta_6$post	0.309***	0.914***
	(0.0312)	(0.0154)
δLoS	0.0296***	
	(0.000456)	
αConstant	7.016***	
	(0.188)	
Observations	11,815	11,711
R-squared	0.290	0.326[13]
Number of hospitals	603	499

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 20 presents the results of the subgroup analysis; columns 1 and 2 display the DDD results for bills and LoS for cataract patients, while columns 2 and 4 display the results for HIVD patients. In the case of cataract, neither bills nor LoS for patients from targeted hospitals and with the targeted disease experienced any significant changes post-policy compared to other patients in the control groups (i.e., patients with similar diseases such as cataract from targeted hospitals and those from untargeted hospitals). For HIVD, bills for patients from targeted hospitals and with the targeted disease decreased by 19.8% post-policy relative to other patients in the control groups (i.e., patients with similar diseases such as HIVD from targeted hospitals and those from untargeted hospitals). Moreover, the LoS for patients from targeted hospitals and with

[13] The fixed effect of Poisson regression does not output the intercept; thus, it is important to calculate the R-square. In the pattern of Least Square Dummy Variable (LSDV) estimation, the Poisson model can output a pseudo-R-square. Hence, we report the pseudo-R-square that is calculated in LSDV estimation to reflect the goodness to fit of the Poisson regression.

the targeted disease remained unchanged post-policy relative to other patients in the control groups. Consequently, the adoption of the PPS appears to have had no significant impact on the bills and health outcomes for cataract patients, while it reduced the bills for HIVD patients.

Similarly to the overall analysis, the subgroup analysis continued to confront the issue of limited sample sizes, and the intervention group did not exhibit a parallel trend compared with any control group in the pre-policy stage, either for bills or for LoS. Therefore, the findings on bills and LoS by DDD estimation may be influenced by the small sample size and the large standard deviation of the data, and thus are not very reliable.

Table 20. DDD results for cataract analysis and HIVD analysis

Diseases	cataract		HIVD	
	(1)	(2)	(3)	(4)
VARIABLES	bills	LoS	bills	LoS
$\beta_0$hospital*disease*post	-0.0565	1.032	-0.198***	1.028
	(0.109)	(0.0759)	(0.0629)	(0.0325)
$\beta_1$hospital*post	-0.197**	0.691***	0.187***	1.031
	(0.0980)	(0.0391)	(0.0499)	(0.0267)
$\beta_2$disease*post	-0.0839	0.897**	-0.0214	1.068***
	(0.0790)	(0.0482)	(0.0246)	(0.0151)
$\beta_3$disease*hospital	-0.119*	0.873***	-0.0685*	1.022
	(0.0621)	(0.0355)	(0.0367)	(0.0186)
$\beta_4$hospital	-0.122	1.645***	0.0116	1.040
	(0.156)	(0.157)	(0.0723)	(0.0402)
$\beta_5$disease	0.480***	0.522***	-0.112***	1.006
	(0.0461)	(0.0166)	(0.0152)	(0.00911)
$\beta_6$post	0.426***	1.038	0.273***	0.900***
	(0.0879)	(0.0602)	(0.0318)	(0.0161)
δlos	0.0232***		0.0313***	
	(0.00114)		(0.000505)	
αConstant	6.903***		7.818***	
	(0.385)		(0.205)	
Observations	3,157	3,141	8,658	8,555
R-squared	0.166	0.320	0.342	0.308
Number of hospitals	100	84	583	480

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

7.5.3 DiD results for cataract patients from targeted hospitals

According to the trends of bills depicted in Figure 26, the trends in bills for the intervention group AC and control group BC of cataract patients from targeted hospitals met the parallel trend assumption at the pre-policy stage. Consequently, we narrowed our focus to a sub-sample of cataract patients from targeted hospitals, performing DiD estimations between these patients (intervention group) and those with similar, untargeted diseases (control group) from the same hospitals.

The DiD analysis, as shown in Table 21, revealed that the bills for cataract patients from targeted hospitals decreased significantly by 16.7% following the PPS reform, compared to patients with untargeted diseases—a finding consistent with the trends observed in the AC and BC groups.

Table 21. DiD results for cataract analysis on patients from targeted hospitals

Diseases	cataract	
	(1)	(2)
VARIABLES	bills	LoS
disease*post	-0.167**	0.956
	(0.0804)	(0.0486)
disease	0.341***	0.466***
	(0.0486)	(0.0127)
post	0.330***	0.663***
	(0.107)	(0.0449)
LoS	0.0241***	
	(0.00148)	
Constant	6.146***	
	(0.506)	
Observations	1,369	1,368
R-squared	0.203	0.289
Number of hospitals	16	15

This table only report the results for key explanatory variables.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

However, Figure 26 indicates that among patients from untargeted hospitals, the bills for those with untargeted diseases (BD group) continued to increase post-PPS reform, whereas the bills for those with targeted diseases (AD group) remained largely stable. In essence, the PPS reform appears to have impacted not only the bills of patients from targeted hospitals but also had spillover effects on the bills of patients with targeted diseases from untargeted hospitals. It is conceivable that untargeted hospitals began reducing their bills for patients in anticipation of the forthcoming PPS reforms, which rendered our DDD estimates for bills insignificant.

Therefore, by focusing solely on patients from targeted hospitals and conducting the DiD estimation, we might conclude that PPS effectively reduces bills for patients with targeted diseases. However, including patients from untargeted hospitals in the DDD estimation revealed that the apparent effects of PPS on bills for patients with targeted diseases from targeted hospitals were neutralized by the spillover effects on patients with targeted diseases from untargeted hospitals.

7.6 Discussion

Based on the DDD estimation, for all patients from targeted hospitals with targeted diseases, their bills decreased by 15.7%, aligning with the expectation that bills would decline following the PPS reform. Conversely, they had the rate 1.093 times higher for LoS, contrary to the anticipated reduction after the PPS reform.

In the subgroup analysis, for cataract patients from targeted hospitals with targeted diseases, neither the bills nor the LoS showed significant changes post-policy compared to other patients. However, for HIVD patients from targeted hospitals with targeted diseases, their bills decreased by 19.8%, while their LoS remained unchanged post-policy relative to other patients. The results indicate that the PPS reform only reduced bills for frail elderly HIVD patients from targeted hospitals and did not reduce LoS for either cataract or HIVD frail elderly patients.

Figures 28 and 29 present forest plots of PPS studies from Chapter 3, updated to include results from this study labeled as Zhao (2023). In these plots, our finding that the PPS reform reduced bills for frail elderly patients aligns with most previous studies; however, our finding that PPS reform increased LoS for these patients diverges from earlier research. Nonetheless, our study is not directly comparable to those in the forest plots, as those studies encompassed patients with various diseases across all age groups, thus estimating the average effect of PPS on a general patient population. Our focus on the frail elderly, who are in poorer health and require more complex clinical pathways, means that the effects of PPS on this subgroup may differ from those on the general population, making direct comparisons inappropriate.

Among the studies reviewed, only (Peng et al., 2021)examined elderly patients in a hospital diagnosed with unilateral displaced femoral neck fractures and found an increase in bills following PPS adoption. However, because this study did not report the standard error or the exact p-value for its results, we could not include it in the forest plots. Their study's intervention, a single-disease PPS reform, paralleled ours, and attributed the failure of cost control among elderly patients to the reform's focus solely on primary diagnoses while overlooking other patient characteristics (e.g., age, gender, disease severity, health- or disease-related complications) and failing to set reasonable tariffs for elderly patients. Our findings that PPS increased LoS for the frail elderly supported the view of (Peng et al., 2021).The frail elderly, often with chronic diseases or complications, require more complex treatment, incur higher costs, and have longer hospital stays, which complicates efforts to reduce their bills or LoS under a single-disease PPS.

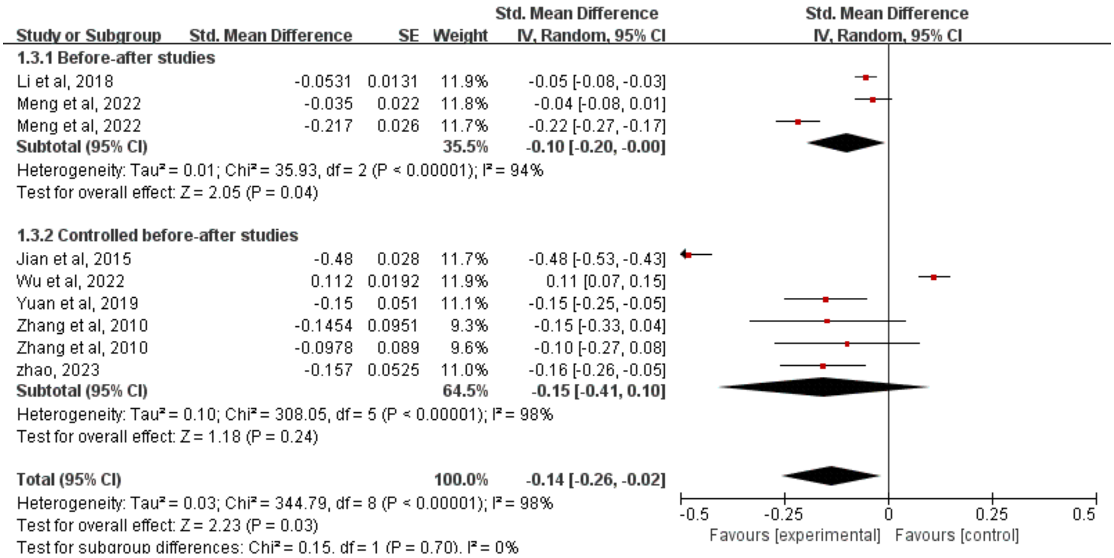


Figure 28. Forest plot of PPS effect on payments including results of this study

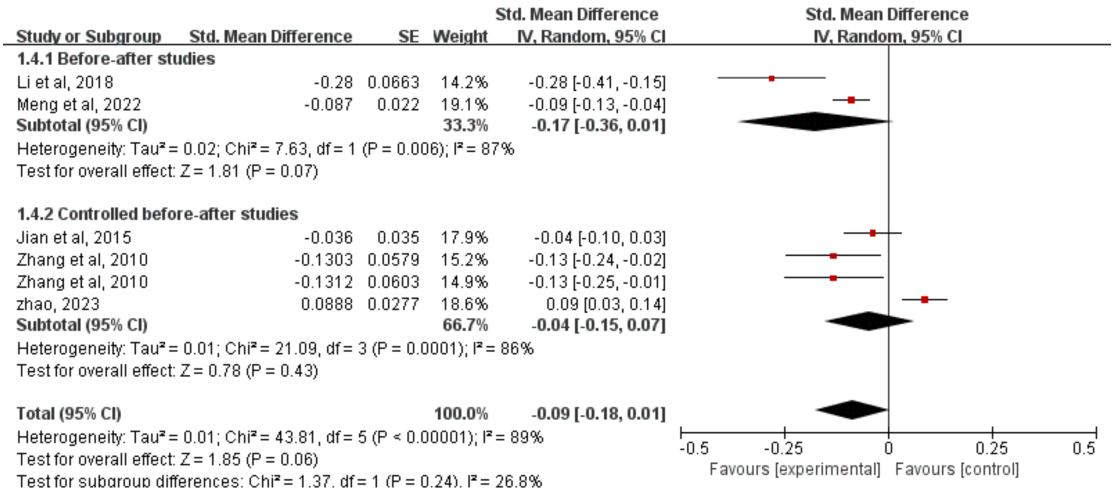


Figure 29. Forest plot of PPS effect on LoS including results of this study

Furthermore, in the DDD estimation, we observed that the bills for cataract patients from targeted hospitals remained unchanged post-reform compared to all other patients. However, in the DiD estimation, which focused solely on patients from targeted hospitals, we noted that bills for cataract patients decreased post-reform relative to patients with untargeted diseases. Therefore, in the DDD estimation, the PPS effect on cataract patients from targeted hospitals may have been negated by the spillover effects from patients with untargeted diseases in targeted hospitals.

Recall from Chapter 6 the study on the PPS reform in Chengdu in 2011, which had a limitation: the absence of data on patients with PPS-targeted diseases who were unaffected by the PPS reform. This gap prevented a more thorough investigation into hospital responses to the reform across different patient subgroups. The possible hospital reactions to the reform were threefold: firstly, hospitals may shift costs from PPS-targeted patients to those with targeted diseases but unaffected by PPS; secondly, hospitals might reduce bills for PPS-targeted patients while keeping bills for unaffected patients with targeted diseases unchanged; thirdly, bills for both groups might decrease due to a spillover policy effect. In this chapter, through DDD and DiD estimations on cataract patients, we observed the spillover effect of the 2018 PPS reform and developed a more comprehensive understanding of its implications for hospital behavior.

The strength of our study lies in the empirical method adopted. Had we merely chosen patients with PPS diseases from targeted hospitals as the treatment group and those from unaffected hospitals as the control group, the two groups would not have been entirely comparable due to the selection of targeted hospitals. Alternatively, selecting patients with PPS diseases from tier 3 hospitals as the treatment group and patients with similar untargeted diseases as the control group might have seemed more comparable, but this approach would have significantly limited the number of observations in our analysis. Our DDD estimation thus helps overcome the heterogeneity between targeted and untargeted hospitals, while also enabling a larger sample size.

A limitation of the study on the PPS reform in Chengdu in 2011, discussed in Chapter 6, was that the data lacked patients with PPS-targeted diseases who were unaffected by the PPS reform. Consequently, we could not conduct a more in-depth investigation of hospitals' reactions to the PPS reform across various patient groups. In this chapter, DDD estimation not only methodologically improved our empirical strategies but also enabled more detailed analyses of different research objectives and provided a clearer depiction of hospital behavior under PPS reform.

However, several limitations predominantly linked to dataset restrictions constrain our study.

Firstly, the data encompasses only elderly individuals, excluding coverage for all 44 diseases. The targeted patients in the data are primarily those with cataract and HIVD—diseases prevalent among the elderly. Given this selected population, the study's scope is limited to investigating PPS effects among the frail elderly, covering only a subset of the population targeted by the policy, and cannot represent the general effects of PPS on all targeted patients. Another limitation is the absence of common pre-policy trends. According to DDD estimation assumptions, there should be a common trend for the treatment and control groups at the pre-policy stage within PPS-targeted hospitals. For patients from untargeted hospitals, a common trend should exist at both pre-policy and post-policy stages as they are not affected by the policy. However, as the descriptive results section illustrates, patients from PPS-targeted hospitals with targeted diseases and those with similar untargeted diseases do not display the expected common trends before the policy. Similarly, patients from untargeted hospitals do not exhibit the required common trends throughout the study period. Additionally, due to the small sample size, there are fluctuations in the trends of outcome variables. Therefore, based on the trend of outcome variables, there should be no observed policy effect in the DDD results. Nonetheless, the DDD results indicate significant decreases in bills and increases in LoS for patients from targeted hospitals with targeted diseases. These DDD results, contradicting the findings from the trend analysis of the treatment and control groups, are not convincing.

These unconvincing results may be attributable to the large standard deviations coupled with a small sample size in the DDD estimation. To increase the sample size and achieve a lower standard deviation, we are considering bootstrap as a potential method. However, according to the study by Bertrand et al. (2004), bootstrap can only compute consistent standard errors when the number of observations in each cross-section of the panel data is sufficiently large. In their research, they implemented bootstrap by drawing replacement matrices to construct a bootstrap sample, applying OLS to this sample, and performing a t-test on the coefficient for the bootstrap sample. The t-distribution of the coefficient in the bootstrap sample approached randomness as the number of observations in each cross-sectional time horizon increased, but the method was inefficient when the observations in each cross-sectional time horizon of the panel were fewer than 50. Unfortunately, our data do not meet this criterion since the number of observations for patients from targeted hospitals and those with targeted diseases in the second half of 2019 was below 50.

7.7 Conclusion

According to the results, following the PPS reform in Chengdu in 2018, among the frail elderly patients, bills for PPS-targeted patients decreased and LoS for PPS-targeted patients increased. The decrease in bills occurred in HIVD patients but was absent in cataract patients; the negligible PPS effect on bills for cataract patients could be explained by a spillover effect. However, these results were biased due to the small sample size and the violation of the parallel trend assumption of DDD, suggesting that these findings may not accurately represent the true impact of the PPS reform.

8. Discussion and conclusion

8.1 Key objectives

This dissertation first provides a general background of the payment system and payment reforms in China. In the background, the dissertation first introduces existing payment methods of FFS, GB and PPS in China and provides the comparison among these payment methods from the aspects of cost control, increasing activities and health service quality, then illustrates the reimbursing process and the relationship between the payer, the hospitals and the patients under different payment methods, and finally describes the development of GB and PPS reforms in China.

Following the background, the dissertation provides a theoretical framework to predict the changes in bills, LoS and service volumes after replacing FFS with GB and to predict the changes in bills, LoS and patient health outcomes after replacing FFS with PPS.

We develop a three-agent theoretical framework, including the patient, payer and hospital to predict how bills, LoS and the number of patients might change following the payment reform. According to the theoretical framework, the payer adopts GB to decrease the total health payment to the hospital; at the same time, the payer encourages hospitals to guarantee their health service quality. When faced with a tighter financial constraint, the hospitals will reduce the supply of health care (measured by average LoS) for patients and decrease costs, and the hospitals might reduce the volume of patients to ensure costs do not exceed their budget. Thus, we predict that bills, average LoS and the number of patients for hospitals will decrease after the GB adoption. And we predict that the bills, OOP and LoS for patients will decrease after the PPS adoption, while 30-day readmissions for patients will increase after the PPS adoption.

The dissertation provides a review of the empirical literature, to synthesize the effects of changing from FFS to either GB or PPS from previous studies in China. Upon introducing the context and data sources in Chengdu in China, the dissertation evaluates three payment reforms in Chengdu, drawing lessons from these empirical findings.

The key objectives of this dissertation are to evaluate the payment reforms in Chengdu in China from the perspective of bills and health service quality, including the GB reform in 2013, the first PPS reform which targeted 10 diseases in 2011, and the second PPS reform which targeted 44 diseases and 101 diagnostic groups in 2018. We develop the theoretical framework to predict policy effects and then investigate the effects empirically.

For the three payments reforms in Chengdu, there are different research objectives for each payment reform. For the GB reform in 2013, the objectives are to investigate the changes in bills and LoS for the patients and the numbers of patients for the hospitals after the reform. For the PPS reform in 2011, the objectives are to investigate the changes in bills, OOPs, LoS and readmissions for the targeted patients after the reform. And the research questions for the PPS reform in 2018 are to investigate the changes in bills and LoS for the frail elderly patients with targeted diseases from targeted hospitals after the reform.

8.2 Empirical methods

We test the predictions of theoretical framework by applying various empirical methods, appropriate to the form of the evaluation and the constraints of the available data.

In the study of GB reform in 2013, we apply ITS analysis to evaluate the GB effect on bills, LoS and the number of patients. Moreover, we use Cumby–Huizinga test (Cumby and Huizinga, 1992) for residual autocorrelation, and we re-estimate the model specifying lags for outcome variables to correctly account for the autocorrelation.

In the study of PPS reform in 2011, we apply DiD estimation to evaluate the PPS effects on bills, OOPs, LoS and 30-day readmissions. To test the robustness of our results, we combine the DiD estimation with PSM in order to make more comparable treatment and control groups, and we apply the 1:1 nearest neighbor matching without replacement.

In the study on PPS reform in 2018, we apply DDD estimation to evaluate the PPS effects on bills and LoS for the frail elderly patients.

8.3 Key findings

Then we have the empirical findings for the GB and PPS reforms in Chengdu.

In the study on GB reform in 2013, we find that GB adoption has immediate effects on bills and the number of patients. Bills increase by 39.8% and the number of patients decrease by 12.7% in the first year of GB adoption. After adjusting for residual correlation, these immediate effects of the policy become insignificant, and the post-policy trend of bills decreased by 17.8%.

In the study on PPS reform in 2011, according to the DiD (DiD+PSM) estimation, bills decreased by 10.0% (10.1%), OOP decreasesd by 10.0% (9.9%), the rate for LoS was 0.935 (0.892) times lower, while the probability of being readmitted in 30 days increased by 2.4% (2.7%). Moreover, our study found the heterogenous performance of tier 2 and tier 3 hospitals under PPS. Both tier 3 and tier 2 hospitals decreased bills, OOP and LoS for patients, while tier 3 hospitals maintained the health outcomes but in tier 2 hospitals readmissions increased by 2.8% following the payment reform.

In the study on PPS reform in 2018, according to the DDD estimation, for all the patients from targeted hospitals and with targeted diseases, their payments decreased by 15.7% and they had the rate 1.093 times higher for LoS after the policy compared with the other patients. For cataract patients from targeted hospitals and with targeted disease, neither the payments nor the LoS experience significant change after the policy compared with the other patients. For HIVD patients from targeted hospitals and with targeted disease, their payments decreased by 19.8% while LoS experienced no change after the policy compared with the other patients. However, these results were biased since the sample size was too small and the parallel trend assumption of DDD was violated, thus, these findings may not reflect a real impact of PPS reform.

8.3 Strength and weakness

The strength of the dissertation is that different methodologies have been adopted to analyze the policy effects.

In the section of literature review, forest plots are applied to synthesize the PPS and GB policy effects of previous studies in the quantitative way.

In the study on GB reform in 2013, ITS estimation allows us to not only investigate the initial change on the outcomes after GB reform but also investigate the changes in the trends of outcomes.

In the study on PPS reform in 2011, we are able to identify intervention and control groups, thus, we adopt the DiD estimation to estimate the average treatment effect on intervention group, and combine the DiD with PSM to assess the robustness of the estimation.

In the study of PPS reform in 2018, since the reform was only targeted to 44 diseases among several tier 3 hospitals while all the tier 2 hospitals were untargeted, there exists heterogeneity between the targeted hospitals and other unaffected hospitals. If we select the patients of PPS diseases from targeted hospitals as treatment group and select the patients of the similar untargeted diseases as control group, the treatment group and control group appear comparable, however, the number of observations in our analysis will be quite small. Hence, rather than only focusing on the targeted hospitals, we include both targeted and untargeted hospitals and conduct DDD estimation to overcome the heterogeneity between targeted hospitals and untargeted hospitals.

There is a general limitation across the thesis' three empirical studies in the lack of an accurate measurement for treatment intensity. In the dissertation, LoS is used as a proxy indicator of intensity. According to previous studies, LoS can be the indicator of treatment intensity (Charunwatthana and Supakankunti, 2014), while there are still other studies claiming that LoS can be the indicator of hospital performance (Boes and Napierala, 2021). Hence, LoS is not a perfect measurement for treatment intensity, since its ambiguity in either relating to treatment intensity or health outcomes. However, due to the limitation of data, we have no other information about treatment intensity.

There are also specific limitations for each empirical study.

For the study on GB reform in 2013, the samples are unbalanced in the data. A large number of hospitals are dropped since the data for some years are missing. Due to this limitation, after adjusting for residual autocorrelation, the number of observations in the analytical sample is quite small, which attenuates the validity of ITS estimation.

Also, due to the lack of information on OOPs, we cannot investigate whether the hospitals transfer the financial burden from reimbursed payments to OOPs. Third, due to the lack of other quality indicators (e.g., the self-reported health status of discharged patients or the readmission rates), we cannot test the impacts on health service quality in the empirical chapter.

For the study on PPS reform in 2011, our data only contains patients who are the enrolled in UEBMI and URBMI and discharged from tier 2 and tier 3 hospitals. This means the analysis does not allow selection of patients with PPS targeted diagnoses but unaffected by PPS policy (e.g. the patients with PPS targeted diagnoses but enrolled in NRCMS) as a control group. This might undermine the comparability of our treatment and control groups, though we applied PSM to reduce heterogeneity. Also, since the data do not contain the patients with PPS targeted diagnoses but unaffected by PPS policy, we cannot investigate whether the hospitals shifted attention from the patients affected by the policy to the patients with same targeted diagnoses but unaffected by the policy.

For the study on PPS reform in 2018, the data consists of elderly people and does not cover all 44 diseases, but cataract and HIVD. Due to the selected population, the study can only investigate the PPS effect among frail elderly people and cannot represent the general PPS effect to all the targeted patients.

Also, in the policy document, cataract and senility were two separated diseases, however, in the data, the disease name was not recorded accurately and both senile cataract and cataract could be recorded as cataract, so that we can only identify the patients under the broad category of cataract.

Finally, the patients from intervention group and control groups did not have the common trend before the policy as they are supposed to and, due to the small samples, there were fluctuations in the trend of outcome variables. Thus, according to the trend of outcome variables, there should not be any policy effect in the DDD results. Even so the DDD results suggested that the PPS adoption decreased bills and increased LoS for the patients from targeted hospitals and with targeted diseases significantly. However, because parallel trends were not observed in the pre-policy period, the DDD results must be interpreted subject to this caution.

8.4 Implications for policy

For the GB study, according to the empirical findings, there are several policy implications: Since Chengdu implemented the 'soft' target GB payment where hospitals can be reimbursed partially if expenditures above the benchmark under some special conditions, the hospital might overstate some information in the annual report to get high reimbursement when their expenditures are over benchmark. To avoid this, the payer might provide a stronger provision on verifying hospital reports, and provide stronger motivation for hospitals to decrease bills. Also, for the number of patients, there is no change for tier 3 hospitals after GB, but a decrease in the number of patients is observed among tier 2 hospitals after GB adoption. Thus, the payer might provide more comprehensive inspection for the hospital performance under GB, especially monitoring of the number of patients of hospitals, and rules to avoid hospitals refusing to admit patients.

Drawing conclusions from the PPS reform in 2011, we can provide two implications. First, to guarantee health service quality, monitoring of quality accompanied with the PPS reform is required, especially for tier 2 hospitals. Second, according to the hierarchy of hospitals in China, tier 2 hospitals are of smaller scale and with lower capacity, so appear more suitable for patients with less severe conditions. Being based on the finding that only tier 2 hospitals cause adverse health outcomes for patient after PPS adoption in Chengdu, other than motivating tier 2 hospitals to improve their health service quality in PPS reform, it is also plausible to reallocate the patients among tier 2 hospitals and tier 3 hospitals according to the relative severity of their conditions.

According to the study in the PPS reform in 2018, the findings among the frail elderly that the LoS for PPS targeted patients increased after the reform implied that the single disease PPS reform was not suitable for some patient groups and did not have the policy effect on them as we predicted., Thus, it is necessary to set the clinical path for the special groups as the frail elderly in PPS reform. However, due to the data limitation, these findings are not as convincing as the finding from the study of the 2011 PPS reform.

Other than specific implications for GB and PPS reforms, we also have the general implications for the multi-method payment reform in Chengdu in China.

The PPS reforms in Chengdu were still in initial stage when conducting these evaluations. The PPS reform in 2011 only targeted 10 diseases which was completely in the form of single-disease PPS. The PPS reform in 2018 increased the targeting diseases to 44 and developed 101 diagnostic groups under the 44 diseases which could be regarded as the transition stage from single-disease PPS to DRGs.

One reason for the limited impact of PPS reforms in Chengdu from 2011 to 2018 was that there was not a formal nationwide DRG grouper for China until the advent of CH-DRG (the formal DRG grouper for China) in 2019. Moreover, since the implementation of DRGs had high requirement on healthcare informatics for hospitals, which would take a longer period for hospitals in China to be prepared with DRGs payment. For instance, the DRGs adoption required the accurate recording for ICD codes, however, the physicians in hospitals in China were not familiar with the rules of ICD coding and there was the lack of professional clinical coders in China. Thus, the health departments of Chinese local governments had to keep providing the coding training for technicians in the hospitals before completely adopting DRGs payments in their regions.

The most obvious limitation of single disease PPS was that the policy effect was only in a small range on several targeted diseases but could not change the hospital behavior in a wider sense. Moreover, according to the study on PPS reforms in 2018 among the frail elderly, we found that the single disease PPS tariff set for general patients may not be suitable for the frail elderly group which made the hospitals difficult to decrease the LoS for these patients. Thus, another obvious limitation of the single-disease PPS reform was that it was not able to consider the characteristics of different patients and was not able to set up the different tariffs for different patient groups.

Being faced with the limitations of single disease PPS reforms, we recommend that the PPS targeting diseases need to be expanded, that the reasonable tariffs for different patient groups (especially for the groups with complications) need to be developed, and that it is necessary for the hospitals in Chengdu to be prepared with DRGs reform and adopt the CH-DRG system in future.

However, being faced with the current lack of professional clinical coders in China, it will take a longer time for hospitals in Chengdu to get prepared with the adoption of DRGs payment. Thus, before the complete adoption of DRGs payment in Chengdu, the feasible pattern of payment reform is to combine the PPS and GB reforms, implementing PPS pilots for targeted diseases under the GB payment scheme applied to all the patients from tier 2 and tier 3 the hospitals, to strongly motivate cost control. At the same time, as Chengdu previously implemented the ‘soft’ target GB payment where hospitals can be reimbursed partially if expenditures above the benchmark under some special conditions, the payer might need to provide a stronger provision to verify hospital reports and reduce reimbursement for the excessive expenditures in the next GB reforms.

As well as cost control, another important aspect of payment reforms is health service quality. In the payment reforms, hospitals need to commit to maintaining and even elevating health service quality. However, the results of this dissertation showed that the PPS reform in 2011 increased 30-day readmissions, implying poorer health outcome for patients. Thus, in the next payment reforms, the payer needs to provide stronger incentives to improve health service quality packaged with the cost control. To this end, we propose two possible approaches as follows.

The first approach is to include health service quality management in PPS and GB payments, for instance to adopt hospital performance when implementing payment reforms. According to the definition of WHO, hospital performance assessment encompasses six interrelated dimensions: clinical effectiveness, safety, patient centredness, responsive governance, staff orientation, and efficiency (Veillard et al., 2005), which can provide comprehensive provision on hospitals to improve the quality, control the expenditure and increase the efficiency. Before 2019, China did not have a very comprehensive assessment scheme and did not conduct comprehensive performance assessment on hospitals, but only required hospitals to submit the statistics on several simple quality indicators (e.g., emergency mortality rates, readmissions) in the annual quality assessment. Consequently, there was the lack of a strong motivation for hospitals to improve quality in the payment reforms.

More recently, alongside with the nationwide DRG-PPS reforms started from 2019, China has initiated the nationwide hospital performance assessment, first implemented among tier 3 public hospitals in 2019 and then expanded to tier 2 hospitals in 2020, and including indicators of quality, safety, payment control, efficiency, patient satisfaction, and sustainable development for hospitals. China now combines hospital performance assessment with DRG-PPS reform: hospitals need to not only report the patients outcomes at hospital level but also report patient outcomes according to a series of performance indicators for every diagnostic group. The payer then assesses the outcomes for every group and compares the outcomes between different hospitals within the same groups. This, therefore, applies DRGs as the quality management tool and strengthens the performance competition within treating the similar diseases among hospitals. Due to the lack of the latest data, we cannot investigate the change in patient health outcomes after the adoption of hospital performance evaluation in Chengdu, but we can infer that the quality assessment of hospital performance evaluation will motivate the hospitals to improve patient health outcomes in the payment reforms.

The second approach is to explore value-based healthcare payment in the payment reforms. Value-based healthcare was first proposed by Porter and Teisberg (2006), which claimed that the value for patients must be the overarching principle in the organization and management of health care delivery systems, and hospitals should be paid based on patient health outcomes under the value-based payment. Pay for performance (P4P) is a form of value-based payment, which offers financial incentives to hospitals for meeting certain performance measures and evaluates process quality and efficiency. Although P4P was not implemented as widely as PPS or GB, it was still adopted in the payment reforms in some areas in China. For instance, in the payment reforms in Guizhou province, the P4P was combined with GB payment to motivate the hospitals to improve the health outcomes for patients. The study of Zhou et al. (2021) showed that the inclusion of P4P in GB payment in Guizhou province helped to reduce the quality risks associated with cost control by improving the process quality among AMI and pneumonia patients, resulting in the increase in aspirin use at discharge, discharge with β-blocker and smoking cessation advice among AMI patients, and the increase in Oxygenation index assessment and pneumonia vaccine use rates among pneumonia patients. Thus, the success of P4P adoption in Guizhou province makes it visible for Chengdu to adopt P4P in the next reforms to motivate quality improvements.

8.5 Implications for research

We have several implications of this dissertation for research, mainly based on the aspects of methodology and data limitations.

In the empirical chapters, we applied ITS, DiD and DDD as the empirical methods according to the different policy exposure of the reforms, with the method of every chapter becoming more complex than the previous one. In chapter 5, we could only investigate the before-after effect of GB reform since all the hospitals in the data were targeted by GB. In chapter 6, with the PPS policy variation by disease, we could identify the intervention and control groups and conduct the DiD estimation, which is an established approach designed for causal inference. In chapter 7, with the PPS policy variation by disease and by hospital, we conducted DDD estimation, which did not only improve the empirical strategy methodologically from the aspect of causal inference but also allowed for the more focused analyses.

That said the empirical studies of this dissertation would have been of higher quality if the data limitations can be overcome.

First, the choices for the measurement for outcome variables were quite limited in our data. For example, in the study on the 2013 GB reform, since there was no information of readmission rates for the hospital, we could not investigate the GB effect on hospital service quality. In the studies on the 2011 PPS reform and the 2018 PPS reform, due to the lack of information on treatment intensity (e.g., the prescription records for the patients), we could only use the LoS to measure the treatment intensity. Actually, the hospitals in China have recorded the treatment details for patients in the front pages of the clinical notes, but the information of patient front page was separated from the claims data and we could not observe the treatment details in our patient data. Thus, one research implication is that the data may already be available and the front page can be merged with the claim data, then it will allow the researchers to take the comprehensive and accurate evaluation on the reforms.

Second, there is scope for improved data quality notably regarding missing or imprecise values for some variables in hospital data. The hospital data was recorded completely manually until 2010 and was gradually recorded in the electronic system from 2010. Thus, part of the observations in the hospital data were coded from the paper records and some observations were missed during the coding process. Also, for the samples after 2011 when the electronical system was adopted, there were still missing values for some variables, suggesting that the datas were not reported in a very precise way. Thus, the other research implication is improving the data quality and reducing the missing values. This would ensure that empirical research would be subject to less systematic error and be more efficient.

9. Bibliography

AVERILL, R. F. 1994. The evolution of case-mix measurement using DRGs: past, present and future. *Case-Based Telematic Systems towards Equity in Health Care,* 14**,** 75-83.http://doi.org/10.3233/978-1-60750-858-8-75.

BAKER, L. C. 1999. Association of managed care market share and health expenditures for fee-for-service Medicare patients. *JAMA,* 281**,** 432-7.http://doi.org/10.1001/jama.281.5.432.

BERTRAND, M., DUFLO, E. & MULLAINATHAN, S. 2004. How Much Should We Trust Differences-In-Differences Estimates?*. *The Quarterly Journal of Economics,* 119**,** 249-275.http://doi.org/10.1162/003355304772839588.

BOES, S. & NAPIERALA, C. 2021. Assessment of the introduction of DRG-based reimbursement in Switzerland: Evidence on the short-term effects on length of stay compliance in university hospitals. *Health Policy,* 125**,** 739-750.https://doi.org/10.1016/j.healthpol.2021.01.010.

CHARUNWATTHANA, W. & SUPAKANKUNTI, S. 2014. Intensity response of Thai public hospitals to payment system change. *International Journal of Economic Policy in Emerging Economies,* 7**,** 335-345.http://doi.org/10.1504/IJEPEE.2014.066651.

CHEN, B. & FAN, V. Y. 2016. Global Budget Payment: Proposing the CAP Framework. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing,* 53**,** 0046958016669016.http://doi.org/10.1177/0046958016669016.

CHEN, Y. J., ZHANG, X. Y., YAN, J. Q., XUE, T., QIAN, M. C. & YING, X. H. 2023. Impact of Diagnosis-Related Groups on Inpatient Quality of Health Care: A Systematic Review and Meta-Analysis. *Inquiry,* 60**,** 469580231167011.http://doi.org/10.1177/00469580231167011.

CHENGDU_BUREAU_OF_STATISTICS 2017. Chengdu Statistical Year Book.

CHENGDU_BUREAU_OF_STATISTICS 2021. Chengdu Statistical Year Book.

CLARKE, A. & ROSEN, R. 2001. Length of stay: how short should hospital care be? *The European Journal of Public Health,* 11**,** 166-170.http://doi.org/10.1093/eurpub/11.2.166.

CUMBY, R. E. & HUIZINGA, J. 1992. Testing the Autocorrelation Structure of Disturbances in Ordinary Least Squares and Instrumental Variables Regressions. *Econometrica,* 60**,** 185-195.http://doi.org/10.2307/2951684.

CUTLER, D. M. 1995. The Incidence of Adverse Medical Outcomes Under Prospective Payment. *Econometrica,* 63**,** 29-50.http://doi.org/10.2307/2951696.

DEMIR, E., CHAUSSALET, T. J., XIE, H. & MILLARD, P. H. 2008. Emergency readmission criterion: a technique for determining the emergency readmission time window. *IEEE Transactions on Information Technology in Biomedicine,* 12**,** 644-649.http://doi.org/10.1109/TITB.2007.911311.

DOU, G., ZHANG, Y., HE, Y., HUANG, Q., YE, Y., ZHANG, X., WANG, W. & YING, X. 2019. Impact of the Global Budget Payment System on Expenditure of Cardiovascular Diseases: An Interrupted Time Series Analysis in Shanghai, China. *Int J Environ Res Public Health,* 16.http://doi.org/10.3390/ijerph16081385.

ELLIS, R. P. & MCGUIRE, T. G. 1986. Provider behavior under prospective reimbursement. Cost sharing and supply. *Journal of Health Economics,* 5**,** 129-151.http://doi.org/10.1016/0167-6296(86)90002-0.

FENG, J., YU, Y. & LOU, Y. 2015. Health demand and health cost growth in China. *Social Sciences in*

China, 3**,** 85-103.https://lib.cqvip.com/Qikan/Article/Detail?id=664232702.

FRIEBEL, R., DHARMARAJAN, K., KRUMHOLZ, H. M. & STEVENTON, A. 2017. Reductions in Readmission Rates Are Associated With Modest Improvements in Patient-reported Health Gains Following Hip and Knee Replacement in England. *Medical Care,* 55**,** 834-840.http://doi.org/10.1097/mlr.0000000000000779.

GUAN, X., ZHANG, C., HU, H. & SHI, L. 2020. The impact of global budget on expenditure, service volume, and quality of care among patients with pneumonia in a secondary hospital in China: a retrospective study. *BMC Public Health,* 20**,** 522.http://doi.org/10.1186/s12889-020-08619-3.

GUTERMAN, S. & DOBSON, A. 1986. Impact of the Medicare prospective payment system for hospitals. *Health Care Financing Review,* 7**,** 97-114.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4191526/.

HAYFORD, T. B. 2012. The impact of hospital mergers on treatment intensity and health outcomes. *Health Services Research,* 47**,** 1008-1029.http://doi.org/10.1111/j.1475-6773.2011.01351.x.

HE, Q., LYU, P., WANG, Y., CHEN, S. & LI, J. 2019. Current Situation and Suggestions for Instant Settlement of Off-site Medical Treatment from the Perspective of Public Hospitals. *Modern Hospital Management,* 17**,** 34-36.http://doi.org/10.3969/j.issn.1672-4232.2019.02.010.

HE, R., MIAO, Y., YE, T., ZHANG, Y., TANG, W., LI, Z. & ZHANG, L. 2017. The effects of global budget on cost control and readmission in rural China: a difference-in-difference analysis. *Journal of Medical Economics,* 20**,** 903-910.http://doi.org/10.1080/13696998.2017.1336448.

HODGKIN, D. & MCGUIRE, T. G. 1994. Payment levels and hospital response to prospective payment. *Journal of Health Economics,* 13**,** 1-29.http://doi.org/10.1016/0167-6296(94)90002-7.

HU, H., ZHAO, L., YONG, Y., NICHOLAS, S., MAITLAND, E., ZHAO, W., YAN, H., MA, Y. & SHI, X. 2023. Medical insurance payment schemes and patient medical expenses: a cross-sectional study of lung cancer patients in urban China. *BMC Health Services Research,* 23**,** 89.http://doi.org/10.1186/s12913-023-09078-3.

HU, M., LU, M., DU, S. & JIAN, W. 2014. The Evaluation of the Pilot Study on Diagnosis Related Groups and Prospective Payment System in Beijing. *China Health Insurance***,** 48-52.http://doi.org/10.3969/j.issn.1674-3830.2014.04.024.

HUA, M., MA, X., MORRISON, R. S., LI, G. & WUNSCH, H. 2018. Association between the availability of hospital-based palliative care and treatment intensity for critically ill patients. *Annals of the American Thoracic Society,* 15**,** 1067-1074.http://doi.org/10.1513/AnnalsATS.201711-872OC.

HUANG, L., WEI, C., HUANG, L., HE, H. & HE, C. 2022. Study on the Effect of DRG Payment on Tuberculosis Patients——Taking a Class III Specialized Hospital of Infectious Disease as an Example. *China Health Insurance*.https://d.wanfangdata.com.cn/periodical/zgylbx202204016.

HUANG, Y., LIU, Y., YANG, X., LI, J. & FANG, P. 2016. Global budget payment system helps to reduce outpatient medical expenditure of hypertension in China. *Springerplus,* 5**,** 1877.http://doi.org/10.1186/s40064-016-3565-7.

JIAN, W., LU, M., CHAN, K. Y., POON, A. N., HAN, W., HU, M. & YIP, W. 2015a. Payment Reform Pilot In Beijing Hospitals Reduced Expenditures And Out-Of-Pocket Payments Per Admission. *Health Affairs,* 34**,** 1745-1752.http://doi.org/10.1377/hlthaff.2015.0074.

JIAN, W., LU, M. & HU, M. 2015b. The Situation and Effectiveness Analysis of the Pilot Implement of Diagnosis Related Groups Payment in Beijing. *China Health Insurance***,** 52-55.http://doi.org/10.3969/j.issn.1674-3830.2015.3.12.

JIAN, W., LU, M., LIU, G., CHAN, K. Y. & POON, A. N. 2019. Beijing's diagnosis-related group payment reform pilot: Impact on quality of acute myocardial infarction care. *Social Science & Medicine,* 243.http://doi.org/10.1016/j.socscimed.2019.112590.

LI, H., CHEN, Y., GAO, H., CHANG, J., SU, D., LEI, S., JIANG, D., HU, X., TAN, M. & CHEN, Z. 2019. Effect of an integrated payment system on the direct economic burden and readmission of rural cerebral infarction inpatients: Evidence from Anhui, China. *International journal of environmental research and public health*, 1554.https://doi.org/10.3390%2Fijerph16091554.

LI, H. M., CHEN, Y. C., GAO, H. X., ZHANG, Y., CHEN, L., CHANG, J. J., SU, D., LEI, S. H., JIANG, D. & HU, X. M. 2018. Effectiveness evaluation of quota payment for specific diseases under global budget: a typical provider payment system reform in rural China. *BMC Health Service Research,* 18.https://doi.org/10.1186/s12913-018-3415-0.

LI, L. & ZHANG, S. 2022. Understanding the Public Policy of Global Budget Payment Reform Improves the Quality of Public Healthcare From the Perspective of Patients in China. *Frontiers in Psychology,* 13, 911197.http://doi.org/10.3389/fpsyg.2022.911197.

LI, X. 2018. Theoretical Analysis of Choice of Public Hospital Behavior under Global Budget. *Chinese Health Economics*.http://www.cn-he.cn/ch/reader/view_abstract.aspx?file_no=1201803103&st=alljournals.

LIANG, D., CHEN, L., LIN, B. & WU, S. 2012. The Different Effects of the Two Kinds of Payment Options for Medical Expenses:Based on Empirical Contrast of the Chancheng District and Shunde District in Foshan. *China Health Insurance*, 41-44.https://www.cqvip.com/QK/70846X/201210/43460493.html.

LIU, C. & WANG, Y. 2024. Research Status and Prospect of Cross-regional Medical Treatment in China Based on Bibliometric Analysis. *Chinese General Practice,* 27, 1525-1532.https://www.chinagp.net/CN/10.12114/j.issn.1007-9572.2023.0364.

MAKRIS, M. & SICILIANI, L. 2013. Optimal Incentive Schemes for Altruistic Providers. *Journal of Public Economic Theory,* 15, 675-699.https://doi.org/10.1111/jpet.12047.

MAURER, P. P. 2004. Hospital readmissions-are they predictable and avoidable? *Swiss medical weekly,* 134, 606-611.http://doi.org/10.4414/smw.2004.10706.

MENG, H. 2015. Practical Exploration of Compound Payment under the Total Budget Control. *China Health Insurance*, 49-50.https://lib.cqvip.com/Qikan/Article/Detail?id=664529080&from=Qikan_Search_Index.

MENG, Q. & TANG, S. 2013. Universal Health Care Coverage in China: Challenges and Opportunities. *Procedia - Social and Behavioral Sciences,* 77, 330-340.https://doi.org/10.1016/j.sbspro.2013.03.091.

MENG, Z., HUI, W., CAI, Y., LIU, J. & WU, H. 2020. The effects of DRGs-based payment compared with cost-based payment on inpatient healthcare utilization: A systematic review and meta-analysis. *Health Policy,* 124, 359-367.https://doi.org/10.1016/j.healthpol.2020.01.007.

MENG, Z., ZOU, K., SONG, S., WU, H. & HAN, Y. 2022. Associations of Chinese diagnosis-related group systems with inpatient expenditures for older people with hip fracture. *BMC Geriatrics,* 22, 169.http://doi.org/10.1186/s12877-022-02865-3.

MULLIN, R., VERTREES, J., FREEDMAN, R., CASTIONI, R. & TINKER, A. 2002. Case-mix analysis across patient populations and boundaries: A refined classification system designed specifically for international use. 3M/Health Information System, Connecticut. *Saint Paul, MN: 3M Health Information Systems*.http://apdrgsuisse.ch/public/fr/a-0209-irdrg.pdf.

PENG, H.-M., XU, Y., CI, P.-W., ZHANG, J., ZHANG, B.-Z. & WENG, X.-S. 2021. A Simple Diagnosis-Related Groups-Based Reimbursement System Is Cost Ineffective for Elderly Patients With Displaced Femoral Neck Fracture Undergoing Hemiarthroplasty in Beijing. *Frontiers in medicine*, 733206-733206.https://doi.org/10.3389/fmed.2021.733206.

PENG, J., ZHANG, M., YU, P. & WANG, N. 2018a. Can single disease payment system based on clinical pathway reduce hospitalization costs in rural area? A case study of uterine leiomyoma in Anhui, China. *BMC Health Services Research,* 18, 990.http://doi.org/10.1186/s12913-018-3807-1.

PENG, Y., LEI, H., WANG, H., WANG, L., XU, J. & JIN, C. 2018b. Progress and inspiration of case-based payment system reform in China. *Chinese Health Resources,* 21, 477-481.http://doi.org/10.13688/j.cnki.chr.2018.18102.

PORTER, M. E. & TEISBERG, E. O. 2006. Redefining health care: creating value-based competition on results.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1592405/.

RUMBALL-SMITH, J., HIDER, P. & GRAHAM, P. 2009. The readmission rate as an indicator of the quality of elective surgical inpatient care for the elderly in New Zealand. *The New Zealand Medical Journal,* 122.https://pubmed.ncbi.nlm.nih.gov/19305447/.

SHLEIFER, A. 1985. A Theory of Yardstick Competition. *RAND Journal of Economics,* 16, 319-327.https://scholar.harvard.edu/files/shleifer/files/theory_yardstick_comp.pdf.

SKINNER, J. 2011. Causes and consequences of regional variations in health care. *Handbook of health economics,* 2, 45-93.https://doi.org/10.1016/B978-0-444-53592-4.00002-5.

SONG, J. & CHEN, X. 2018. Thoughts and suggestions of establishing the medical alliance in the tertiary hospital. *Soft Science of Health,* 32, 7-9.http://doi.org/10.3969/j.issn.1003-2800.2018.11.002.

VEILLARD, J., CHAMPAGNE, F., KLAZINGA, N., KAZANDJIAN, V., ARAH, O. A. & GUISSET, A. L. 2005. A performance assessment framework for hospitals: the WHO regional office for Europe PATH project. *International Journal for Quality in Health Care,* 17, 487-496.http://doi.org/10.1093/intqhc/mzi072.

WAGSTAFF, A., YIP, W., LINDELOW, M. & HSIAO, W. C. 2009. China's health system and its reform: a review of recent studies. *Health Econ,* 18 Suppl 2, S7-23.http://doi.org/10.1002/hec.1518.

WANG, X. 2015. Impact Evaluation of Provider Payment Reform under New Rural Cooperative Medical Scheme in Gansu, China.https://doi.org/10.1016/j.jval.2014.08.035.

WEI, T. & FENG, W. 2019. Impact analysis of the payment reform of "same disease with the same price in the same level of hospital" diagnosis-related group based on L City, Guizhou Province: an empirical study *The Lancet,* 394.https://doi.org/10.1016/S0140-6736(19)32399-2.

WU, J., HE, X. & FENG, X. L. 2022. Can case-based payment contain healthcare costs? - A curious case from China. *Social Science & Medicine,* 312, 115384.https://doi.org/10.1016/j.socscimed.2022.115384.

WU, Y. & GUO, Y. 2016. Analysis on the Influence of total Prepaid Health Insurance System on the Running of Specialized Hospital. *China Health Insurance*, 52-55.https://d.wanfangdata.com.cn/periodical/zgylbx201607019.

XIANG, L., ZHONG, Z. & JIANG, J. 2022. The Response of Different-Levels Public Hospitals to Regional Global Budget with a Floating Payment System: Evidence from China. *International Journal of Environmental Research and Public Health,* 19.http://doi.org/10.3390/ijerph192315507.

XIE, X., GUAN, W. & LI, Q. 2015. Evaluation and Consideration on Health Insurance Payment of Shunyi District. *China Health Insurance*, 50-

53.https://d.wanfangdata.com.cn/periodical/zgylbx201508022.

YANG, Y., MAN, X., YU, Z., NICHOLAS, S., MAITLAND, E., HUANG, Z., MA, Y. & SHI, X. 2022. Managing Urban Stroke Health Expenditures in China: Role of Payment Method and Hospital Level. *International journal of health policy and management*, 2698-2706.https://doi.org/10.34172/ijhpm.2022.5117.

YIP, W. & HSIAO, W. 2009. China's health care reform: A tentative assessment. *China Economic Review,* 20, 613-619.https://doi.org/10.1016/j.chieco.2009.08.003.

YIP, W., HSIAO, W., MENG, Q., CHEN, W. & SUN, X. 2010. Realignment of incentives for health-care providers in China. *The Lancet,* 375, 1120-1130.http://doi.org/10.1016/S0140-6736(10)60063-3.

YIP, W. C.-M., HSIAO, W. C., CHEN, W., HU, S., MA, J. & MAYNARD, A. 2012. Early appraisal of China's huge and complex health-care reforms. *The Lancet,* 379, 833-842.https://doi.org/10.1016/S0140-6736(11)61880-1.

YU, H. 2015. Universal health insurance coverage for 1.3 billion people: What accounts for China's success? *Health Policy,* 119, 1145-52. https://doi.org/10.1016/j.healthpol.2015.07.008.

YUAN, S., LIU, W., WEI, F., ZHANG, H., WANG, S., ZHU, W. & MA, J. 2019. Impacts of hospital payment based on diagnosis related groups (DRGs) with global budget on resource use and quality of care: A case study in China. *Iranian journal of public health*, 238-246.https://doi.org/10.18502/ijph.v48i2.818.

ZHANG, J. 2010. The impact of a diagnosis-related group-based prospective payment experiment: the experience of Shanghai. *Applied economics letters,* 17, 1797-1803.http://doi.org/10.1080/13504850903317347.

ZHANG, J., XIONG, X. & LI, J. 2013. The Influence of Global Budget upon the Medical Expense and the Utilization of Medical Services. *China Health Insurance*, 27-32.http://doi.org/10.3969/j.issn.1674-3830.2013.11.007.

ZHANG, M. & HU, M. 2015. Application of DRGs in Medical Insurance Management. *China Health Insurance*, 51-53.http://doi.org/10.3969/j.issn.1674-3830.2015.04.022.

ZHAO, Y. 2015. Mechanism of Prepaid Method of Medical Insurance in Controlling Medical Expense *Chinese Hospital Management,* 35, 45-47.http://www.zgyygl.com/ch/reader/view_abstract.aspx?file_no=2201410172&flag=1.

ZHENG, L., JIA, L. & LI, Q. 2013. Analysis of the Medical Resources Allocation in Secondary and Tertiary Hospitals of Anhui Province in 2011 *Chinese Journal of Social Medicine,* 30, 316-318.http://doi.org/10.3969/j.issn.1674-3830.2013.11.007.

ZHOU, W., JIAN, W., WANG, Z., PAN, J., HU, M. & YIP, W. 2021. Impact of global budget combined with pay-for-performance on the quality of care in county hospitals: a difference-in-differences study design with a propaensity-score-matched control group using data from Guizhou province, China. *BMC Health Service Research,* 21, 1296.http://doi.org/10.1186/s12913-021-07338-8.

10. Appendix

10.1 Appendix for chapter 5

Table A 1. The results of ITS estimation

VARIABLES	(1) bills	(2) LoS	(3) Number of patients
year	0.0117	0.0331	0.0333**
	(0.0147)	(0.231)	(0.0169)
after	0.398***	-0.958	-0.127*
	(0.0607)	(1.046)	(0.0730)
After*year	0.0167	-0.182	-0.0180
	(0.0257)	(0.448)	(0.0311)
tier	0.0330	-1.032	0.848***
	(0.0564)	(0.967)	(0.0675)
Employee number	-0.00145***	-0.0317***	0.00218***
	(0.000161)	(0.00274)	(0.000194)
Bed number	0.00199***	0.0334***	-0.00107***
	(0.000140)	(0.00241)	(0.000171)
equipment number	0.000125**	0.00450***	-0.000586***
	(6.08e-05)	(0.00103)	(7.30e-05)
Constant	5.648***	12.54***	11.50***
	(0.0424)	(0.692)	(0.0489)
Observations	982	981	1,000
Number of id	401	400	401

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

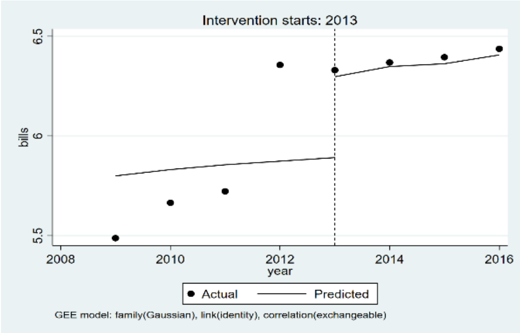


Figure A 1-1. The results of bills

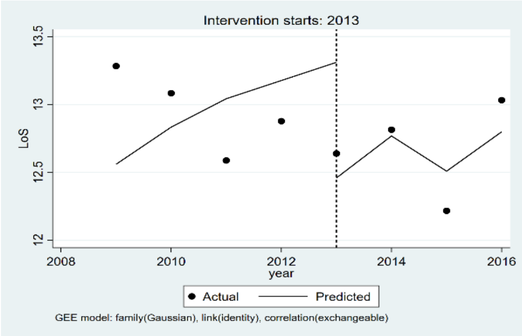


Figure A 1-2. The results of LoS

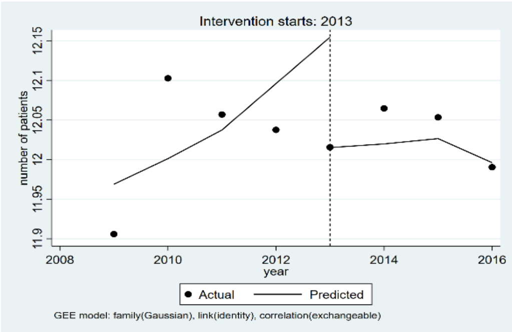


Figure A 1-3. The results of number of patients

Figure A 1. The results of ITS estimation

Table A 2. The results of ITS estimation after adjusting residual correlation

VARIABLES	(1) bills	(2) LoS	(3) Number of patients
year	0.119*	0.545	0.00138
	(0.0684)	(0.692)	(0.0708)
after	-0.0250	-0.653	0.0919
	(0.130)	(1.747)	(0.174)
After*year	-0.178**	-0.986	0.0829
	(0.0755)	(0.854)	(0.0865)
tier	-0.102	-1.293	0.888***
	(0.114)	(1.577)	(0.142)
Employee number	-0.00141***	-0.0234***	0.00170***
	(0.000266)	(0.00393)	(0.000366)
Bed number	0.00193***	0.0236***	-0.00128***
	(0.000237)	(0.00336)	(0.000351)
equipment number	0.000208**	0.00396***	-0.000340***
	(9.35e-05)	(0.00144)	(0.000123)
Constant	6.138***	13.99***	11.70***
	(0.133)	(1.427)	(0.157)
Observations	158	228	155
Number of id	75	75	40

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

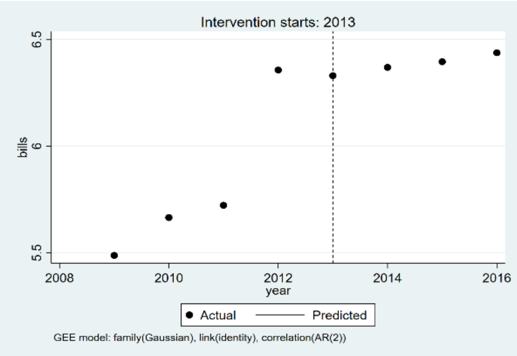


Figure A2-1. The results of bills

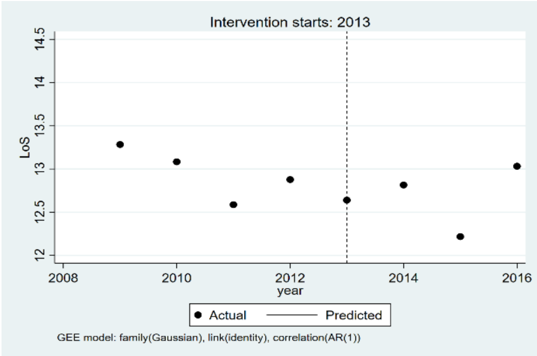


Figure A2-2. The results of LoS

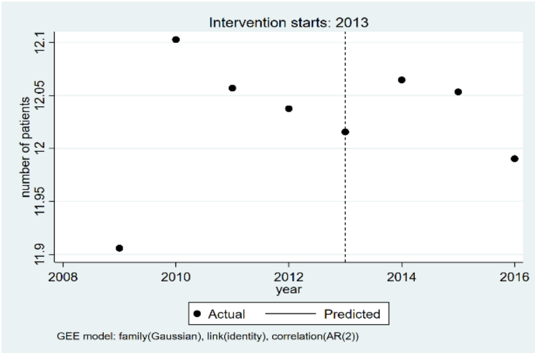


Figure A2-3. The results of number of patients

Figure A 2. The results of ITS estimation after adjusting residual correlation

Table A 3. The results of ITS estimation for tier 3 hospitals

VARIABLES	(1) bills	(2) LoS	(3) Number of patients
year	0.0808***	0.318	-0.0353
	(0.0223)	(0.327)	(0.0222)
after	0.277***	-1.418	-0.100
	(0.0746)	(1.252)	(0.0691)
After*year	-0.103***	-0.790	0.0444
	(0.0314)	(0.514)	(0.0299)
employee number	-0.000666***	-0.00926***	0.000416***
	(0.000142)	(0.00212)	(0.000134)
bed number	0.00107***	0.0118***	-0.000346***
	(0.000111)	(0.00186)	(9.99e-05)
equipment number	3.87e-05	0.000305	0.000109**
	(5.23e-05)	(0.000761)	(4.96e-05)
Constant	5.977***	12.71***	12.88***
	(0.0731)	(1.026)	(0.0747)
Observations	269	269	269
Number of id	167	167	167

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

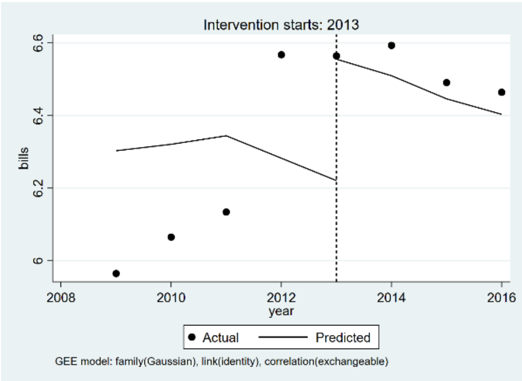


Figure A3-1. The results of bills

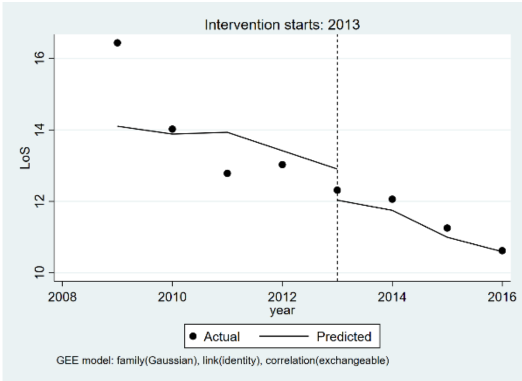


Figure A3-2. The results of LoS

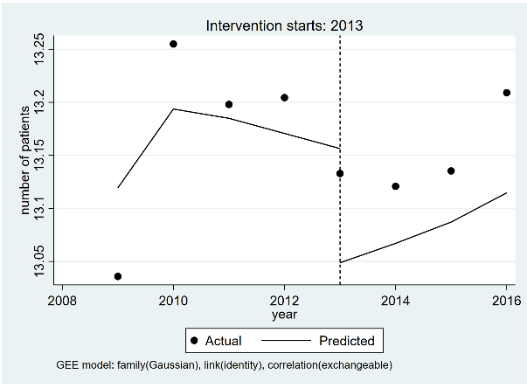


Figure A3-3. The results of number of patients

Figure A 3. The results of ITS estimations for tier 3 hospitals

Table A 4. The results of ITS estimation for tier 2 hospitals

VARIABLES	(1) bills	(2) LoS	(3) Number of patients
year	-0.0119	0.0753	0.0245*
	(0.0165)	(0.259)	(0.0148)
after	0.380***	-3.029***	0.0150
	(0.0704)	(1.160)	(0.0640)
After*year	0.0722**	0.0506	-0.0511*
	(0.0304)	(0.507)	(0.0277)
employee number	-0.00241***	-0.0825***	0.00484***
	(0.000321)	(0.00530)	(0.000292)
bed number	0.00426***	0.0884***	-0.00275***
	(0.000263)	(0.00436)	(0.000240)
equipment number	-0.000483**	0.00361	0.000718***
	(0.000232)	(0.00381)	(0.000211)
Constant	5.417***	12.89***	10.97***
	(0.0605)	(0.981)	(0.0547)
Observations	713	712	713
Number of id	372	371	372

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

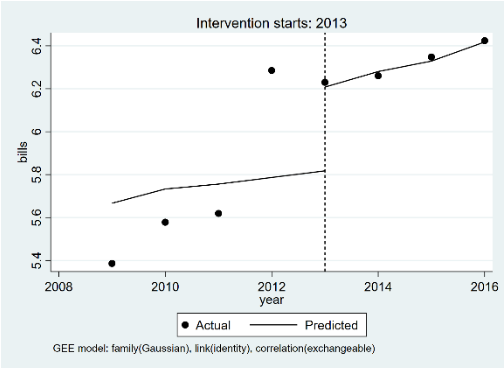


Figure A4-1. The results of bills

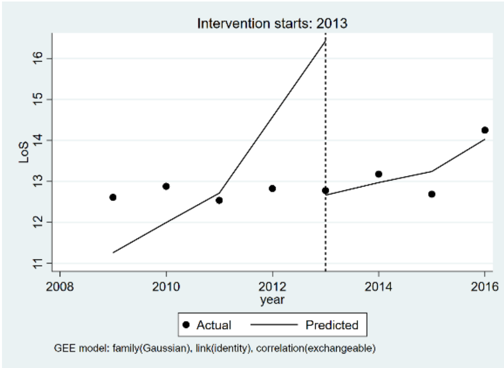


Figure A4-2. The results of LoS

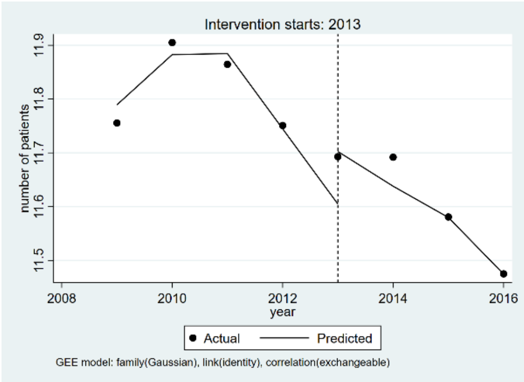


Figure A4-3. The results of number of patients

Figure A 4. The results of ITS estimation for tier 2 hospitals

10.2 Appendix for chapter 6

Table A 5. The results of DiD estimation

VARIABLES	(1) bills	(2) OOPs	(3) los	(4) 30-day readmissions
PPS*post	-0.0993***	-0.100***	0.935***	0.0243***
	(0.0171)	(0.0188)	(0.0133)	(0.00814)
PPS	-0.0905***	-0.0945***	0.892***	-0.0109
	(0.0162)	(0.0178)	(0.0121)	(0.00771)
post	0.303***	0.242***	0.940***	7.86e-05
	(0.0253)	(0.0278)	(0.0202)	(0.0124)
los	0.0533***	0.0522***		0.000381
	(0.000592)	(0.000652)		(0.000283)
age	0.0136***	0.0126***	1.006***	-0.000528
	(0.000891)	(0.000983)	(0.000752)	(0.000428)
agesq	-0.000111***	-0.000127***	1.000	1.06e-05**
	(9.61e-06)	(1.06e-05)	(7.80e-06)	(4.61e-06)
Gender (female=1)	-0.000270	-0.00769	1.025***	0.00256
	(0.00719)	(0.00793)	(0.00615)	(0.00346)
Insurance (URBMI=1)	0.0299***	0.360***	0.936***	6.69e-05
	(0.00771)	(0.00850)	(0.00596)	(0.00371)
Tier (tier 3=1)	0.0842***	0.419***	0.994	-0.0134*
	(0.0164)	(0.0180)	(0.0133)	(0.00788)
Time dummies (by quarter)	controlled	controlled	controlled	controlled
Top 3 digits of ICD codes	controlled	controlled	controlled	controlled
Constant	7.784***	6.807***		0.0130
	(0.0353)	(0.0389)		(0.0169)
Observations	18,438	18,415	18,428	18,071
R-squared	0.467	0.416	0.122	0.026
Number of hospitals	167	167	157	166

Table A 6. The results of DiD+PSM estimation

VARIABLES	(1) bills	(2) OOPs	(3) los	(4) 30-day readmissions
PPS*post	-0.101***	-0.0985***	0.892***	0.0266**
	(0.0220)	(0.0245)	(0.0251)	(0.0106)
PPS	-0.0416**	-0.0396*	0.954*	-0.0164
	(0.0208)	(0.0232)	(0.0255)	(0.01000)
post	0.293***	0.231***	1.040	0.00765
	(0.0334)	(0.0372)	(0.0460)	(0.0165)
los	0.0529***	0.0517***		0.000253
	(0.000781)	(0.000872)		(0.000377)
age	0.0168***	0.0161***	1.007***	-0.00110*
	(0.00120)	(0.00134)	(0.00137)	(0.000581)
agesq	-0.000128***	-0.000143***	1.000	1.48e-05**
	(1.31e-05)	(1.46e-05)	(1.46e-05)	(6.35e-06)
Gender (female=1)	-0.00196	-0.0157	1.000	0.00486
	(0.00961)	(0.0107)	(0.0120)	(0.00466)
Insurance (URBMI=1)	0.0680***	0.386***	1.004	0.00145
	(0.0102)	(0.0113)	(0.0123)	(0.00494)
Tier (tier 3=1)	0.0850***	0.414***	0.980	-0.0236**
	(0.0214)	(0.0239)	(0.0246)	(0.0104)
Time dummies (by quarter)	controlled	controlled	controlled	controlled
Top 3 digits of ICD codes	controlled	controlled	controlled	controlled
Constant	7.634***	6.655***		0.0305
	(0.0461)	(0.0514)		(0.0222)
Observations	11,220	11,201	5,738	11,018
R-squared	0.483	0.423	0.113	0.023
Number of hospitals	162	162	146	162

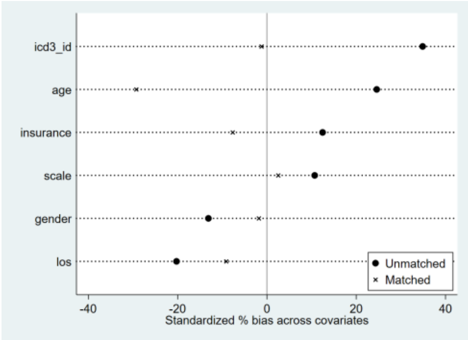


Figure A5-1. Balance test

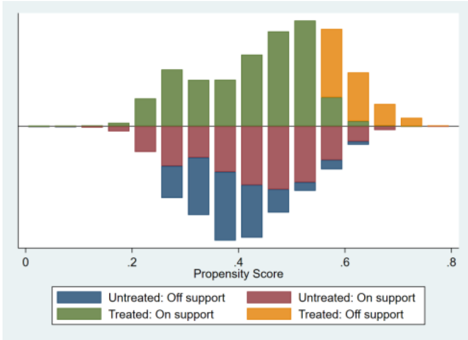


Figure A5-2. Propensity score

Figure A 5. Balance test and propensity score (with LoS in covariates)

Table A 7. Balance test of PSM (with LoS in covariates)

Variable	Unmatched Matched	Mean Treated	Mean Control	t-test t	t-test p>t	V(T)/V(C)
age	U	46.499	42.208	16.61	0.000	1.12*
	M	41.146	46.252	-15.53	0.000	0.75*
gender	U	1.4636	1.529	-8.82	0.000	1.00
	M	1.5226	1.5317	-0.92	0.360	1.00
LoS	U	7.7243	8.9077	-13.59	0.000	0.92*
	M	8.241	8.7743	-4.37	0.000	0.99
insurance	U	1.5723	1.5103	8.39	0.000	0.98
	M	1.5204	1.5586	-3.86	0.000	1.01
tier	U	165.35	160.15	7.19	0.000	0.95*
	M	161.96	160.74	1.26	0.207	0.99
top 3 digit ICD-10 codes	U	6.5829	5.7648	24.09	0.000	2.25*
	M	6.0643	6.0929	-0.55	0.579	2.13*

* if variance ratio outside [0.96; 1.04] for U and [0.95; 1.05] for M

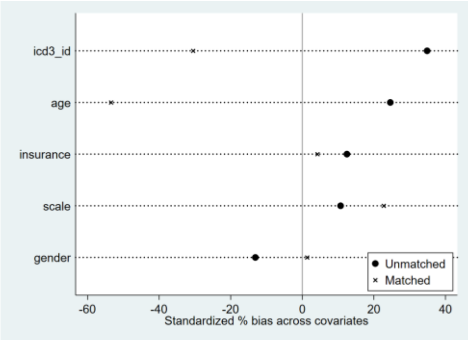


Figure A6-1. Balance test

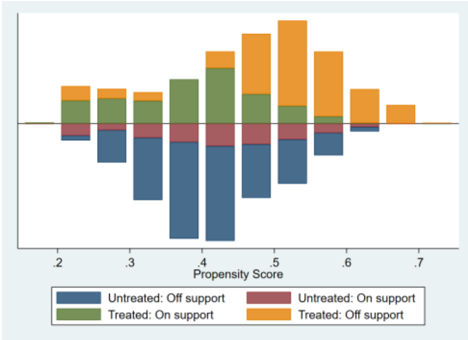


Figure A6-2. Propensity score

Figure A 6. Balance test and propensity score (without LoS in covariates)

Table A 8. Balance test of PSM (without LoS in covariates)

Variable	Unmatched Matched	Mean Treated	Mean Control	t-test t	t-test p>t	V(T)/V(C)
age	U	46.499	42.208	16.61	0.000	1.12*
	M	36.224	45.549	-16.3	0.000	0.55*
gender	U	1.4636	1.529	-8.82	0.000	1.00
	M	1.5815	1.5748	0.4	0.690	1.00
insurance	U	1.5723	1.5103	8.39	0.000	0.98
	M	1.4996	1.4784	1.23	0.219	1.00
tier	U	165.35	160.15	7.19	0.000	0.95*
	M	162.11	151.06	6.56	0.000	0.95
top 3 digit ICD-10 codes	U	6.5829	5.7648	24.09	0.000	2.25*
	M	5.4596	6.1755	-7.49	0.000	1.81*

* if variance ratio outside [0.96; 1.04] for U and [0.95; 1.05] for M

Table A 9. The results of DiD estimation after taking out anticipation effect

VARIABLES	(1) bills	(2) OOPs	(3) los	(4) 30-day readmissions
PPS*post	-0.0627***	-0.0648***	0.988	0.0205**
	(0.0200)	(0.0221)	(0.0168)	(0.00954)
PPS	-0.117***	-0.120***	0.854***	-0.00888
	(0.0192)	(0.0212)	(0.0141)	(0.00915)
post	0.286***	0.226***	0.918***	0.00185
	(0.0257)	(0.0284)	(0.0200)	(0.0126)
los	0.0533***	0.0523***		0.000368
	(0.000592)	(0.000652)		(0.000283)
age	0.0136***	0.0126***	1.006***	-0.000530
	(0.000892)	(0.000983)	(0.000752)	(0.000428)
agesq	-0.000111***	-0.000127***	1.000	1.06e-05**
	(9.62e-06)	(1.06e-05)	(7.80e-06)	(4.61e-06)
Gender (female=1)	0.000162	-0.00726	1.025***	0.00247
	(0.00720)	(0.00793)	(0.00615)	(0.00346)
Insurance (URBMI=1)	0.0296***	0.360***	0.936***	0.000116
	(0.00772)	(0.00850)	(0.00596)	(0.00371)
Tier (tier 3=1)	0.0840***	0.418***	0.992	-0.0135*
	(0.0164)	(0.0180)	(0.0133)	(0.00788)
Time dummies (by quarter)	controlled	controlled	controlled	controlled
Top 3 digits of ICD codes	controlled	controlled	controlled	controlled
Constant	7.796***	6.819***		0.0122
	(0.0357)	(0.0393)		(0.0171)
Observations	18,438	18,415	18,428	18,071
R-squared	0.467	0.416	0.122	0.026
Number of hospitals	167	167	157	166

Table A 10. The results of DiD estimation for tier 3 hospitals

VARIABLES	(1) bills	(2) OOPs	(3) los	(4) 30-day readmissions
PPS*post	-0.103***	-0.124***	0.922***	0.0278
	(0.0388)	(0.0415)	(0.0258)	(0.0175)
PPS	0.226***	0.258***	0.879***	0.00847
	(0.0693)	(0.0742)	(0.0241)	(0.0316)
post	-0.110***	-0.122***	0.862***	-0.0216
	(0.0379)	(0.0406)	(0.0446)	(0.0171)
los	0.0445***	0.0462***		0.000810**
	(0.000904)	(0.000968)		(0.000409)
age	0.0159***	0.0137***	1.004***	-0.000544
	(0.00149)	(0.00159)	(0.00111)	(0.000676)
agesq	-0.000126***	-0.000130***	1.000***	1.32e-05*
	(1.61e-05)	(1.73e-05)	(1.16e-05)	(7.33e-06)
Gender (female=1)	-0.00324	-0.0146	1.013	-0.00894
	(0.0128)	(0.0137)	(0.00961)	(0.00583)
Insurance (URBMI=1)	0.0344**	0.425***	0.951***	0.000920
	(0.0136)	(0.0145)	(0.00972)	(0.00621)
Time dummies (by quarter)	controlled	controlled	controlled	controlled
Top 3 digits of ICD codes	controlled	controlled	controlled	controlled
Constant	8.142***	7.389***		0.0130
	(0.0812)	(0.0870)		(0.0367)
Observations	7,177	7,166	7,175	7,023
R-squared	0.418	0.384	0.133	0.036
Number of hospitals	42	42	40	42

Table A 11. The results of DiD estimation for tier 2 hospitals

VARIABLES	(1) bills	(2) OOPs	(3) los	(4) 30-day readmissions
PPS*post	-0.0962***	-0.0750***	0.937***	0.0287***
	(0.0185)	(0.0208)	(0.0158)	(0.00938)
PPS	0.290***	0.178***	0.893***	-0.00938
	(0.0273)	(0.0306)	(0.0142)	(0.0144)
post	-0.0854***	-0.0937***	0.979	-0.00882
	(0.0172)	(0.0193)	(0.0247)	(0.00870)
los	0.0610***	0.0567***		-5.76e-05
	(0.000783)	(0.000878)		(0.000397)
age	0.0115***	0.0122***	1.008***	-0.000527
	(0.00110)	(0.00123)	(0.00104)	(0.000559)
agesq	-9.51e-05***	-0.000129***	1.000***	9.00e-06
	(1.18e-05)	(1.32e-05)	(1.07e-05)	(6.01e-06)
Gender (female=1)	0.000307	-0.00522	1.033***	0.00924**
	(0.00841)	(0.00943)	(0.00802)	(0.00429)
Insurance (URBMI=1)	0.0324***	0.317***	0.920***	-0.000260
	(0.00919)	(0.0103)	(0.00756)	(0.00469)
Time dummies (by quarter)	controlled	controlled	controlled	controlled
Top 3 digits of ICD codes	controlled	controlled	controlled	controlled
Constant	7.659***	6.760***		0.00784
	(0.0403)	(0.0452)		(0.0204)
Observations	11,261	11,249	11,252	11,048
R-squared	0.519	0.449	0.118	0.023
Number of hospitals	146	146	137	145

10.3 Appendix for chapter 7

Table A 12. DDD results for overall analysis

	(1)	(2)
VARIABLES	bills	LoS
hospital*disease*post	-0.157***	1.093***
	(0.0525)	(0.0303)
hospital*post	0.0297	0.973
	(0.0440)	(0.0218)
disease*post	-0.0351	1.075***
	(0.0250)	(0.0146)
disease*hospital	0.0425	0.751***
	(0.0293)	(0.0116)
hospital	-0.0189	1.266***
	(0.0677)	(0.0443)
disease	-0.0279*	0.924***
	(0.0154)	(0.00797)
post	0.309***	0.914***
	(0.0312)	(0.0154)
LoS	0.0296***	
	(0.000456)	
age	0.0276***	0.959***
	(0.00506)	(0.00254)
age_sq	-0.000181***	1.000***
	(3.44e-05)	(1.81e-05)
gender (female=0)	0.00833	0.926***
	(0.00946)	(0.00506)
tier (tier 2=0)	0.238***	0.949*
	(0.0535)	(0.0286)
		1.093***
frail degree	controlled	controlled
time dummies (by half year)	controlled	controlled
Constant	7.016***	
	(0.188)	
Observations	11,815	11,711
R-squared	0.290	0.326
Number of hospitals	603	499

Table A 13. DDD results for cataract analysis and HIVD analysis

Diseases	cataract		HIVD	
	(1)	(2)	(3)	(4)
VARIABLES	bills	los	bills	los
hospital*disease*post	-0.0565	1.032	-0.198***	1.028
	(0.109)	(0.0759)	(0.0629)	(0.0325)
hospital*post	-0.197**	0.691***	0.187***	1.031
	(0.0980)	(0.0391)	(0.0499)	(0.0267)
disease*post	-0.0839	0.897**	-0.0214	1.068***
	(0.0790)	(0.0482)	(0.0246)	(0.0151)
disease*hospital	-0.119*	0.873***	-0.0685*	1.022
	(0.0621)	(0.0355)	(0.0367)	(0.0186)
hospital	-0.122	1.645***	0.0116	1.040
	(0.156)	(0.157)	(0.0723)	(0.0402)
disease	0.480***	0.522***	-0.112***	1.006
	(0.0461)	(0.0166)	(0.0152)	(0.00911)
post	0.426***	1.038	0.273***	0.900***
	(0.0879)	(0.0602)	(0.0318)	(0.0161)
LoS	0.0232***		0.0313***	
	(0.00114)		(0.000505)	
age	0.0307***	0.951***	0.00601	0.980***
	(0.0102)	(0.00589)	(0.00552)	(0.00299)
age_sq	-0.000204***	1.000***	-5.18e-05	1.000***
	(6.89e-05)	(4.37e-05)	(3.76e-05)	(2.07e-05)
gender (female=0)	0.0181	0.964**	0.00378	0.940***
	(0.0185)	(0.0146)	(0.0103)	(0.00557)
tier (tier 2=0)	0.288***	1.265***	0.240***	0.903***
	(0.105)	(0.0954)	(0.0603)	(0.0308)
frail degree	controlled	controlled	controlled	controlled
time dummies (by half year)	controlled	controlled	controlled	controlled
Constant	6.903***		7.818***	
	(0.385)		(0.205)	
Observations	3,157	3,141	8,658	8,555
R-squared	0.166	0.320	0.342	0.308
Number of hospitals	100	84	583	480

Table A 14. DiD results for cataract analysis on patients from targeted hospitals

	(1)	(2)
VARIABLES	bills	LoS
disease*post	-0.167**	0.956
	(0.0804)	(0.0486)
disease	0.341***	0.466***
	(0.0486)	(0.0127)
post	0.330***	0.663***
	(0.107)	(0.0449)
LoS	0.0241***	
	(0.00148)	
age	0.0649***	0.960***
	(0.0139)	(0.00732)
age_sq	-0.000439***	1.000**
	(9.53e-05)	(5.51e-05)
gender (female=0)	0.0307	0.977
	(0.0302)	(0.0206)
frail degree	controlled	controlled
time dummies (by half year)	controlled	controlled
Constant	6.146***	
	(0.506)	
Observations	1,369	1,368
R-squared	0.203	0.289
Number of hospitals	16	15

Table A 15. PPS tariffs for 101 diagnostic groups of 44 diseases

No.	Disease	Main operations	Anesthesia method	Quantity	PPS tariff
1	Nodular goiter	Total thyroidectomy	General Anesthesia		10380
2	Nodular goiter	Subtotal thyroidectomy	Nerve block anesthesia	Unilateral	9370
3	Nodular goiter	Subtotal thyroidectomy	General Anesthesia	Bilateral	10840
4	Nodular goiter	Partial thyroidectomy	Nerve block anesthesia	Unilateral	9370
5	Nodular goiter	Partial thyroidectomy	General Anesthesia	Bilateral	10840
6	Benign thyroid tumors	Partial thyroidectomy	Nerve block anesthesia	Unilateral	9590
7	Benign thyroid tumors	Partial thyroidectomy	General Anesthesia	Bilateral	10840
8	Benign thyroid tumors	Subtotal thyroidectomy	Nerve block anesthesia	Unilateral	9370
9	Benign thyroid tumors	Subtotal thyroidectomy	General Anesthesia	Bilateral	10840
10	Thyroid Tumors	Removal of thyroid adenoma	Nerve block anesthesia	Unilateral	11430
11	Thyroid Tumors	Removal of thyroid adenoma	General Anesthesia	Bilateral	12160
12	Primary acute angle-closure glaucoma	External trabeculotomy with trabeculectomy	Local anesthesia	Single Eye	4880
13	Primary acute angle-closure glaucoma	External trabeculotomy with trabeculectomy	Local anesthesia	Both eyes	6600
14	Age-related cataracts	Cataract extracapsular extraction + IOL implantation	Local anesthesia	Single Eye	4710
15	Age-related cataracts	Cataract extracapsular extraction + IOL implantation	Local anesthesia	Both eyes	7610
16	Age-related cataracts	Cataract ultrasound emulsion aspiration + IOL implantation	Local anesthesia	Single Eye	7380
17	Age-related cataracts	Cataract ultrasound emulsion aspiration + IOL implantation	Local anesthesia	Both eyes	10120
18	Age-related cataracts	Cataract Ultrasonic Emulsion Extraction	Local anesthesia	Single Eye	4910

19	Age-related cataracts	Cataract Ultrasonic Emulsion Extraction	Local anesthesia	Both eyes	4910
20	Cataracts	Cataract extracapsular extraction + IOL implantation	Local anesthesia	Single Eye	4710
21	Cataracts	Cataract extracapsular extraction + IOL implantation	Local anesthesia	Both eyes	7610
22	Cataracts	Cataract ultrasonic phacoemulsification extraction + IOL implantation	Local anesthesia	Single Eye	7380
23	Cataracts	Cataract ultrasonic phacoemulsification extraction + IOL implantation	Local anesthesia	Both eyes	10120
24	pterygium	pterygium excision + corneal transplantation	Local anesthesia	Single Eye	2050
25	pterygium	pterygium excision + corneal transplantation	Local anesthesia	Both eyes	2140
26	Chronic dacryocystitis	Nasal lacrimal sac anastomosis	Local anesthesia	Bilateral	6280
27	Chronic suppurative otitis media	Tympanoplasty (including auditory chain reconstruction, tympanic membrane repair, and lesion exploration surgery; including types 1-5)	Local anesthesia	Bilateral	11030
28	Chronic tonsillitis	Tonsillectomy	General Anesthesia	Bilateral	7940
29	Chronic sinusitis	Transnasal endoscopic sinus surgery (including frontal sinus, septal sinus, and pterygoid sinus)	Local anesthesia	1 sinus	7140
30	Chronic sinusitis	Transnasal endoscopic sinus surgery (including frontal sinus, septal sinus, and pterygoid sinus)	Local anesthesia	More than 2 sinuses	9280
31	Maxillary cysts	Removal of jaw bone cysts	General Anesthesia		7900
32	Benign lung tumors	Lung wedge resection	General Anesthesia	Unilateral	30830
33	Benign lung tumors	Lung wedge resection	General Anesthesia	Bilateral	32180

34	Benign lung tumors	Transthoracoscopic lung wedge resection	General Anesthesia	Unilateral	32500
35	Benign lung tumors	Transthoracoscopic lung wedge resection	General Anesthesia	Bilateral	36670
36	Spontaneous pneumothorax	Transthoracoscopic pleural fixation	General Anesthesia		28170
37	Acute suppurative appendicitis	Trans-laparoscopic appendectomy	General Anesthesia		6480
38	Acute suppurative appendicitis	Appendectomy	Continuous epidural block anesthesia		6480
39	Acute simple appendicitis	Appendectomy	Lumbar anesthesia with continuous epidural block		5740
40	Acute simple appendicitis	Trans-laparoscopic appendectomy	General Anesthesia		5740
41	Anal fissure	Surgical treatment of common perianal diseases	Continuous epidural block anesthesia		5930
42	Colonic polyps	Special treatment via colonoscopy (microwave therapy)	General anesthesia without intubation		3310
43	Colonic polyps	Special treatment via colonoscopy (laser, electrodesiccation)	General anesthesia without intubation		3310
44	Ascending colonic polyps	Special treatment via colonoscopy (microwave therapy)	General anesthesia without intubation		3310
45	Ascending colonic polyps	Special treatment via colonoscopy (laser, electrodesiccation)	General anesthesia without intubation		3310
46	Descending colon polyps	Special treatment via colonoscopy (microwave therapy)	General anesthesia without intubation		3310
47	Descending colon polyps	Special treatment via colonoscopy (laser, electrodesiccation)	General anesthesia without		3310

			intubation		
48	Sigmoid colon polyp	Special treatment via colonoscopy (microwave therapy)	General anesthesia without intubation		3310
49	Sigmoid colon polyp	Special treatment via colonoscopy (laser, electrodesiccation)	General anesthesia without intubation		3310
50	Gallbladder polyps	Cholecystectomy	Continuous epidural block anesthesia		6160
51	Unilateral or unspecified inguinal hernia without obstruction or gangrene	Inguinal hernia repair	Nerve block anesthesia	Unilateral	7600
52	Unilateral or unspecified inguinal hernia without obstruction or gangrene	Tension-free inguinal hernia repair	Nerve block anesthesia	Unilateral	7600
53	Inguinal hernia	Inguinal hernia repair	Nerve block anesthesia	Unilateral	7600
54	Inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Unilateral	7600
55	Direct inguinal hernia	Inguinal hernia repair	Nerve block anesthesia	Unilateral	7600
56	Direct inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Unilateral	7600
57	Recurrent inguinal hernia	Inguinal hernia repair	Nerve block anesthesia	Unilateral	8710
58	Recurrent inguinal hernia	Inguinal hernia repair	Nerve block anesthesia	Bilateral	12550
59	Recurrent inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Unilateral	8710
60	Recurrent inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Bilateral	12550
61	Recurrent direct inguinal hernia	Inguinal hernia repair	Nerve block anesthesia	Unilateral	8710
62	Recurrent direct inguinal hernia	Inguinal hernia repair	Nerve block anesthesia	Bilateral	12550
63	Recurrent direct inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Unilateral	8710
64	Recurrent direct inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Bilateral	12550
65	Recurrent inguinal	Inguinal hernia repair	Nerve block	Unilate	8710

	hernia		anesthesia	ral	
66	Recurrent inguinal hernia	Inguinal hernia repair	Nerve block anesthesia	Bilateral	12550
67	Recurrent inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Unilateral	8710
68	Recurrent inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Bilateral	12550
69	Bilateral inguinal hernia	Inguinal hernia repair	Nerve block anesthesia	Bilateral	11440
70	Bilateral inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Bilateral	11440
71	Bilateral direct inguinal hernia	Inguinal hernia repair	Nerve block anesthesia	Bilateral	11440
72	Bilateral direct inguinal hernia	Tension-free inguinal hernia repair	Nerve block anesthesia	Bilateral	11440
73	Bilateral inguinal hernia without obstruction or gangrene	Tension-free inguinal hernia repair	Nerve block anesthesia	Bilateral	11440
74	Bilateral inguinal hernia (one side direct hernia, side hernia)	Tension-free inguinal hernia repair	Nerve block anesthesia	Bilateral	11440
75	Varicose veins of the lower extremities	Saphenous vein high ligation with ten stripping	Continuous epidural block anesthesia	Unilateral	5630
76	Varicose veins of the lower extremities	Saphenous vein high ligation + stripping	Continuous epidural block anesthesia	Bilateral	7620
77	Thrombosed external hemorrhoids	Thrombosed external hemorrhoidectomy	Local anesthesia		3400
78	Chronic cholecystitis or combined gallbladder stones	Trans-laparoscopic cholecystectomy	Continuous epidural block anesthesia		9730
79	Benign prostatic hyperplasia	Transurethral electrodesiccation of the prostate	Continuous epidural block anesthesia		11330
80	Ureteral calculus	Transurethral ureteroscopic ultrasonic lithotripsy for lithotripsy	Lumbar anesthesia with continuous epidural block		6320
81	Ureteral calculus	Transurethral ureteroscopic laser lithotripsy for lithotripsy	Lumbar anesthesia with continuous		10810

			epidural block		
82	Ureteral calculus	Transurethral ureteroscopic pneumatic ballast lithotripsy for lithotripsy	Continuous epidural block anesthesia		7290
83	Testicular syringomyelia	Traffic syringomyelia repair	Continuous epidural block anesthesia		6510
84	Testicular syringomyelia	Testicular Sphincter Reversal	Continuous epidural block anesthesia		5030
85	Uterine smooth muscle tumor	Laparoscopic combined with negative total hysterectomy	General Anesthesia		11770
86	Uterine smooth muscle tumor	Transabdominal total hysterectomy	Lumbar anesthesia with continuous epidural block		11440
87	Uterine smooth muscle tumor	Transvaginal total hysterectomy	Lumbar anesthesia with continuous epidural block		9670
88	Uterine smooth muscle tumor	Subtotal transabdominal hysterectomy	Lumbar anesthesia with continuous epidural block		9670
89	Uterine smooth muscle tumor	Transhysteroscopic submucosal myomectomy	Lumbar anesthesia with continuous epidural block		7430
90	Benign Ovarian Tumor	Transabdominal unilateral ovarian cyst removal	Lumbar anesthesia with continuous epidural block		9920
91	Lumbar disc herniation	Percutaneous Laser Lumbar Disc Removal	General Anesthesia	1 intervertebral disc	12520
92	Rouge cyst	Carotid cyst excision	Nerve block anesthesia	Unilateral	7000
93	Rouge cyst	Carotid cyst excision	Nerve block	Bilater	9120

			anesthesia	al	
94	Removal of internal fracture fixation device	Removal of internal fixation device for fracture	Local anesthesia	Removal of an internal fixation device	4050
95	Removal of internal fracture fixation device	Removal of internal fixation device for fracture	Local anesthesia	Removal of more than one internal fixation device	7310
96	Osteoarthritis of the knee joint	Arthroscopic knee debridement	Continuous epidural block anesthesia	Bilateral	14050
97	Fracture of the lower end of the radius (Collet's fracture)	Fracture manipulation and reconstruction	General Anesthesia	Unilateral	3790
98	Acute mastitis	Superficial breast abscess incision and drainage	Nerve block anesthesia	Unilateral	4070
99	Acute mastitis	Deep breast abscess incision and drainage	Nerve block anesthesia	Unilateral	4070
100	Benign skin tumors of the breast	Mastectomy of breast masses	Nerve block anesthesia	Unilateral	4860
101	Benign skin tumors of the breast	Mastectomy of breast masses	Nerve block anesthesia	Bilateral	5990