Policy Evaluation on the Integration of Urban and Rural Residents Basic Medical Insurance in China

Liran Li

Thesis submitted for

The Degree of Doctor of Philosophy

Health Econometrics

School of Health and Related Research

Faculty of Medicine, Dentistry and Health

The University of Sheffield

July 2023

Abstract

Background: China's health system has been undergoing a series of major reforms, covering various areas such as drug pricing, hospital management, and health insurance. The health insurance system, which existed in three distinct forms for employees, urban residents, and rural residents, underwent a series of crucial reforms at the turn of the 21st century. The recent advancement in these reforms involved the integration of basic medical insurance for urban and rural residents. This integration process began with extensive pilots and was completed in 2020 after a nationwide rollout in 2016. Previous research has primarily examined the impact of insurance enrolment, including improved patient accessibility of healthcare, reduced out-of-pocket (OOP) expenses for patients, and equitable care provision among diverse insured groups. However, these studies were either limited to a specific type of insurance or involved comparisons among different insurance types. There were only a few studies available that focused on integrated insurance in pilot areas. To the best of my knowledge, there is a lack of comprehensive research on the entire integration process and its policy effects on health service use and/or health expenses.

Objective: This research aims to understand China's urban-rural health insurance integration policies and to evaluate their effects. Specifically, this comprehensive evaluation of the policies includes:

1. To summarise the different regional integration policies and to discuss the potential benefits and problems of integration;
2. To estimate the effect of integration on health service use and costs, at the national level;
3. To explore the differences in policy effects among different populations in socio-economic subgroups, at the national level;
4. To assess the impact of different integration policies on health service use and costs at the city level, using ischaemic heart disease as an example;
5. To investigate the mechanisms of action of different integration policies at the city level and how they differ, using ischaemic heart disease as an example.

Method: In this thesis, a mixed methods approach is employed, combining both qualitative and quantitative methods. The qualitative part involves document analysis, which focuses on the first objective of the thesis. Policy documents are gathered from public government sources and websites, and a summary and comparison of these documents are conducted. Specifically, the timing of policy implementation is summarised for 338 cities in China, along with the timeline of implementation for 25 provincial capital regions. Furthermore, the insurance terms for urban and rural residents before and after the integration are detailed, summarised, and compared for four selected cities (Beijing, Tianjin, Chengdu, and Shanghai). These terms include premiums, deductibles, reimbursement rates, and ceilings for outpatient and hospitalisation reimbursements. The first quantitative analysis part focuses on the second and third objectives of this thesis. Data from the China Health and Retirement Longitudinal Study (CHARLS), a national survey of middle-aged and elderly individuals, are used to analyse the nationwide average policy effects and to examine population heterogeneity based on socio-economic factors. The staggered Difference-in-Difference (DID) method is employed as the main analytical approach, while the Difference-in-Difference-in-Difference (DDD) method is used as a supplemental analysis approach. Subsequently, the second quantitative analysis part addresses the fourth and fifth objectives of this thesis. Data from Electronic Medical Records (EMR) of Ischemic Heart Disease (IHD) patients are used to examine variations in policy effects across cities. The DID method is employed as the main analytical approach, and Propensity Score Matching (PSM) combined with DID is used as a sensitivity analysis approach. Additionally, analyses of dynamic effects and quantile effects are conducted to explore the mechanisms of policy action at the city level.

Findings: The qualitative analysis shed light on the implementation process of the integrated urban and rural resident health insurance scheme. It was found that the integrated insurance was initially introduced in 2007 and subsequently underwent a gradual pilot phase across the country until 2016. Following this pilot phase, more than half of the cities completed their integration in 2017 and 2018. Finally, by the end of 2020, the integration was fully implemented in all cities. The varying timing of integration among cities is likely to be a significant factor contributing to the variations in integration terms and potential effects.

In the four case cities examined, findings indicated that the insurance premiums increased rapidly, which might weaken the willingness of rural residents to enrol in the insurance and might decrease their satisfaction with it. One possible reason is that on average, the difference in expenses between insured individuals (insurance fee) and uninsured individuals (treatment fee) is not great. Moreover, there are various limitations on the actual reimbursement of insurance, such as deductibles, ceilings, and whether certain drugs used are included in the reimbursable list. As a result, many insured residents may experience a disparity between their initial expectations and the actual reimbursement amount when making claims. The changes in outpatient reimbursement terms are more consistent across cities, which are likely to encourage more visits and reduce costs. Another notable benefit of integration is the substantial increase in the number of institutions and reimbursable drugs available to patients. However, changes in outpatient reimbursement terms also indicate that integration may be more advantageous for urban than rural residents. Regarding hospitalisation reimbursement, there are fewer commonalities in the integration terms across cities. Therefore, differences in terms across cities, between urban and rural, as well as variations in health status and economic conditions among different population groups, may all contribute to insignificant integration effects on hospitalisation at the national level.

In the survey-based quantitative analysis, the results demonstrate that integration has a positive influence on the probability of outpatient visits and reduces the OOP costs of individuals in both outpatient visits and hospitalisation at the national level. However, it is notable that the integration policy effects vary significantly among different socioeconomic subgroups. Specifically, rural residents experience greater benefits compared to urban residents, and residents with higher incomes benefit more than those with lower incomes. Furthermore, in the EMR-based quantitative analysis, the results reaffirm the conclusions drawn from the qualitative analysis. It is that integration policies in various cities similarly affect outpatient fees. However, substantial differences merge when considering the impact on hospitalisation services and fees. In the subsequent two deeper analyses, the results of dynamic effect analysis point to three different patterns: policy effect delay, reversal, and recession. These effects may be a result of various interactions among hospital administrators, doctors, and patients. When

examining the effects across patients with different quantiles of total fees, the results show that the policy effects grow stronger as patients' pre-integration total fees (which reflect the severity of their disease) increase. This finding supports the notion that integration helps address the unmet medical needs of patients who face financial constraints. Additionally, the results also suggest that there may be cases where some cities' policies have increased the deductible and reduced treatment for patients with milder conditions.

Conclusion: This thesis provided a comprehensive evaluation of China's latest completed reform, the integration of urban and rural basic medical insurance, by combining quantitative and qualitative analyses. The thesis highlights five key findings:

1. The integration policies and effects of outpatients were similar across cities, but the integration policies and effects of hospitalisation were even more different;
2. Nationally, the integration policies promoted outpatient visits and reduced outpatient and hospitalisation OOP costs;
3. Urban-rural healthcare inequality was reduced through integration, but inequality among different income groups was expanded;
4. Differences in the impact of integration in different cities on hospitalisation may be attributed to differences in doctor behaviours, reimbursement terms, and reimbursement catalogues;
5. For hospitalisation, integration effectively released patients' unmet medical needs constrained by economic factors.

This research offers several advantages compared to previous studies, such as the use of mixed methods and more advanced and comprehensive causal identification methods and analysis frameworks. However, certain limitations exist, such as the omission of considering the medical supply side and the interaction of multiple policies, and the inherent defects of data types. Future research endeavours could concentrate on expanding data sources, improving causal identification methods, and examining the interaction between different policies.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ATE	Average Treatment Effect
ATT	Average Treatment Effect on Treated
BLUE	Best Linear Unbiased Estimator
CHARLS	China Health and Retirement Longitudinal Study
CI	Concentration Index
CIA	Conditional Independence Assumption
CIC	Change-in-Change
DDD	Difference-in-Difference-in-Difference
DID	Difference-in-Difference
DRG	Diagnosis Related Group
EMR	Electronic Medical Record
FE	Fixed-Effect
GIS	Government Insurance Scheme
GLM	Generalized Linear Model
HTA	Health Technology Assessment
IV	Instrumental Variables
LATE	Local Average Treatment Effect
LIS	Labour Insurance Scheme
NRCMS	New Rural Cooperative Medical Scheme
OLS	Ordinary Least Square
OOP	Out-of-Pocket
PRC	People's Republic of China
PSM	Propensity Score Matching
QDID	Quantile Difference-in-Difference
(R)CMS	(Rural) Cooperative Medical Scheme/System
RCT	Randomised Controlled Trial
RD	Regression Discontinuity
RR	Reimbursement Rate
SARS	Severe Acute Respiratory Syndrome
SCM	Synthetic Control Method
SHI	Social Health Insurance
SUTVA	Stable Unit Treatment Value Assumption
UEBMI	Urban Employee Basic Medical Insurance
UHC	Universal Health Coverage
URBMI	Urban Resident Basic Medical Insurance
URRBMI	Urban and Rural Resident Basic Medical Insurance
WHO	World Health Organization

Acknowledgement

First, I would like to thank my supervisors, Prof Mónica Hernández Alava, Prof Julie Balen, and Dr Hannah Jordan, for their invaluable guidance and unwavering support throughout these years of research. In addition, I am also thankful to my friends at the University of Sheffield, who have been a source of constant support and joy along the way. In particular, I am very grateful to Lichang Liu, Qian Zhao, and Jiafeng Xiong for their companionship and shared laughter during the ups and downs of this journey of life. Finally, my parents, they are the most important. Their unwavering support has been the cornerstone of my journey. They have provided me with not just financial assistance, but also emotional and intellectual support. Their trust in me has allowed me to live a life full of freedom and pursue my passions without worry. I am forever grateful for their love and encouragement.

Liran Li
July 2023

Declaration

I declare that this thesis is a presentation of my original work, and I am the solo author. This work has not previously been presented for an award at this, or any other, University, Information included from other sources (e.g. journal articles and books) has been fully acknowledged and correctly referenced.

Liran Li

July 2023

Chapter 1 Introduction

This introductory chapter starts with a brief description to China's health insurance system. Then, previous studies on health insurance in China are briefly reviewed and research gaps are highlighted. Finally, the research aims and objectives are stated, and the framework of this thesis is given.

1.1 Background

China's healthcare system has garnered global scholarly attention due to its great achievement. Despite being classified as a low and middle-income country by the World Bank, China has nearly achieved Universal Health Coverage (UHC) (Yu, 2015). UHC refers to "all people have access to the full range of quality health services they need, when and where they need them, without financial hardship"[1]. To achieve UHC, the Chinese government has carried out considerable reforms to the health insurance system in the past few decades, aligning with the nation's economic and societal developments.

In the early 21st century, China's health insurance landscape transformed into three distinct schemes based on employment and geographical disparities. The most recent reform is the integration of urban and rural resident basic medical insurance. Evaluating the extent to which and how this reform has accomplished its intended objectives is an imperative concern for policymakers and researchers alike. This thesis, therefore, aims to examine the impact of integration reform on residents' healthcare services use and associated expenses, addressing both the existence of an effect and the underlying mechanisms at play.

Following the establishment of the People's Republic of China (PRC) in 1949, a comprehensive healthcare system has been implemented, significantly improving the health of the nation. Yet,

[1] https://www.who.int/health-topics/universal-health-coverage#tab=tab_1

the health security system[2] differed between urban and rural areas (World Bank, 1997). In urban areas, two employment-based health insurance schemes, the Labour Insurance Scheme (LIS) and the Government Insurance Scheme (GIS), were implemented, covering almost all urban residents (World Bank, 1997). In rural areas, a rural cooperative medical system (RCMS) was implemented, earning praise from the World Bank and WHO (Xingzhu and Huaijie, 1992). It was a notable model among developing countries in addressing health funding challenges and enhancing overall healthcare efficiency (Caldwell, 1986).

In 1978, China began its economic reform, namely, the "Reform and Opening Up". It brought about tremendous changes in China's social economy. Consequently, the old health security system, reliant on the publicly-owned economy, gradually collapsed. By 2000, China's healthcare sector was regarded as one of the most inequitable in the world, with the equity of individual health expenses being ranked fourth from last (WHO, 2000). To address these problems, China carried out healthcare reforms in the 1990s and finally formed a new health insurance system[3] in the early 2000s (Figure 1-1). By 2012, over 1.3 billion people were enrolled in the three basic medical insurance schemes, achieving a health coverage of 95% (The State Council of China, 2012).

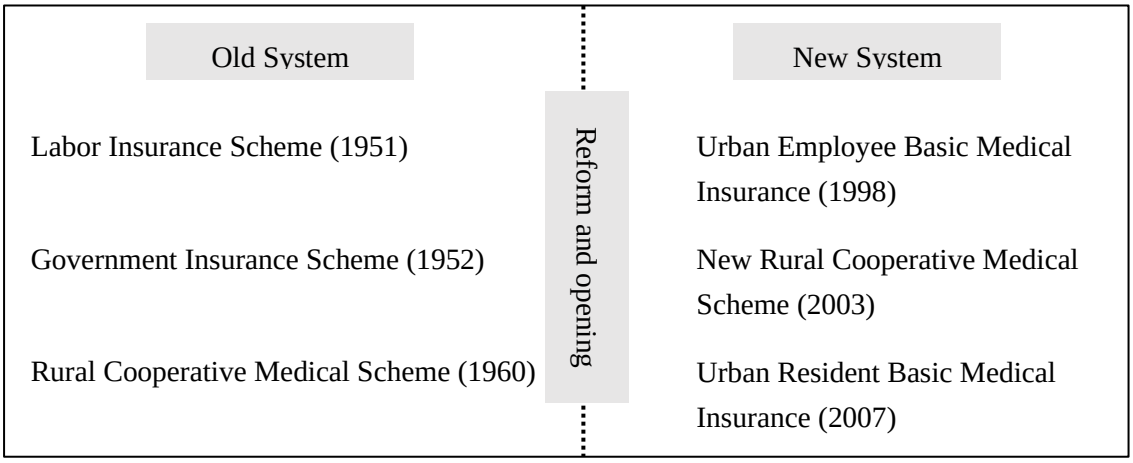

Figure 1-1 China's health insurance reform at the end of the 20th century

[2] The definitions and distinctions among the healthcare system, health security system, and health insurance system in China are explained in detail in Appendix A.

[3] Insurance schemes in the old and new systems are not one-to-one matched. More details about the changes in insurance schemes are given in Chapter 2.

Despite the formation of the new health insurance system, urban-rural inequity remained prominent in China during the early 2000s. Due to the dual social system of urban-rural separation[4], the new health insurance system adopted three schemes for varying populations: urban employees, urban unemployed residents, and rural residents. These schemes exhibited great variations in payment standards and reimbursement scope. Notably, rural residents, facing lower quality and less access to primary healthcare, as well as inferior living conditions, exhibit higher morbidity rate and lower health status compared to urban residents (Zhang, 2003; Xie, 2009; Xi *et al.*, 2014). Despite a higher expected need for health services among rural residents, their actual use remains lower, possibly due to financial constraints arising from lower average disposable incomes. Moreover, due to the lower incomes, the risk of health service expenses consuming most household income, pushing impoverished households deeper into poverty, is also greater in rural areas (Gao *et al.*, 2002; Zhang, 2003).

Another issue was duplicate insurance. Originally, the three insurance schemes were designed to cover different populations so that everyone could be protected by insurance. However, different institutions independently managed these three health insurance schemes. This management structure allowed individuals, especially migrants, to enrol in and claim benefits from more than one insurance scheme concurrently. The situation of duplicate insurance was very common and led to a serious consequence: duplicate claims wasted huge financial subsidies (Changxue *et al.*, 2016).

In response to these challenges, in 2009, the Chinese government introduced a further reform, the "New Healthcare Reform" (The State Council of China, 2009). An important part of this reform was the exploration of a unified urban-rural resident health insurance system. The integration of urban-rural resident health insurance was implemented at the city level because rural areas were administratively governed by urban institutions. Pilot projects for health insurance integration began in some areas in 2009, with subsequent expansions based on local

[4] China's registered permanent residence system (Hukou) differentiates the agricultural and the non-agricultural registered permanent residence strictly. Education, employment, and social security are linked to hukou, and the two types of hukou are treated differently.

conditions. By 2016, approximately one-fifth of cities had successfully integrated urban and rural health insurance systems.

Based on pilot experiences in different cities, a government-funded study (Changxue *et al.*, 2016) provided a framework for integration (Figure 1-2). On 12 January 2016, the State Council in China issued: "The Opinions of the State Council on the Integration of the Basic Medical Insurance System for Urban and Rural Residents" (The State Council of China, 2016) (the full text of the Chinese version of this document is provided in Appendix B). This document outlined official advice and requirements for the implementation of the Urban and Rural Resident Basic Medical Insurance (URRBMI). Subsequently, the integration commenced nationwide. By the beginning of 2018, 23 provinces (of 31 in mainland China) had completed their integrations of basic medical insurance for urban and rural residents[5]. By the end of 2020, all regions successfully concluded their integration processes.

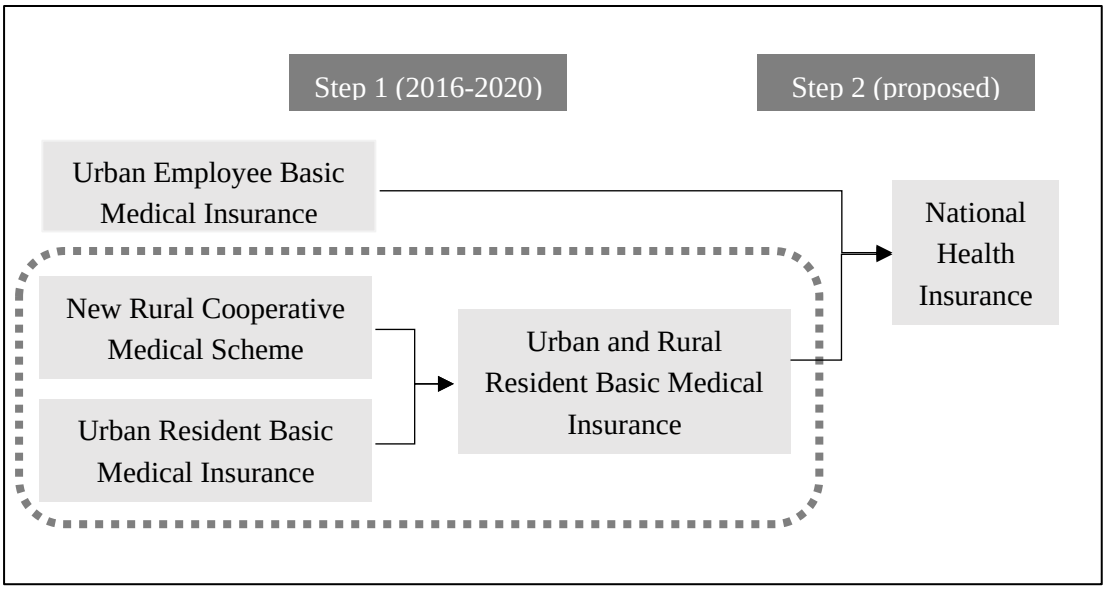

Figure 1-2 Further reform of health insurance in China

[5] Since the integration is at the city level, not the province level, the completion of a province means insurances in all cities in that province are integrated, but it does not mean that there is only one insurance scheme per province. For those that were ongoing, insurances in some cities in that province have not yet been integrated.

1.2 Previous Studies of Health Insurance in China

There are many studies examining social health insurance in China, involving the effect on the accessibility of health services and the financial protection of patients (Wagstaff and Lindelow, 2008; Lei and Lin, 2009; Wagstaff *et al.*, 2009; Yip and Hsiao, 2009; Meng *et al.*, 2012; Long *et al.*, 2013). However, due to the use of different analytical methods, selection of various outcome indicators, and a variety of study populations, results are inconsistent and policy conclusions are, therefore, difficult to draw. In addition, there are also several studies focusing on equity (Ma, Song and Gu, 2016; Su *et al.*, 2018). However, most of the above studies on health insurance looked at one type of insurance or compared different types of insurance, such as employee insurance versus resident insurance, or social insurance versus commercial insurance. These studies cannot provide any causal explanation for insurance integration. More recently, studies compared the effects of integrated resident insurance with other insurance schemes, and some have assessed the policy effect of integration in pilot areas (Ma, Zhao and Gu, 2016; Xie, Hu and Chen, 2017). However, to the best of my knowledge, no studies have taken into account the actual policy effect of the comprehensive implementation of integration on health services use and health expenditure, and the regional and population heterogeneity of the effects of integration (Figure 1-3).

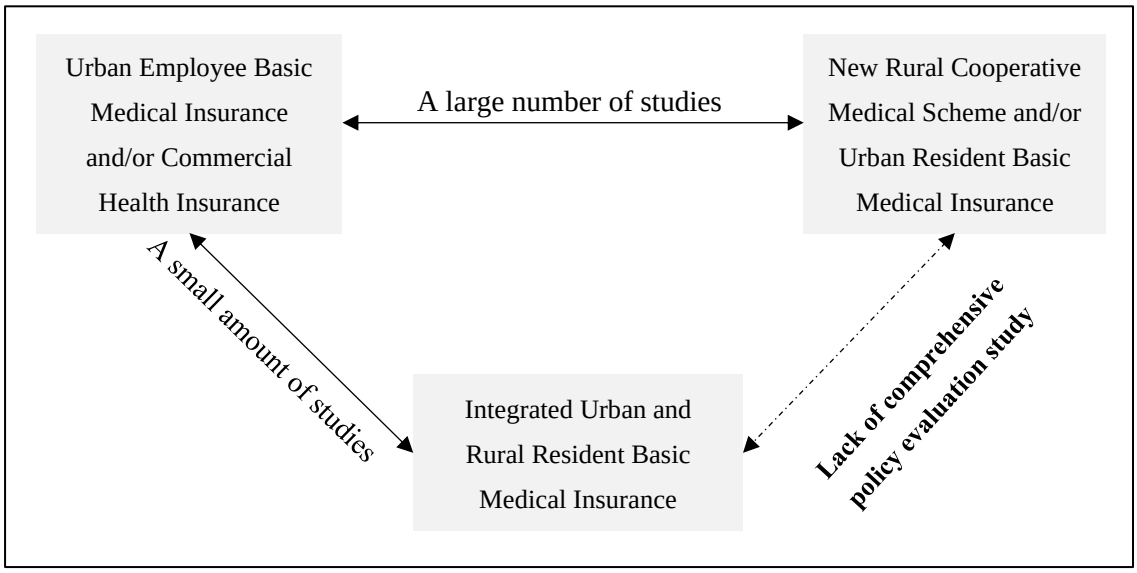

Figure 1-3 Research gap on health insurance in China

A comprehensive policy evaluation study is necessary to better understand whether and how the reform has been effective in achieving its policy goals and to assess what remains to be done in the future. Local governments have offered different integration plans. Although only one institution manages the integrated urban and rural resident insurance, many regions still offer different payment plans and benefit packages within one insurance scheme. The latest research highlighted that nearly half of the interviewees were dissatisfied with the integration, and many interviewees argued that integration reforms had failed to significantly reduce the large gap between urban and rural residents (Shan *et al.*, 2018). Therefore, a better understanding and evaluation of these integration policies themselves is both critical and timely. In addition, research on different pilot areas presents differential outcomes, some even contrary to expectations (Xie, Hu and Chen, 2017; Dai *et al.*, 2018). Therefore, exploring regional differences in policy effects is a way to better understand policy implementation rather than just the overall average integration effect.

1.3 Aims, Objectives, and Thesis Framework

Based on the current government policy goals and existing research results, this research aims to discuss the diverse urban-rural health insurance integration policies implemented in different cities, drawing on document analysis. It also aims to evaluate the effect, variations, and underlying mechanisms of policies through econometric analysis. The thesis focuses on answering key questions: What are the commonalities and disparities observed among the various integration plans? To what degree and through what means does the integration influence individuals' utilisation of healthcare services and their health expenditure?

To answer questions above, the following research objectives are set:

1) To summarise the different regional integration policies and to discuss the potential benefits and problems of integration;
2) To estimate the effect of integration on health service use and costs, at the national level;
3) To explore the differences in policy effect among the different populations in socio-economic subgroups, at the national level;

4) To assess the impact of different integration policies on health service use and costs at the city level, using ischaemic heart disease as an example;
5) To investigate the mechanisms of action of different integration policies at the city level and how they differ, using ischaemic heart disease as an example.

This thesis includes five parts, encompassing a total of ten chapters (Figure 1-4). Following this introductory chapter, Chapters 2 and 3 provide detailed introductions of the evolution of China's health insurance system, as well as demographic, economic, and healthcare trends in China. These chapters aim to familiarise international readers, unacquainted with the Chinese background, with a comprehensive understanding of the current situation in China.

Chapters 4, 5, and 6 constitute the literature review part. They discuss economics and healthcare theories pertinent to this study, causal identification strategies and analysis models, and previous empirical studies on health insurance integration. These chapters jointly establish the specific research objectives and construct the framework for subsequent analysis in this thesis.

Chapters 7, 8, and 9 form the analysis part. Chapter 7 presents a qualitative analysis addressing the first research objective. It discusses the potential benefits and problems of integration in terms of premiums, outpatient reimbursement, and hospitalisation reimbursement. This is mainly accomplished through a comparison of insurance terms before and after integration in four selected cities (Chengdu, Tianjin, Beijing, and Shanghai). Chapter 7 not only provides a rationale for subsequent quantitative analysis but also offers possible explanations for the results derived from quantitative analysis. Chapter 8 undertakes a quantitative analysis using data from the China Health and Retirement Longitudinal Study (CHARLS) survey, focusing on the second and third objectives. It analyses the integration effects at the national level and explores the heterogeneity of effects among urban-rural and income subgroups, Chapter 9 conducts another quantitative analysis, employing data from Electronic Medical Records (EMR). It addresses the fourth and fifth objectives, analysing different policy effects in six anonymous cities and examining their respective mechanisms.

The final chapter is the conclusion, which summarises and discusses the main findings and contributions of the three analyses. It also discusses some potential measures to address current integration issues, highlights the strengths and limitations of this study, and concludes with suggestions for future research.

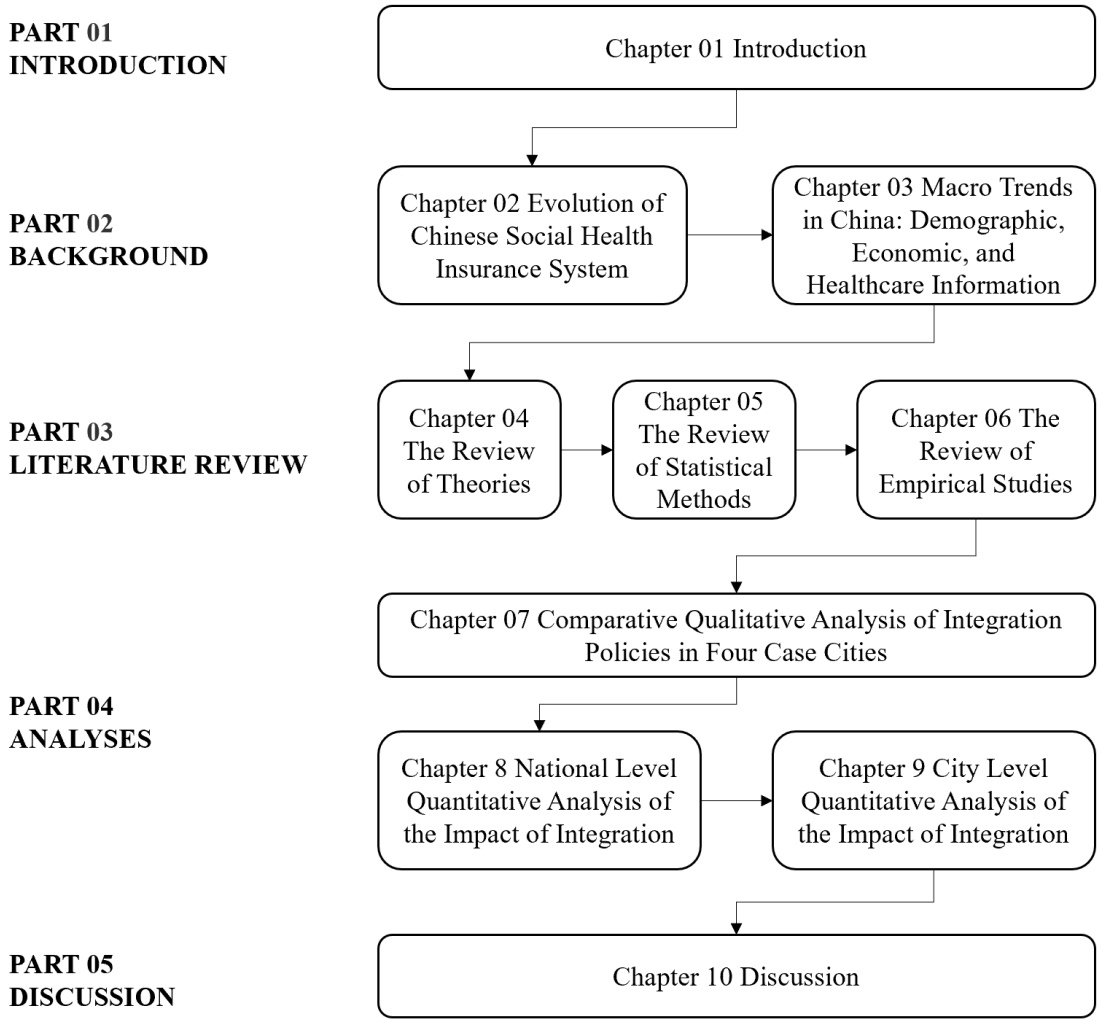

Figure 1-4 Thesis framework

Chapter 2 Evolution of Chinese Social Health Insurance System

The health insurance system in China has experienced great changes over the past seven decades. Figure 2-1 shows the evolution of the Chinese social health insurance system. In urban areas, the system has transitioned from the Government Insurance Scheme (GIS) and the Labour Insurance Scheme (LIS) to the Urban Employee Basic Medical Insurance (UEBMI) and the Urban Resident Basic Medical Insurance (URBMI). In rural areas, the Rural Cooperative Medical Scheme (RCMS) was replaced by the New Rural Cooperative Medical Scheme (NRCMS). These three basic health insurance schemes collectively form the core of China's health security system. However, they were separately designed for urban and rural populations, leading to separate management, varied funding sources, and distinct benefits. This fragmentation of the system significantly contributed to the inequity in access to health services and disparities in financial protection among individuals covered by different schemes (Meng *et al.*, 2015). To address these issues between urban and rural areas, the Chinese government endeavoured to integrate basic health insurance schemes.

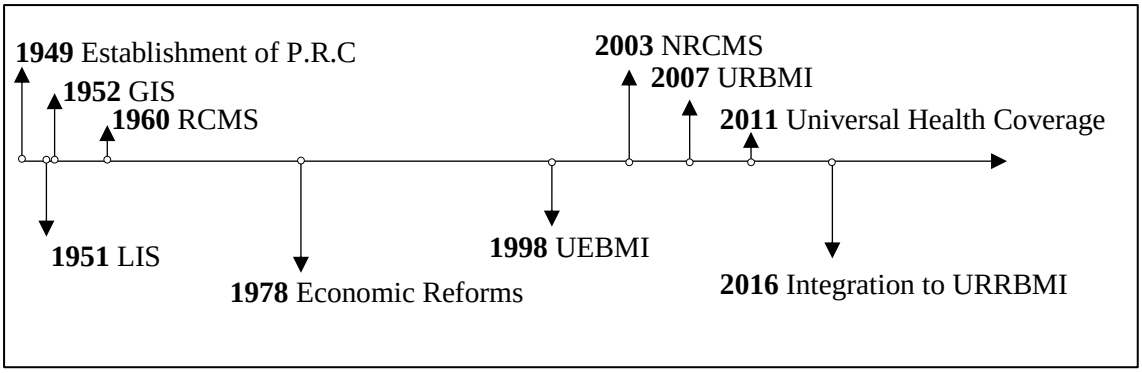

Figure 2-1 The evolution of the social health insurance system in China

This chapter discusses the establishment, achievements, and challenges encountered by China's health insurance system in different periods. It intended to provide readers, especially those unfamiliar with China, with a clear picture of the current state of China's health insurance system.

2.1 The Old Health Insurance System

China's social health insurance system was established in the 1950s (Liu, 2002) and was compatible with China's economic system at that time. In the era of China's planned economy, urban residents were assigned to specific working units (government or state-owned/collective-owned enterprises), while rural residents were affiliated with agricultural cooperatives[6]. Therefore, the initial social health insurance system in China contained three different categories based on each person's workplace: the Government Insurance Scheme (GIS), the Labour Insurance Scheme (LIS), and the Rural Cooperative Medical Scheme (RCMS), as shown in Figure 2-2.

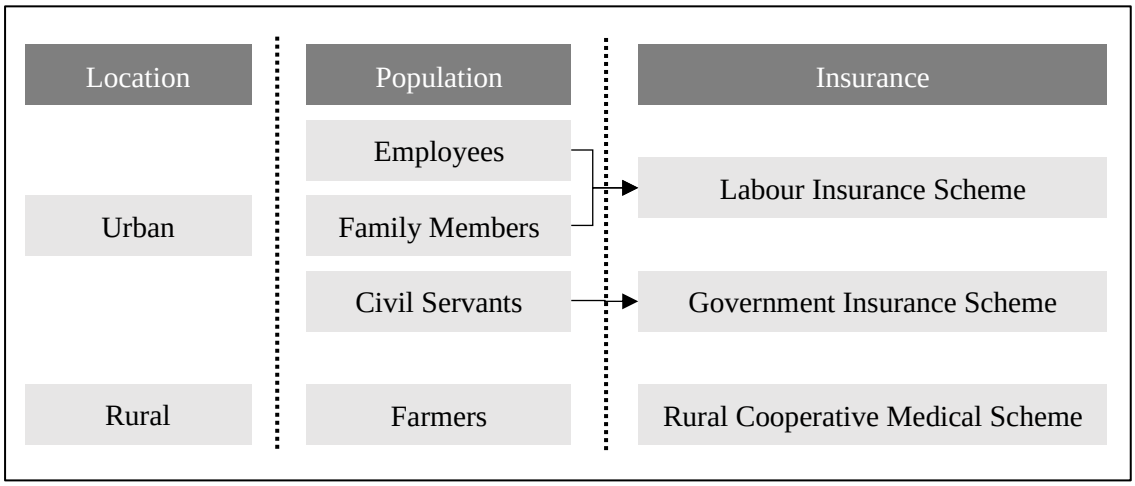

Figure 2-2 The old health insurance system in China

LIS was introduced in 1951. Under the LIS, all insurance funds came from the enterprise income, insuring all employees and their families. Employees enjoyed free health services, such as visits and drugs, while family members benefited from healthcare services at half the regular fee. GIS was established in 1952. It drew insurance funds from the government budget, covering civil servants, public institution staff, their dependents, students in higher education, and soldiers. RCMS commenced in 1955 and expanded nationwide in 1960. Funds came from

[6] Agricultural Cooperative, where production resources (land, machinery, farm animals) are pooled, and members farm jointly. All income is owned by the cooperative. Due to the egalitarian distribution method (everyone gets the same), there is a risk of low motivation of farmers for production.

premiums and the income of agricultural cooperatives (15% to 20%). RCMS covered all rural residents in the cooperatives. In the 1970s, the coverage rate of RCMS reached more than 90% (Xueshan *et al.*, 1995; Wagstaff and Lindelow, 2008).

During this period, there was a part-time medical practice scheme in Chinese rural areas known as "barefoot doctors". Farmers, with basic medical knowledge and skills, could farm half-time and practice medicine half-time approved by the grass-roots government. This scheme was vigorously promoted by the government in the 1960s and 1970s, and was hailed by the WHO as a successful example of providing primary healthcare services in developing countries (Zhang, 1994; Weiyuan, 2008; Zhang and Unschuld, 2008). Although barefoot provided only primary healthcare services like basic medicine and immunisation, life expectancy improved greatly, from 40 years in 1949 to 65.5 years in 1980 (Miller, Eggleston and Zhang, 2011). During this period, rural and urban residents in China had equal access to primary healthcare services (Miller, Eggleston and Zhang, 2011).

In 1978, China initiated major economic system reforms, transitioning from a planned to a market economy. Although the government did not officially announce any privatisation policies, they dramatically reduced political support and subsidies to public institutions. The plummeting subsidies compelled doctors and healthcare institutions, the suppliers of healthcare services, to be more profit oriented. This resulted in the de facto privatisation of China's healthcare system (Geyndt, Zhao and Liu, 1992; Blumenthal and Hsiao, 2005; Yu, 2015). The number of barefoot doctors declined rapidly after the economic reform. By the late 1980s, barefoot doctors either became private doctors or ceased working as doctors (Zhu *et al.*, 1989; Zhang and Unschuld, 2008). As for hospitals, although they were state-owned, subsidies from the government decreased by 90%. The rapidly reduced subsidies meant that hospitals had to earn more from patients[7], which led to the over-utilisation of higher-priced drugs and diagnostic

[7] Unlike the free medical care in the UK National Health Service, all services in Chinese hospitals are paid for by patients themselves or by insurance.

tests[8] (Hsiao, 1995; Martineau, Gong and Tang, 2004; Ma, Lu and Quan, 2008; Tao *et al.*, 2010; Eggleston, 2012). And then, it led to a rapid rise in individual health expenses (Watts, 2006).

Meanwhile, due to market-oriented economic reforms, state-owned enterprises were challenged by private enterprises due to reduced revenue. In the 1990s, millions of employees were made redundant to reduce the burden of state-owned enterprises and improve enterprises' competitiveness in the market. Consequently, they lost their health insurance coverage (Du, 2009; Yu, 2015). By 1998, half of the urban population lacked insurance coverage (Wagstaff and Lindelow, 2008). In addition, for employers, the increasing financial burden from GIS and LIS resulted in non-private enterprises reducing the healthcare benefits offered to their employees. Finally, the LIS was crippled.

Turn the attention to rural areas, in the economic reform, the "Household Responsibility System[9]" replaced the agricultural cooperative. Without cooperatives, the RCMS almost collapsed in the 1980s. By 1999, the health coverage rate in rural areas slumped from over 90% to only 7% (Eggleston, 2012). Thus, insurance coverage was declining, health spending was increasing, and the health gap between urban and rural populations was expanding. Even though many urban residents also lost coverage, the challenge was notably more pronounced for rural residents. Between 1980 and 2002, most of the population in China did not have public health insurance; the proportion of Out-of-pocket (OOP) expenses in total health expenditure increased 1.94 times from 1978 to 2001; the mortality rates of under five-year-olds in rural areas were 33 in 1000, 2.54 times than that in urban areas in 2003 (Eggleston, 2012; Yip and Hsiao, 2015; Chen, 2016).

[8] In the past, China implemented a Medicine Markup Policy. Hospitals were allowed to sell drugs at a price higher than the purchase price. Therefore, high-priced drugs enabled more profits. This policy has now been abolished.

[9] The Household responsibility system was introduced in 1979. Unlike in the agricultural cooperative, only a small amount of the income of the farmers is paid to the government, and most of the income belongs to the farmers themselves. This system greatly improved the productivity of farmers.

2.2 The New Health Insurance System Prior to the Integration

In response to the great healthcare challenges faced in the late 20th century, the Chinese government implemented a series of reforms to the healthcare system, including the social health insurance system reforms. This new social health insurance system consists of three new health insurance schemes: the UEBMI, the URBMI, and the NRCMS. The structure of the new health insurance system is shown in Figure 2-3. Insurance schemes in old and new systems are not matched one-to-one. The LIS and the RCMS have disappeared, but the GIS still exists in the central government and institutions under its direct management.

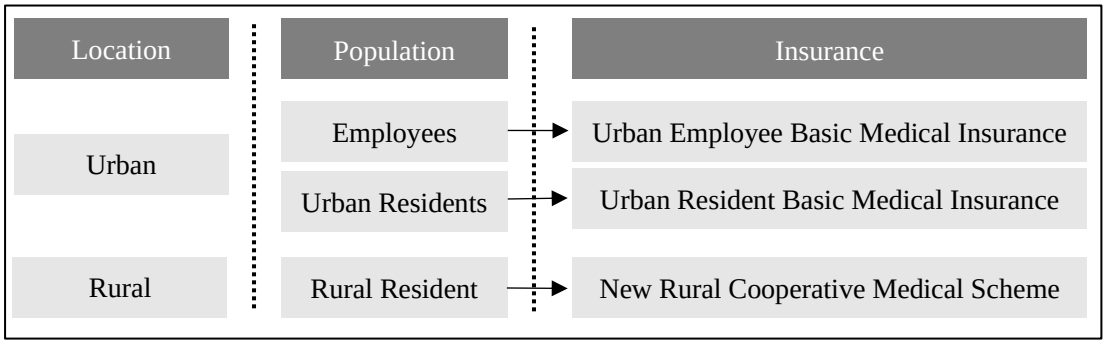

Figure 2-3 The new health insurance system in China

2.2.1 Establishment of the New Health Insurance System

The UEBMI was introduced to replace the LIS in 1998 (The State Council of China, 1998). The UEBMI covered all employees in both private and state-owned enterprises. However, unlike the LIS, the UEBMI did not cover the family members of employees. Most of the local government civil servants transferred to the UEBMI after 2010, and only the officers of the central government and its directly affiliated agencies remained in the GIS.

Fundings of the UEBMI came from the payroll taxes of the employer (6%-10%) and the employee (2%). The premium from employees went to their individual saving accounts, which could only be used for health expenses. The premium from employers went to the insurance funding pool. Money in the funding pool could be used for both hospitalisation and outpatient reimbursement, and money in individual savings accounts could be used to pay for the self-

payment portion of the bill. This scheme was designed and implemented at the city level, therefore the premiums, deductibles, co-payments, and reimbursement caps varied from city to city. In 2012, there were 333 UEBMI schemes across the country (Meng *et al.*, 2015).

The NRCMS was introduced in 2003 due to the outbreak of Severe Acute Respiratory Syndrome (SARS). At the end of the 20th century, many researchers were shocked by the decline in health coverage, and some attempted to assist the government to rebuild the health insurance system in rural China (Hao *et al.*, 1998; Carrin *et al.*, 1999). However, these studies did not receive the attention of the government immediately. In 2003, the outbreak of SARS caused huge economic losses in China (Lee and McKibbin, 2004). The government realised the insufficiency of China's healthcare system and then made large-scale investments in the public health system, including major subsidies for the promotion of the NRCMS.

The NRCMS was a family-based voluntary insurance scheme. Government subsidy was the main source of funding and farmers only needed to pay a small premium. The scheme was designed and implemented at the county level. Thus, the premiums, deductibles, co-payments, and reimbursement ceilings varied from county to county. In 2012, there were an estimated 2,852 NRCMS schemes across the country (Meng *et al.*, 2015). A study summarised four main package designs of NRCMS (Table 2-1) (Chen, 2016).

Table 2-1 Models of different NRCMS designs

Mode	Outpatient	Hospitalisation	Percentage
1	Paid by family saving accounts (Deductibles and caps apply)	Reimbursed according to a formula (Deductibles and caps apply)	65%
2	Reimbursed through collective funds (No deductibles and caps)	Reimbursed according to a formula (No deductibles and caps)	7%
3	Only reimburse expenses for critical disease (Deductibles and caps apply)	Only reimburse expenses for critical disease (Deductibles and caps apply)	11%
4	Not covered	Reimbursed according to a formula (Deductibles and caps apply)	17%

(Source: Chen, 2016)

In addition to rural residents and urban employees, urban residents, such as unemployed adults, children, and the elderly, constitute another group of people. Following the SARS outbreak in 2003, many in this group lacked health coverage, meaning that they were unable to pay for their medical expenses. In 2005, a government subordinate Think Tank reported widespread dissatisfaction and publicly criticised the government's policies (Ge and Gong, 2007). In 2007, urban resident basic medical insurance pilot projects commenced in 79 cities. Funds came from government subsidies and individual premiums. Initially, it only covered hospitalisation expenses, later expanding to cover outpatient expenses. In 2012, there were 333 URBMI schemes nationwide (Meng *et al.*, 2015). To date, there has been no research conducted that helps characterise the different designs of URBMI.

2.2.2 Achievements after New Health Insurance System Introduced

Since the inception of the three new health insurance schemes, China's health insurance coverage has steadily grown year by year. Finally, China almost achieved UHC again in 2011. This rapid expansion can be attributed to two main reasons: 1) increasing government subsidies; and 2) political orders to local governments.

In 2009, the Chinese government announced a "New Healthcare Reform", involving an investment of 850 billion CNY (around 85 billion GBP) (Freeman and Boynton, 2011). A big portion of this fiscal infusion was directed towards subsidising residents to enrol in NRCMS and URBMI. Five years after the initiation of the NRCMS, subsidies quadrupled (Wang, 2009). By 2011, 75% of the funding of NRCMS and 85% of the funding of URBMI came from government subsidies (Yip *et al.*, 2012).

Meanwhile, the mandate for expanding health insurance coverage was delegated to local governments as a compulsory task (political orders). Local officials had to spend considerable time and effort on health insurance enrolment, preventing negative impacts on their future career promotion (Yu, 2015). As a result, 95% of the population was covered by social health insurance at the end of 2011 (The State Council of China, 2012). Figure 2-4 shows the changes in the number of people covered by different social health insurance schemes.

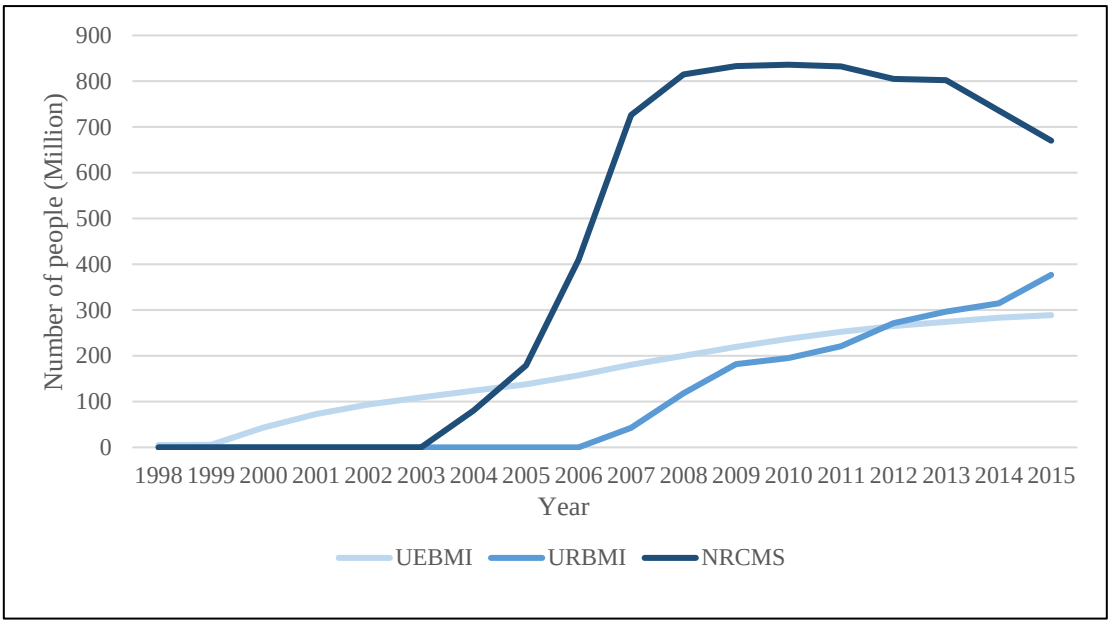

Figure 2-4 The number of people covered by social health insurance in China

(Data sources: Ministry of Human Resources and Social Security of the People's Republic of China, National Health and Family Planning Commission of the People's Republic of China, National Health Security Administration; Note: the integration of URBMI and NRCMS in some areas leads to the decline of the number of people in NRCMS)

The growth of government investment was reflected in increased public health expenditure, including government health expenditure and social health expenditure[10]. On one hand, it reduced the price of health services and drugs (subsidies to hospitals reduce nominal selling prices); on the other hand, it increased insurance benefits (reimbursement rate and ceiling increased, so patients' OOP costs decreased). From 1999 to 2015, the proportion of personal health expenditure in total health expenditure constantly decreased from more than 50% to 29.97% (Figure 2-5) and the annual healthcare institution visits increased from 2.08 billion to 7.7 billion (Figure 2-6).

2.2.3 Challenges in the Current Health Insurance System

Despite the vast success of China's new healthcare system, some old challenges remain pervasive (such as high OOP costs) and some new challenges have also emerged (such as duplicate insurance, and poor chronic disease control).

[10] Government expenditure refers to direct financial investment/subsidy in public health/medical institutions; social expenditure includes the reimbursement of social insurance, donations, etc.

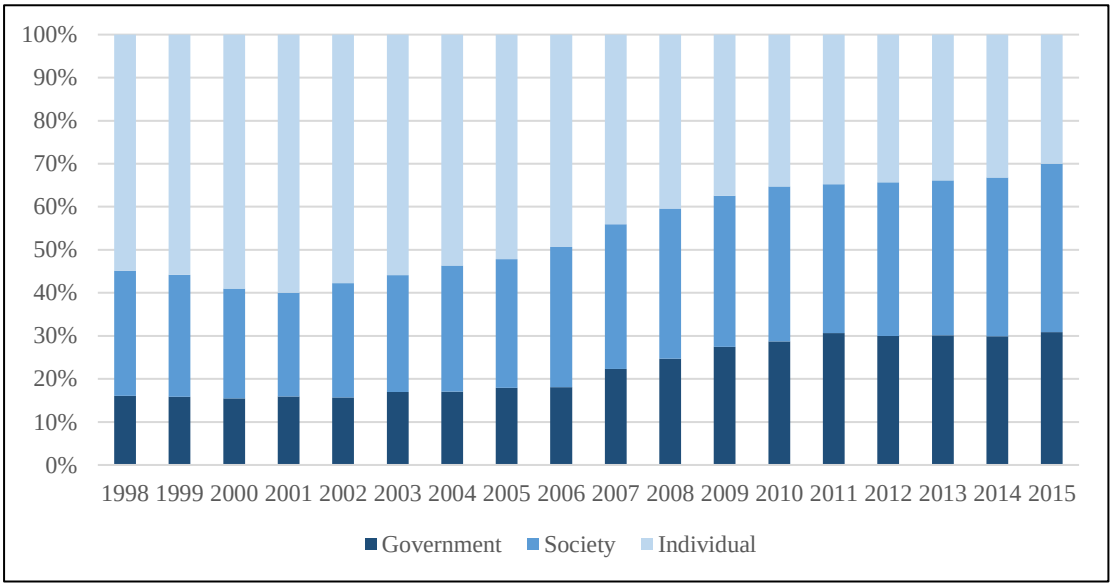

Figure 2-5 The distribution of annual total health expenditure in China

(Data sources: National Health and Family Planning Commission of the People's Republic of China)

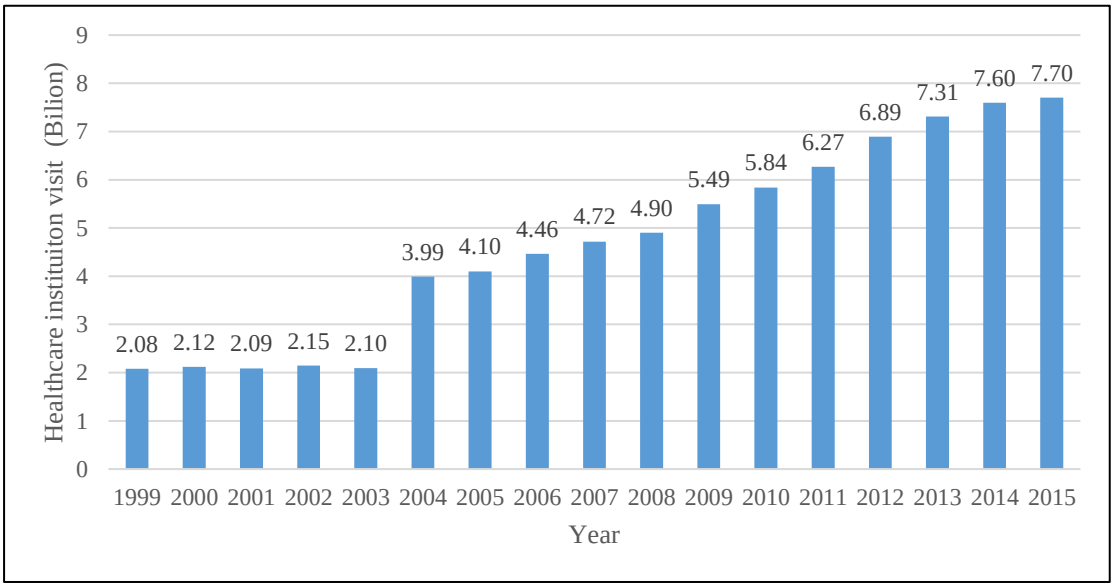

Figure 2-6 The number of annual healthcare institution visits in China

(Data sources: National Health and Family Planning Commission of the People's Republic of China)

Fragmentation of the health insurance system is the main problem. Firstly, the three new insurance schemes have substantial variations in the dimensions of coverage. Insurance coverage dimensions include breadth (% of the population covered), depth (% of health costs covered), and scope (the type of health services/drugs covered), as defined by the WHO (WHO, 2005) (Figure 2-7). The NRCMS has the greatest breadth, but less depth and scope than the other two schemes (Meng *et al.*, 2012; Yu, 2015). It is because high benefits cannot be sustained with low premiums. The annual premiums of both NRCMS and URBMI are around 100 to 150

CNY (about 15 GBP), which cannot support generous benefits packages. Many studies questioned whether the expansion of health insurance has reduced patients' financial risks because the patient's OOP medical expenses and the probability of catastrophic expenditure[11] have not decreased (Trujillo, Portillo and Vernon, 2005; Wagstaff and Lindelow, 2008; Lei and Lin, 2009; Q. Sun *et al.*, 2009; X. Sun *et al.*, 2009; Babiarz *et al.*, 2010; Meng *et al.*, 2012).

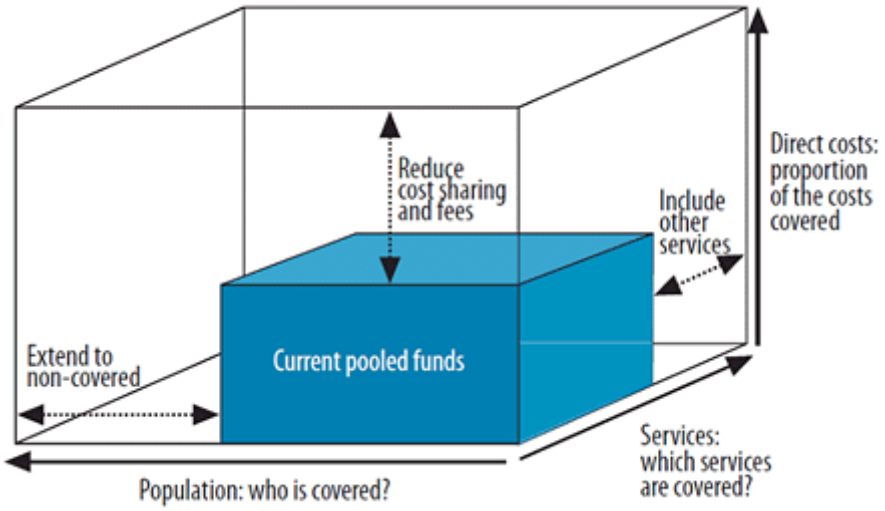

Figure 2-7 Three dimensions of health coverage
(Source: https://apps.who.int/iris/handle/10665/20302)

Secondly, the lack of flexibility among health insurance schemes poses a challenge for migrants. Migrant workers, who work in urban areas but come from rural areas, are very common in China. There were approximately 236 million migrant workers out of a total workforce of 767 million in 2012 (Gan, 2013). Although working and residing in cities, these migrants are identified as rural people based on their "Hukou" (initially determined by place of birth) and are enrolled in NRCMS. While "Hukou" can be transferred, stringent requirements prevail, particularly in more economically developed regions. For example, the transfer often requires three to five years of contractual work, a criterion challenging for temporary migrant workers without formal contracts. Consequently, when migrant workers are in cities, NRCMS cannot

[11] Catastrophic expenditure is an indicator in health research, which refers to whether annual medical expenses exceed a certain percentage of household disposable income. Different percentages are used in different studies.

be used. They can only receive reimbursement when they return to their hometowns, which may be far away and may take several days (Zhang, Nikoloski and Mossialos, 2017).

This issue leads to two main outcomes: 1) they forgo seeking healthcare due to the inability to afford full medical expenses; and 2) they also enrol in URBMI, which is duplicate insurance (also called overlapping insurance or double insurance). Another motivation for duplicate insurance is that some residents intentionally join two insurance schemes to maximise benefits. For example, in some regions, outpatient expenses can be reimbursed by NRCMS but not by URBMI, and the reimbursement rate of hospitalisation expenses in URBMI is higher than that in NRCMS, so residents enrol in NRCMS for outpatient expenses reimbursements and in URBMI for hospitalisation expenses reimbursements. There is also a possibility that a patient seeks reimbursements for the same expenses twice. A case in a government survey (Ministry of Finance of the People's Republic of China, 2014) showed that a patient spent 36,663.66 CNY (about 3,666 GBP) for one month of hospitalisation in 2013. The patient was reimbursed 18,548.14 CNY[12] (about 1,855 GBP) through URBMI, and then 20,165 CNY (about 2,016 GBP) through NRCMS. The total reimbursement was 38,713.14 CNY (about 3,871 GBP), which was higher than his actual medical cost. Duplicate insurance imposes additional financial burdens on the government. The number of people with duplicate insurance in China accounted for about 10% of the total number of insured people. According to the per capita insurance subsidy of 120 CNY (about 12 GBP) in 2010, the financial expenditure of duplicate insurance subsidy is 12 billion CNY (about 1.2 billion GBP) (Changxue *et al.*, 2016).

In addition, with economic development and ageing, the illness spectrum of Chinese residents has changed. Non-infectious chronic diseases have become the leading cause of death in China. In 2012, 86.6% of the total deaths were caused by non-infectious chronic diseases. Cardiovascular diseases, cancer, and chronic respiratory diseases were the main causes of death, accounting for 79.4% of the total deaths (National Health and Family Planning Commission, 2015). Chronic diseases have a long period of illness and a high cost of treatment, which

[12] The Per Capita Disposable Income in China in 2013 was 18,310.75 CNY, of which 26,467 for urban residents and 9,429.56 for rural residents (Data source: National Bureau of Statistics of China).

imposes a large economic burden on patients. Moreover, while chronic disease management focuses on prevention and control, basic medical insurance in most areas focuses on post-morbidity treatment and does not involve long-term disease management[13].

2.3 Integration to Unified National Health Insurance

There are huge differences in benefit packages and funding between urban and rural health insurance schemes, which is considered to be an important factor causing urban-rural inequality (Zheng, 2014; Meng *et al.*, 2015). Therefore, exploring a unified social health insurance scheme by 2020 was proposed as a major objective of China's health system reform (Meng *et al.*, 2015).

The integration of social health insurance comprises two aspects: vertical and horizontal. Vertical integration means that the funding pool and management of the three insurance schemes should be transferred from their current level (county level for NRCMS, city level for URBMI and UEBMI) to the provincial level, and finally to the national level. Horizontal integration means that the funding pool, information, and management of three insurance schemes should be merged, and a unified insurance scheme established. Integrating health insurance can bring several benefits: reducing human resources costs, improving administrative efficiency, reducing the government's financial burden, expanding the funding pool, improving healthcare access, narrowing the gap of healthcare due to regional disparities, and ultimately reducing health inequalities to some extent (Zhu, 2019).

By the beginning of 2016, nine regions, including Tianjin, Shanghai, Zhejiang, Shandong, Guangdong, Chongqing, Ningxia, Qinghai, and Xinjiang Production and Construction Corps, completed the integration of their urban-rural health insurance schemes (Figure 2-8). On 12 January 2016, the State Council of China published "The Opinions of the State Council on the Integration of the Basic Medical Insurance System for Urban and Rural Residents", which gave guidance on the integration of urban and rural basic medical insurance (The State Council of

[13] In the past, for some diseases, such as diabetes, treatment in the chronic phase is not reimbursed by insurance and only treatment in the acute phase is. This also results in some patients not being able to follow the daily maintenance treatment as required by their doctor.

China, 2016). This document mainly pointed toward six aspects of policy integration: integration of coverage population, financing policy, benefits package, drug and service reimbursement catalogue, designated institution, and funding management (details are provided in Appendix B).

By the beginning of 2018, 23 provinces (of a total of 31, excluding Hong Kong, Macao and Taiwan) implemented the integration of basic medical insurance for urban and rural residents (Figure 2-9). On 31 May 2018, China's National Healthcare Security Administration was formally established and appointed to manage all social security schemes. In the second half of 2018, four government institutions (National Healthcare Security Administration, Ministry of Finance of the People's Republic of China, Ministry of Human Resources and Social Security of the People's Republic of China, and National Health Commission of the People's Republic of China) jointly issued a statement entitled "Notice on implementation of integration of urban and rural resident basic medical insurance in 2018" and emphasised that the integration of urban and rural resident basic medical insurance must be completed in 2019 (National Healthcare Security Administration, 2018). By 2020, all cities have completed insurance integration.

Figure 2-8 Provincial administrative region completed integration by 2016

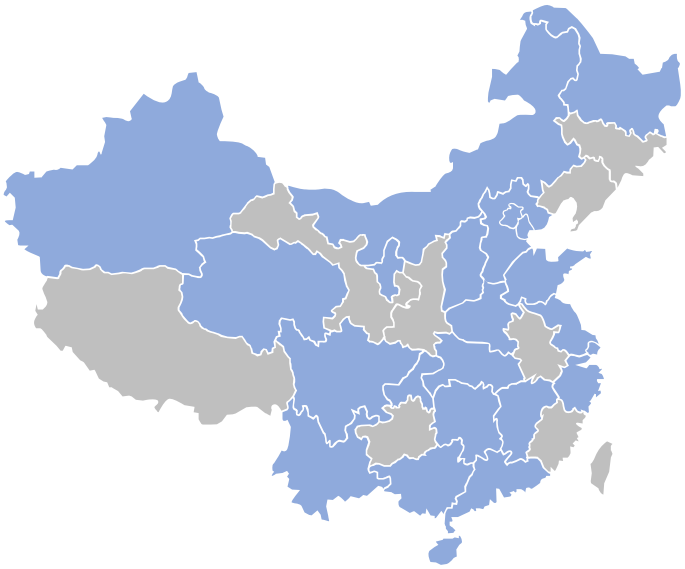

Figure 2-9 Provincial administrative region completed integration by 2018

Chapter 3 Macro Trends in China: Demographic and Healthcare Information

3.1 Introduction

Analysing the insurance integration policy effects on health service use and fees necessitates an understanding of the existing conditions and trends in health resource supply, health service use and fees in China. In addition, before analysing the policy heterogeneity and its effects, it is necessary to grasp the differences in demographic and economic characteristics, as well as health service use and health fees across diverse populations. Thus, this chapter has two purposes, one is to intuitively show the development trend of China's economy and healthcare, and the other is to intuitively compare the huge regional differences in China (including comparisons with other countries). This chapter aims to offer readers, especially those international readers who are not familiar with the Chinese background, a comprehensive understanding of the current situation in China. Meanwhile, it can also provide some inspiration for subsequent analysis design.

This chapter describes, summarises, and compares the current situation and trends in the country and the provinces through three key dimensions: population, economy and health expenditure; the availability of health resources; and the use of health services and associated fees. Section 3.2 describes the data sources and indicators used. Section 3.3 describes the time trends of each indicator at the national level. Sections 3.4 and 3.5 compare different indicators between urban and rural areas and across provinces, respectively. Section 3.6 is the conclusion.

3.2 Methodology

3.2.1 Data Sources

This chapter synthesises macro data from multiple sources, including open government documents, China Statistical Yearbooks, China Health Statistical Yearbooks, World Bank

Open Data, and Organization for Economic Co-operation and Development (OECD) Data. The earliest implementation of integrated resident insurance was in 2007. The implementation of a policy is related to the situation of years before implementation, not the year of implementation. Therefore, the data used in this chapter spans from 2006 to 2018 (Some data were only updated to 2017, while some were updated to 2019). For comparison, data from OECD countries is included to facilitate comprehension for international readers. All policy contents are sourced from government documents published online or peer-reviewed research articles.

It is important to clarify that the "implementation year" refers to the year when residents start to enjoy the insurance benefits, not the year of policy publication or insurance registration. Typically, residents are required to sign up for insurance three to six months before enjoying benefits.

3.2.2 Indicators

All indicators related to several aspects mentioned above, such as the economy and health resources, are described, summarised, and compared. Although health statuses, such as (healthy) life expectancy and chronic disease condition, are important indicators but not included in this study due to the absence of annual official data in China[14]. Table 3-1 provides a summary of all macro trend indicators. Only the key data are discussed in the main text, and detailed tables are given in Appendix C.

Demographics, economy, and health expenditure

Demography and economy serve as basic aspects used to capture the profile of a country. For population, this study examines the total population, ageing rate, and birth rate. They reflect the current and future demographic structure and the changing trends of the whole society. A previous study in China has already shown that medical expenses after age 60 will account for more than 80% of lifetime medical expenses (Jing, 2007). Thus, these indicators indirectly reflect the changes in the whole society's medical expenses.

[14] A national survey is conducted every 5 years, and data can be seen in the Statistical Yearbook.

Table 3-1 Summary of macro trend indicators

Categories		Indicators
Population and Economy	Population	Total Population (million)
		Aging Rate (%)
		Birth Rate (‰)
	Economy	Gross Domestic Product, GDP (¥, billion)
		Per Capita GDP (¥)
		Per Capita Disposable Income, PCDI (¥)
		Per Capita Consumption Expenditure, PCCE (¥)
		Expenditure of Health Care and Health services (¥)
Health Expenditure		Total Health Expenditure (¥, billion)
		Government Health Expenditure (¥, billion)
		Social Health Expenditure (¥, billion)
		Out-Of-Pocket Health Expenditure (¥, billion)
		Per Capita Health Expenditure (¥)
Healthcare Resources	Institutions	The Number of Hospitals (thousand)
		The Number of Tertiary Hospitals
		The Number of Grass-root Institutions (thousand)
		The Number of Inhabitants per Hospitals and Grass-Root Institutions
		The Number of Inhabitants per Tertiary Hospitals (thousand)
	Personnel	The Number of Licensed (Assistant) Doctors (thousand)
		The Number of Registered Nurse (thousand)
		The Number of Licensed (Assistant) Doctors per 1000 Inhabitants
		The Number of Registered Nurse per 1000 Inhabitants
	Facilities	The Number of Sickbeds (thousand)
		The Number of Sickbeds in Hospitals (thousand)
		The Number of Sickbeds in Basic Medical Institutions (thousand)
		The Number of Sickbeds per 1000 Inhabitants
Health Service Use and Cost	Visits	Health Institutions Visits (times, billion)
		Hospital Visits (times, billion)
		Grass-root Institutions Visits (times, billion)
		Average Number of Visits per year (times)
		Health Institutions Hospitalisations (persons, million)
		Hospitalisations in Hospitals (persons, million)
		Hospitalisations in Grass-root Institutions (persons, million)
		Annual Hospitalisation Rate (%)
	Costs	Average Outpatient Costs in Hospital per time (¥)
		Average Hospitalisation Costs in Hospital per person (¥)

The economic level is closely related to the amount and growth of a country's health expenditure, and economic development implies the ability of the government to invest in healthcare. This study measures the economy in terms of aggregate and per capita. Gross Domestic Product (GDP) refers to the final products at market prices produced by all residents in a country during a certain period. It is regarded as the best indicator of a country's economy. Per Capita GDP, Per Capita Disposable Income, and Per Capita Consumption Expenditure can reflect the economy from the "per capita" aspect.

The Total Health Expenditure (THE) is the most direct indicator to measure the total expenditure on health services. In China, the THE is divided into three parts: government, social and individual (internationally, it usually consists of two parts: public and private; in the UK, it is divided into government and non-government). The Government Health Expenditure (GHE) in China refers to the government's direct expenditure on health-related affairs, such as fiscal expense on public health services, medical subsidies, health-related administration and management of health insurance. The Social Health Expenditure (SHE) refers to all society inputs, such as social health insurance (excluding government subsidies), commercial health insurance, private healthcare institutions, and donation. Take resident insurance as an example, in the annual insurance fund, the premium paid by individuals belongs to SHE while the government subsidy to the resident insurance belongs to GHE. The Individual Health Expenditure (IHE) only refers to the cash expenditure on health services by residents, it is equal to the Out-Of-Pocket expenses.

Supply of Healthcare Resources

The supply of healthcare resources reflects the ability to provide health services. The increasing medical resources enhance the availability of health services. If there is no change in medical resources, the patient's behaviour may not change even if the insurance reimbursement term is improved. This study measures healthcare resources from three aspects: healthcare institutions, medical personnel, and facilities.

In China, healthcare institutions refer to all qualified health-related institutions, including four subgroups: hospitals, grass-root level healthcare institutions, specialised public health institutions, and other healthcare institutions (Figure 3-1). Among them, specialised public health institutions include the Centre for Disease Control and Prevention, health education institution, blood collection and supply institution, first aid centre, etc., and other healthcare institution includes medical research institution, medical training institution, medical statistical information centre, etc. Not all healthcare institutions are directly related to patients. Therefore, the number of hospitals and the number of grass-roots healthcare institutions are specially selected to reflect the ease of patients' visits. In addition, according to the classification of the hospital level, the number of tertiary hospitals is selected to reflect the situation of "high-quality" medical resources.

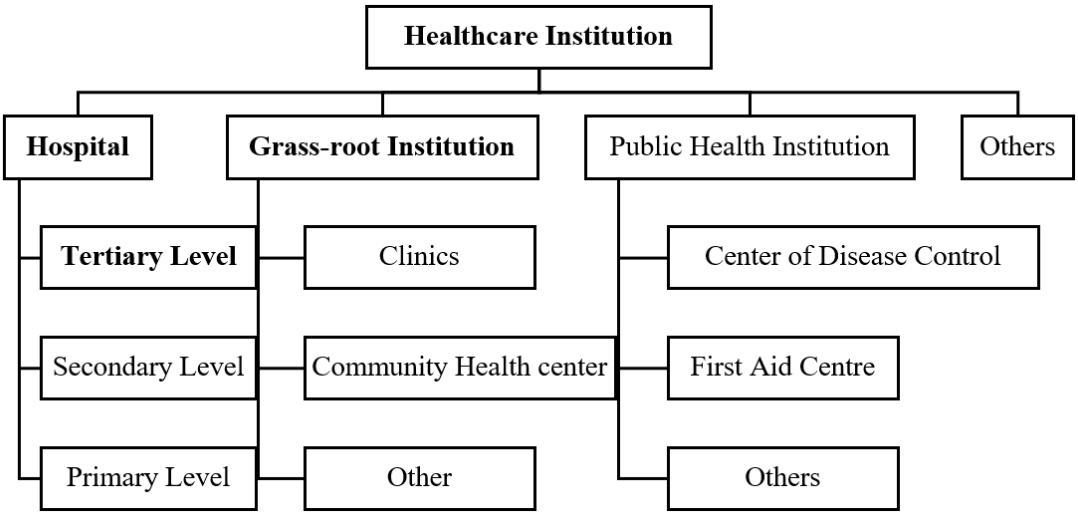

Figure 3-1 Classification of healthcare institutions in China

Medical personnel refer to all employees in healthcare institutions, including medical technical personnel (such as doctors, nurses, and pharmacists), village doctors and assistants, other technical personnel (technicians of device repair, research, teaching, etc.), and managerial and service staff. Only licensed (assistant) doctors and nurses are directly related to patients, so they are selected as indicators. In statistical yearbooks, indicators of medical facilities include beds, equipment and building area. Only the number of beds is selected to use because it is directly related to the patient's hospitalisation.

When comparing, no matter whether international comparison or provincial comparison, due to the difference in population, the comparison of the total amount of resources cannot reflect the accessibility of medical resources well. Therefore, the ratio of medical resources to population is a better indicator, including the number of inhabitants per hospital and basic medical institution, the number of inhabitants per Tertiary Hospital, the number of medical technical personnel per 1000 inhabitants, the number of licensed (assistant) doctors per 1000 inhabitants, the number of registered nurses per 1000 inhabitants, and the number of sickbeds per 1000 inhabitants.

Health Service Use and Costs

Health service use in this study refers to healthcare institution visits and hospitalisations. Visits include not only outpatient services but also emergency and other visits, such as regular health management consulting. Selected indicators include health institution visits, hospital visits, grass-root institutions visits, the average number of visits per year, hospitalisation in health institutions, hospitalisation in hospitals, hospitalisation in grass-root institutions, and the annual hospitalisation rate.

Costs refer to all expenses paid to institutions. Thus, it involves the Out-Of-Pocket (OOP) costs for patients and the reimbursement costs by insurance. This study uses average costs in hospital per visit for outpatients and average costs in hospital per person for hospitalisations.

3.3 Results: National Trends

3.3.1 Demographics, Economy, and Health Expenditure

Overall, China's total population has remained stable while ageing rapidly. An important reason for the deepening of the ageing population is the sharp decline in the birth rate. This ageing trend is expected to contribute to a rapid rise in the medical cost of the whole society. From 2006 to 2018, the total population slightly increased from 1.31 billion to 1.39 billion. The ageing rate increased by about 50%, from 7.9% to 11.9%, but the birth rate dropped to 10.48‰ (Figure 4-2). The birth rate was relatively stable abound 12‰ in the early years. Two peaks

appeared in 2014 (12.37‰) and 2016 (12.95‰) due to the Two-Child Fertility Policy in 2013 and the Open Two-Child Policy in 2015. The Two-Child Fertility Policy means that in a family, as long as either the husband or the wife is from a one-child family, they are allowed to have two children. The Open Two-Child Policy means that all families are allowed to have two children. After 2016, the birth rate fell sharply from 12.95‰ to 10.94‰ in 2018 and 10.48‰ in 2019.

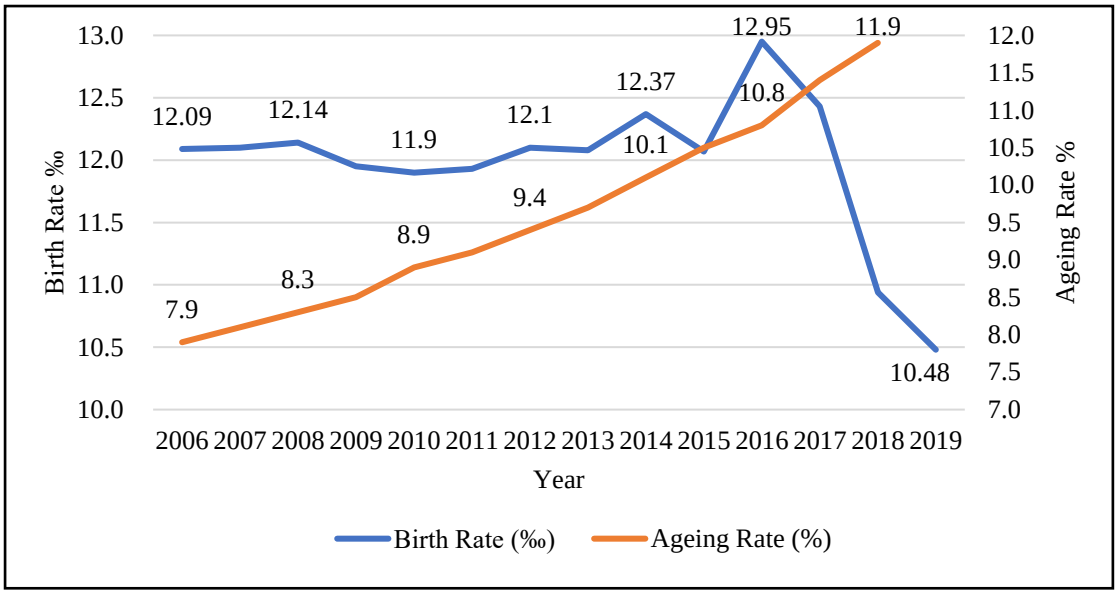

Figure 3-2 The trend of birth and ageing rate in China, 2006 to 2019

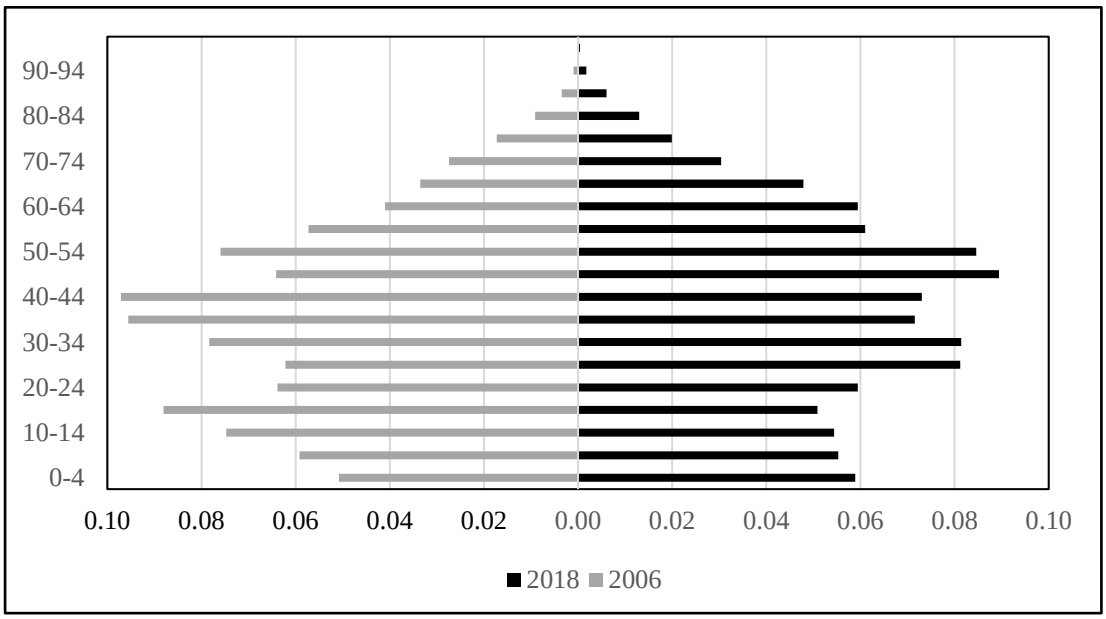

Figure 3-3 Population pyramid in China, 2006 and 2018

As a result, China's demographic structure changed (Figure 3-3). Compared to 2006, the majority of the population changed from 40-44 years old to 45-50 years old, while toddlers, children and adolescents were all less than 6% in 2018. It is foreseeable that the rapid rise of medical costs in the whole society and the increasing burden on the young to support the elderly are inevitable.

Over the past decade, China's economy has witnessed rapid development in both total volume and per capita measures. The total GDP in 2018 was 4.19 times that in 2006 (from 21,944 billion CNY to 91,928 billion CNY) while the per capita GDP was 3.94 times that of 2006 (from 16,738 CNY to 66,006 CNY). However, internationally, China was still not a high-income developed country. In 2018, China surpassed Mexico and Turkey, with per capita GDP reaching 9.77 thousand USD. However, it was still only 8.38% of Luxembourg (the highest), 15.56% of the United States (6th), 20.53% of Germany (13th), 22.75% of the United Kingdom (16th), 23.56% of France (19th), 24.87% of Japan (20th), and 31.15% of South Korea (22nd)[15].

With the development of the economy, people's income and expenditure also increased significantly. As China changed its statistical method of income and expenditure in 2013, the income and expenditure data of residents before 2013 are not provided[16]. From 2013 to 2018, China's per capita disposable income (PCDI) increased by 54%, from 18,311 CNY to 28,228 CNY, while the per capita consumption expenditure (PCCE) also increased by 50%, from 13,220 CNY to 19,854 CNY. In terms of medical expenditure, it increased by 84.76% from 912 (6.90% of PCCE) in 2013 to 1,685 in 2018 (8.49% of PCCE). Although the growth rate of medical expenditure was high, medical expenditure was still not a large proportion of total expenditure.

[15] Based on the World Bank Open Data. World Bank, International Monetary Fund (IMF), and OECD all provide an international comparison of GDP (per capita). However, IMF and OECD's data of China are far from the China Statistical Yearbook data. Thus, the World Bank data, which is similar to the China Statistical Yearbook data, is used.

[16] Before 2013, China used different statistical methods to calculate the income and expenditure of urban and rural residents separately. After 2013, China conducted statistical surveys on the use of unified statistical methods for all residents.

The proportion of medical expenditure in disposable income may be more reflective of the medical burden. In China, it was 4.98% in 2013 and 5.97% in 2018. In general, OOP medical expenses exceeding 10% of disposable income are deemed as "high medical expenses" (Łuczak and García-Gómez, 2012; Law *et al.*, 2013; K. Baird, 2016). According to this definition, on average, China's medical burden is not too heavy. However, this threshold would be lower for poor families. Studies in different countries have shown that high OOP medical expenses are more common among the poor, people in poor health, and the elderly (Cunningham, 2009; Marshall, McGarry and Skinner, 2010; K. E. Baird, 2016; Pandey *et al.*, 2018). In addition, it is important to note that a low medical burden may be because the actual medical burden is not high (the patients' medical needs have been met), or it may be that the patient chooses not to receive health services due to high out-of-pocket costs (patients' medical needs are not being met). An American literature review study (Eaddy *et al.*, 2012) has shown that as cost-sharing increases, adherence to treatment decreases, which leads to worse treatment outcomes.

More specifically, from 2006 to 2018, China's THE increased by 501%, from 984 million CNY to 5,912 million CNY. It grew faster than GDP, so its share in GDP increased from 4.52% to 6.57%. This is consistent with the WHO's global summary: "The health sector continues to expand faster than the economy" (Xu *et al.*, 2019). Internationally, China's total medical expenditure is not low. The United Kingdom's total health expenditure in 2018 was 214.4 billion GBP, accounting for 10.0% of the GPD[17], while China's total health expenditure in the same year was 5,912 billion CNY (around 591 billion GBP), accounting for 6.57% of the GPD. However, since China had a much larger population than the United Kingdom, China's total health expenditure in 2018 was only 4,237 CNY (about 424 GBP) per capita, compared with 3,227 GBP in the United Kingdom[18]. The proportion of IHE shows a declining trend from 49.31% in 2006 to 28.61% in 2018, while the proportion of GHE shows a growing trend from 18.07%

[17] Data Source: Healthcare expenditure, UK Health Accounts: 2018. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/bulletins/ukhealthaccounts/2018

[18] Data Source: Healthcare expenditure, UK Health Accounts: 2018. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/bulletins/ukhealthaccounts/2018

to 27.72%. This reflects the priority that the Chinese government attaches to healthcare in its development. (Details are provided in Appendix C)

3.3.2 Supply of Healthcare Resources

The substantial investment by the government has led to a significant increase in the number of hospitals and tertiary hospitals in China. From 2006 to 2018, hospitals grew by 73.68%, rising from 19 thousand to 33 thousand. Tertiary hospitals also experienced a huge increase of 143.83%, from 1,045 to 2,548 (Figure 3-4). However, despite this growth, the structure of healthcare institutions remained unchanged due to the relatively low proportion of hospitals (2%) compared to grass-root institutions (95%).

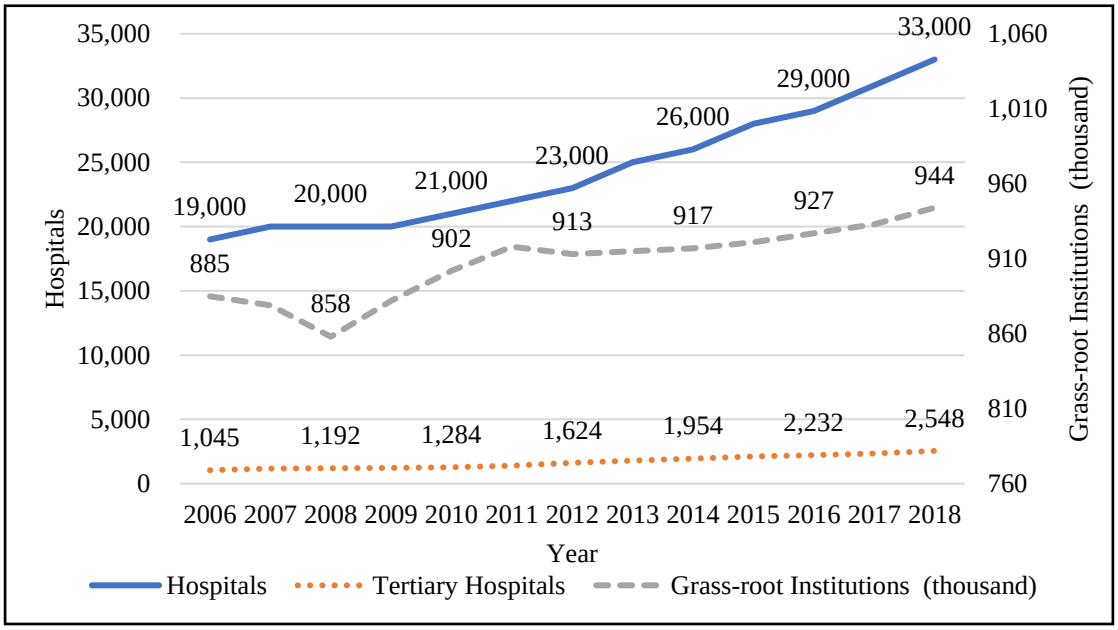

Figure 3-4 The trend of the number of healthcare institutions in China, 2006-2018

The government's emphasis on the medical field is also reflected in the substantial growth of medical personnel and sickbeds (Figures 3-5 and 3-6). From 2006 to 2018, licensed doctors and assistant doctors increased by 71.84%, from 2,099 thousand to 3,607 thousand. Registered nurses increased by 187.45%, from 1,426 thousand to 4,099 thousand. In the past, the number of nurses in China was relatively small, and the ratio of nurses to doctors only reached and exceeded 1:1 until 2013. The number of total sickbeds increased by 139%, from 3,512 thousand to 8,404 thousand, of which hospital sickbeds increased by 155%, from 2,560 thousand to 6,520

thousand, and grass-root institutions' sickbeds increased by108%, from 762 thousand to 1,584 thousand.

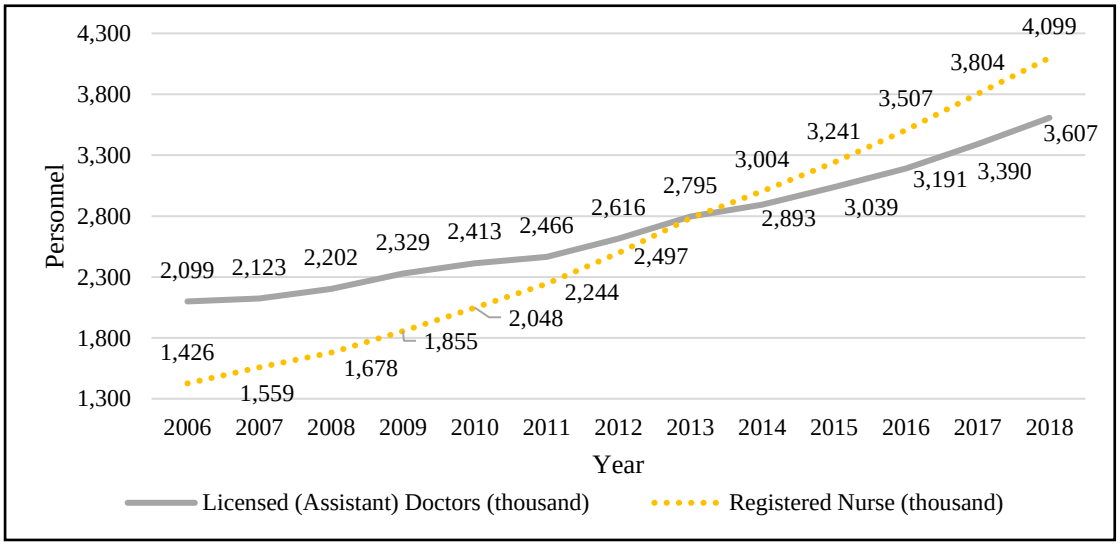

Figure 3-5 The trend of the number of doctors and nurses in China, 2006-2018

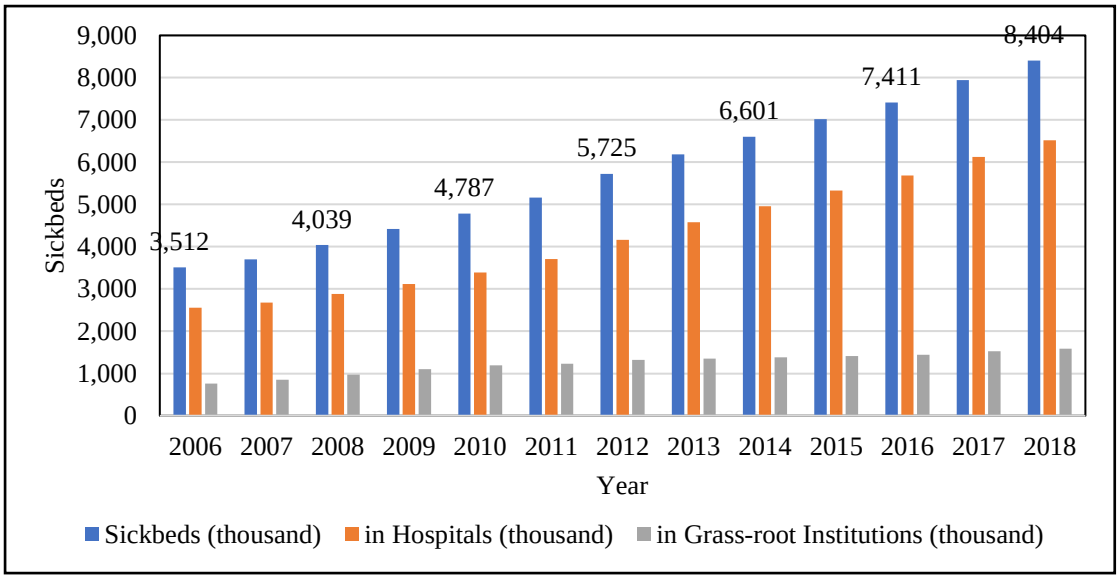

Figure 3-6 The trend of the number of sickbed in China, 2006-2018

On a per capita basis, the accessibility of medical resources also improved (Figure 3-7). Although the number of inhabitants per hospital and grass-root institution remained stable (around 1,450 people), the number of inhabitants per Tertiary hospital largely decreased by 56.44%, from 1,258 thousand people to 548 thousand people. It reflected the increase in the accessibility to high-quality health services. Meanwhile, the number of licensed (assistant) doctors per 1000 inhabitants increased by 61.88%, from 1.6 to 2.59, while the number of

registered nurses per 1000 inhabitants increased by 169.72%, from 1.09 to 2.94. The number of sickbeds per 1000 inhabitants increased by 123.33%, from 2.7 to 6.03.

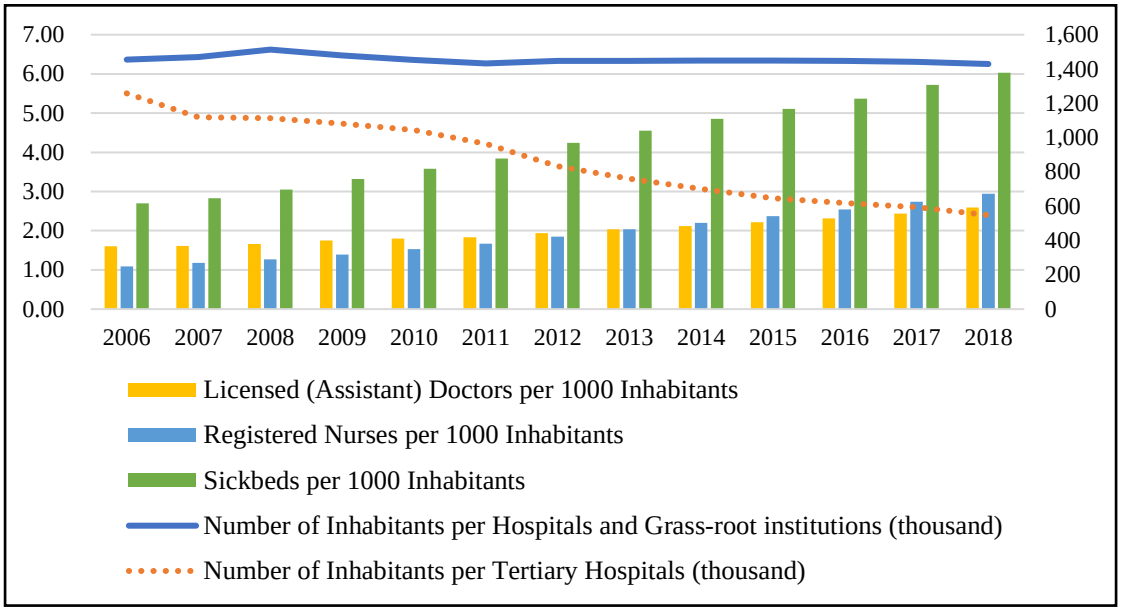

Figure 3-7 The trend of healthcare resources per capita in China, 2006-2018

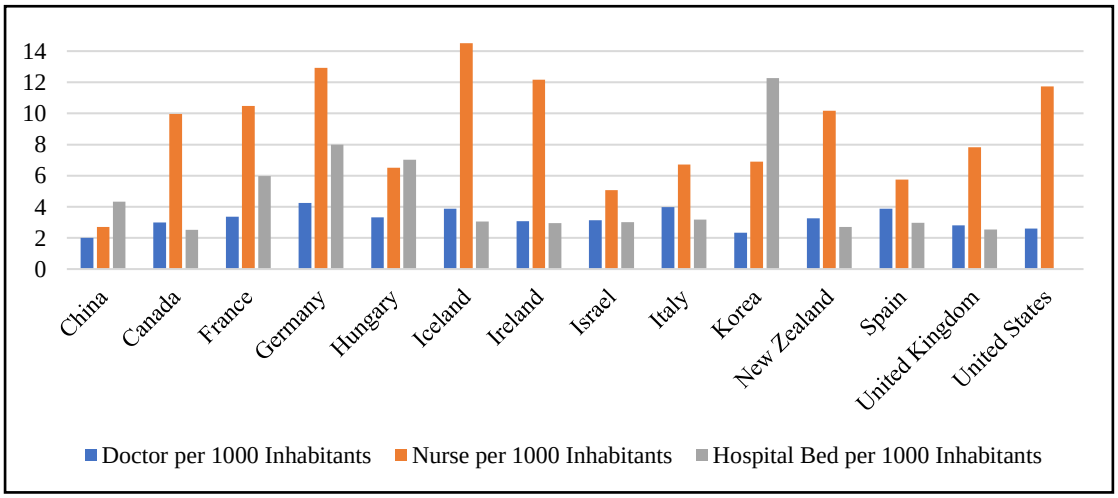

Figure 3-8 The healthcare resources among China and selected OECD countries, 2017

However, despite these improvements, China's per capita medical resources still lag behind developed countries (Figure 3-8). There were 2.01 doctors per 1,000 inhabitants in China, compared to 4.25 in Germany and 2.81 in the UK. For nurses, the number was 2.7 per 1,000 inhabitants in China, while the number were 12.93 in Germany and 7.83 in the UK. The number

of sickbeds per 1,000 inhabitants performed the best, which was 4.3 in 2017, superior to Canada, New Zealand, Italy, etc., but only 35.37% of South Korea and 54.25% of Germany.

3.3.3 Health Service Use and Fee

Overall, both visits and hospitalisations were on the rise, and the increase in hospitalisations was very large; both outpatient and hospitalisation costs were steadily increasing year by year (Figures 3-9/10/11). These hint at the importance of controlling for temporal trends in subsequent quantitative analyses.

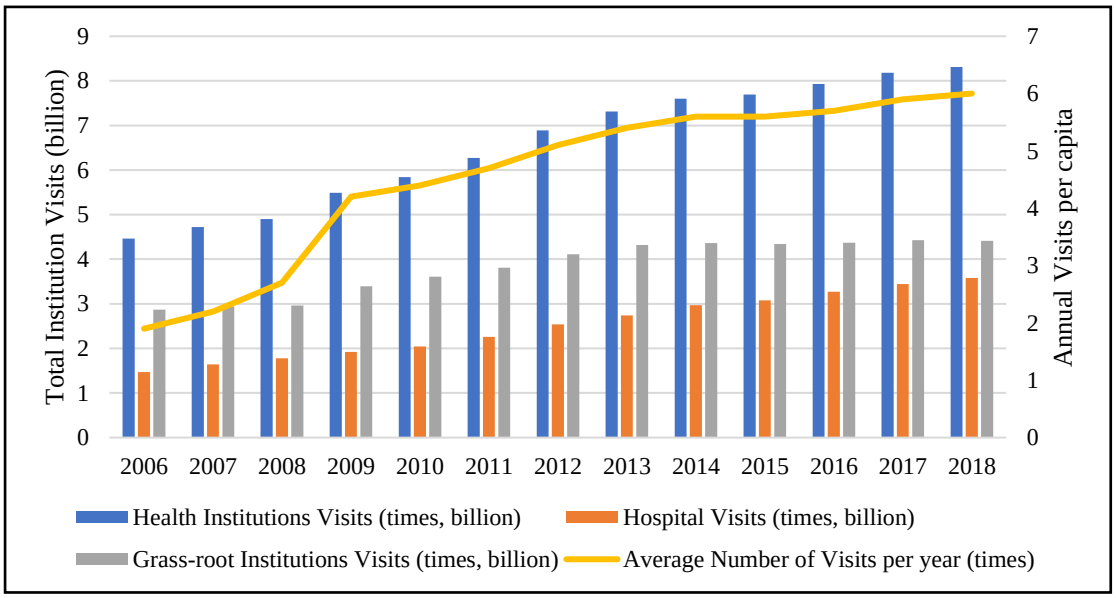

Figure 3-9 The trend of healthcare institution visits in China, 2006-2018

From 2006 to 2018, the total healthcare institution visits rose by 86.32%, from 4.46 billion times to 8.31 billion times, of which grass-root institution visits accounted for 50% to 60%, rose by 53.66%, from 2.87 billion to 4.41 billion times, while the hospital visits accounted for 30% to 40%, rose by 143.54%, from 1.47 billion to 3.58 billion times. The annual visits per person increased by 215.79%, from 1.9 times to 6 times per year. There was a significant jump in 2009 (from 2.7 to 4.2 times), probably due to the "New Healthcare Reform".

For hospitalisation (Figure 3-10), the total hospitalisations increased by 222.78%, from 79 million patients per year to 255 million patients per year, of which hospitalisations in hospitals

accounted for 60% to 80%, rose by 257.14%, from 56 million to 200 million, while hospitalisations in grass-root institutions accounted for 10% to 30%, rose by 257.14%, from 19 million to 44 million. The hospitalisation rate increased steadily by almost 2 times, from 6.14% to 18.27%.

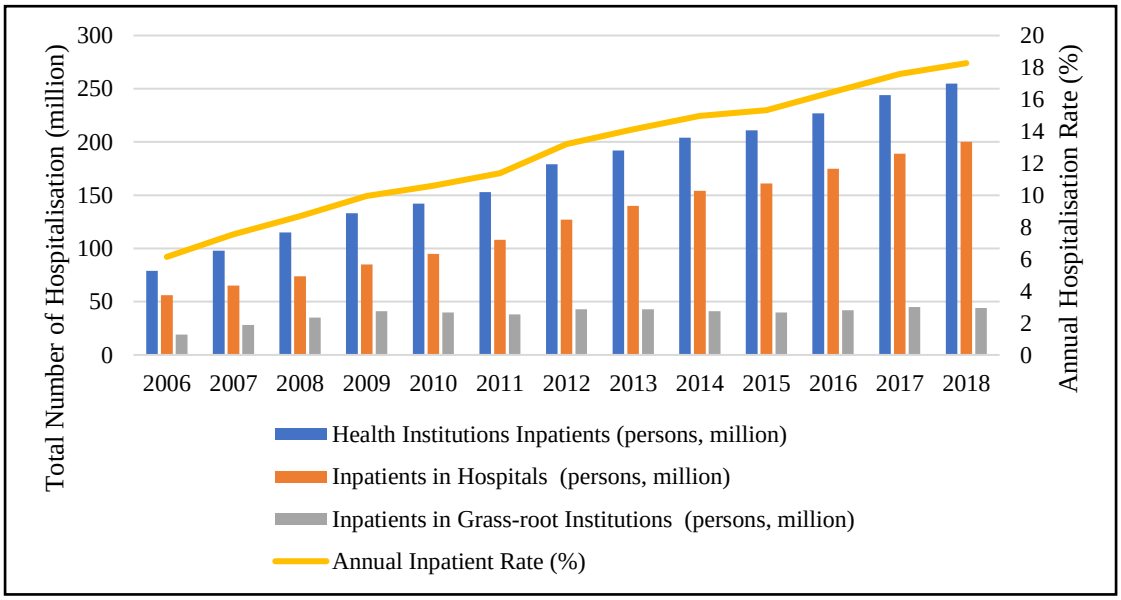

Figure 3-10 The trend of hospitalisations in healthcare institution in China, 2006-2018

The average outpatient costs increased by 112.4%, from 129 CNY per time to 274 CNY per time, while the average hospitalisation costs increased by 99%, from 4,669 CNY per patient to 9,292 CNY per patient (Figure 3-11).

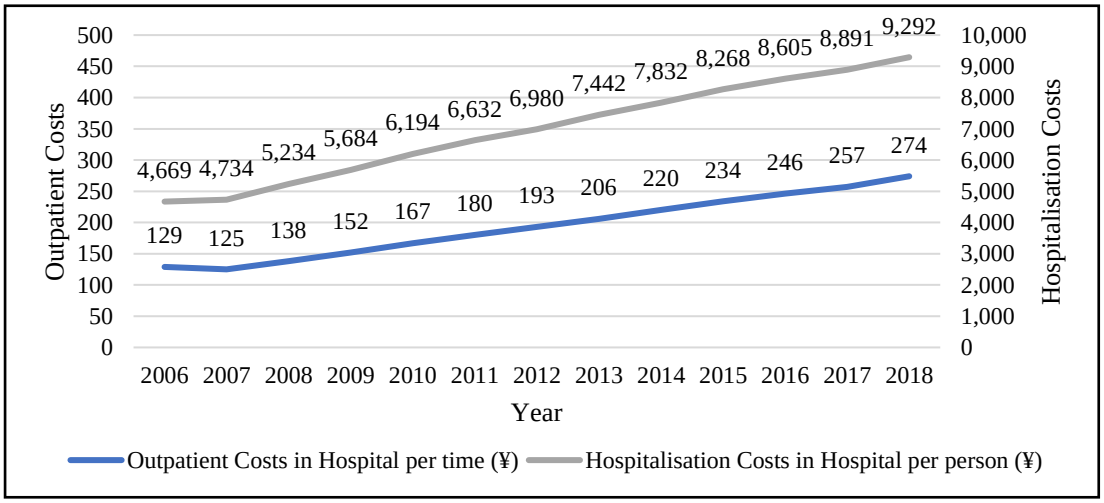

Figure 3-11 The trend of costs in hospital in China, 2006-2018

3.4 Results: Urban-Rural Differences

China's urban population surpassed the rural population in 2011, and the urbanisation rate increased from 44.34% in 2006 to 59.58% in 2018. During the 13 years, the disposable income of urban residents increased by 238%, from 11,620 CNY to 39,251 CNY. Prior to 2013, China did not count the disposable income of rural residents. From 2013 to 2018, the disposable income of rural residents increased by 55%, while that of urban residents increased by 48% during the same period.

For urban residents, medical expenditures have risen steadily with income overall, but the proportion of medical expenditures of PCDI did not change much, fluctuating around 5%. The proportion of medical expenses dropped significantly and continuously from 2010 to 2013, and then rose again in 2014. This decline may be due to the large-scale financial subsidies of the government's healthcare reform in 2009, which caused an increase in medical expenditure to be lower than the increase in income. For rural residents, although medical expenditure was much lower than urban residents (rural medical expenditure in 2018 was 60.62% of urban medical expenditure), their medical burden was higher due to lower income (8.48% for rural vs. 5.21% for urban in 2018). (Details are provided in Appendix C)

There is a huge gap in medical resources between urban and rural areas in China. No matter what kind of indicators, no matter what time, the indicator value of the urban area was 2-3 times that of the rural area (Figure 3-12). The overall urban-rural disparity of licensed doctors per 1,000 inhabitants was on the rise. It rose from 2.03 times in 2006 to a peak of 2.38 times in 2011, and then down to 2.2 times in 2018. Although, compared with urban areas, there are additional doctors (village doctors) in rural areas, which to some extent reduces the gap between urban and rural areas. However, by 2018, rural doctors accounted for only a quarter of the licenced doctors, which means that there was still a gap between urban and rural areas. The urban-rural disparity in the number of nurses and beds per 1,000 people had generally declined, of which the nurse gap decreased from 3.36 times to 2.8 times while the sickbeds gap decreased from 2.45 times to 1.9 times.

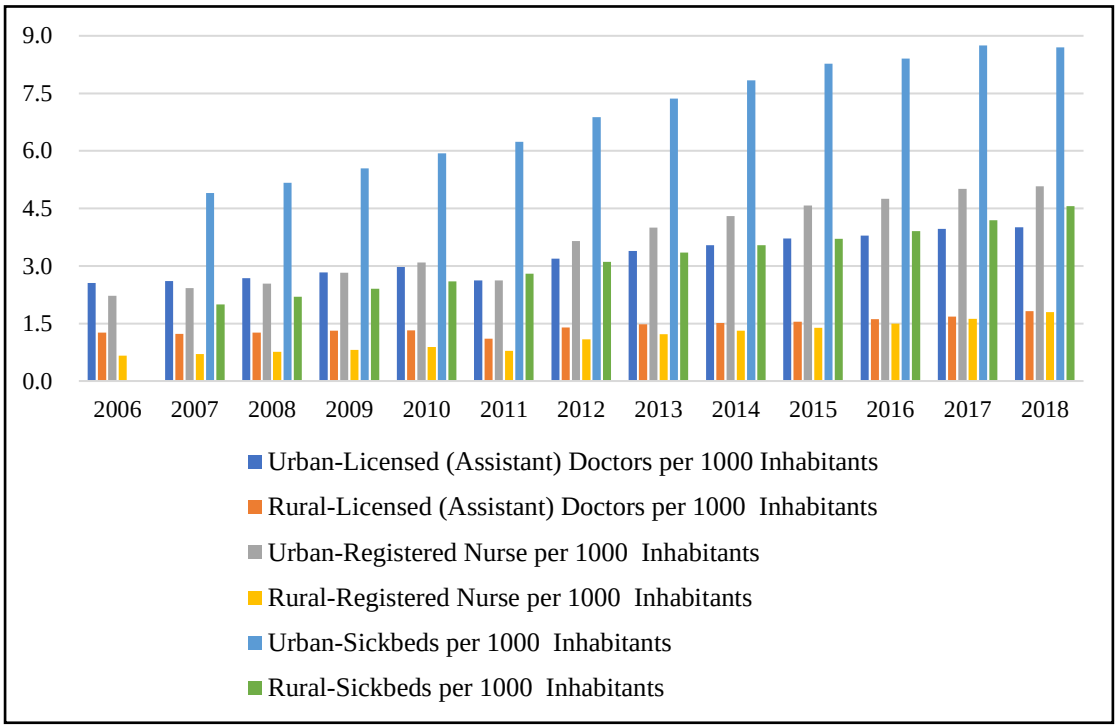

Figure 3-12 The trend and comparison of healthcare resources per capita between urban and rural China, 2006-2018

3.5 Results: Provincial Differences

There are great regional imbalances in China, such as population imbalance and economic imbalance, due to disparities in natural climate and geographical location. To provide a visual representation of these imbalances, data maps are used. There are 34 provincial administrative divisions in China. However, due to the difference in the health system, this study only targets 31 provincial administrative regions in China mainland, excluding Hong Kong, Macao and Taiwan. Thus, these three regions are also not highlighted/marked on the later data maps. Different colours and depths on the maps aid in the presentation of data. Blue donates positive indicators, such as GDP and income (darker indicating better), while red donates negative indicators, such as the population and costs (darker indicating worse).

3.5.1 Demographics, Economy, and Health Expenditure

In 2018, a stark economic imbalance prevailed among China's provinces (Figure 3-13). The eastern and southern coastal areas, as well as the central areas, demonstrated relative

development, while the western and northern areas were relatively backward. The GDP of the poorest province, Xizang, accounted for only 15.5% of that of the richest province, Guangdong (155 billion CNY vs. 9,995 billion CNY). Population distribution echoed this economic trend, with the East, South, and Middle regions teeming with inhabitants, in contrast to the sparser West and North (Figure 3-14). Guangdong, the most populous province, had 32.98 times the population of Xizang, the least populous province (113.46 million vs. 3.44 million).

Because of the extreme imbalance of population distribution, the per capita economy may be a better indicator to reflect economic activities. The top 3 rich provincial regions were Beijing, Shanghai, and Jiangsu, while the top 3 poor regions were Gansu, Heilongjiang, and Guangxi (Figure 3-15). An international comparison may more clearly reflect the economic imbalance of China's provinces. Beijing, the capital of China and the provincial administrative region with the highest per capita GDP, had a per capita GDP of 153 thousand CNY (about 21 thousand USD), which was similar to Greece's (20.32 thousand USD) and half of Britain's (42.94 thousand USD). Gansu, the poorest province in terms of per capita GDP, was only 31 thousand CNY (about 4 thousand USD), which was less than half of Turkey (9.37 thousand USD), the poorest country in the OECD.

Similar to economic development, high-income areas were concentrated on the southeast coast, and due to better economic conditions, residents in these high-income areas had higher consumption expenditure as well as healthcare expenditure. In terms of per capita disposable income, residents in Shanghai (64 thousand CNY), Beijing (62 thousand CNY), and Zhejiang (46 thousand CNY) were the top 3 richest, while residents in Xizang (18 thousand CNY), Gansu (17 thousand CNY), and Guizhou (17 thousand CNY) were the top 3 poorest (Figure 3-16).

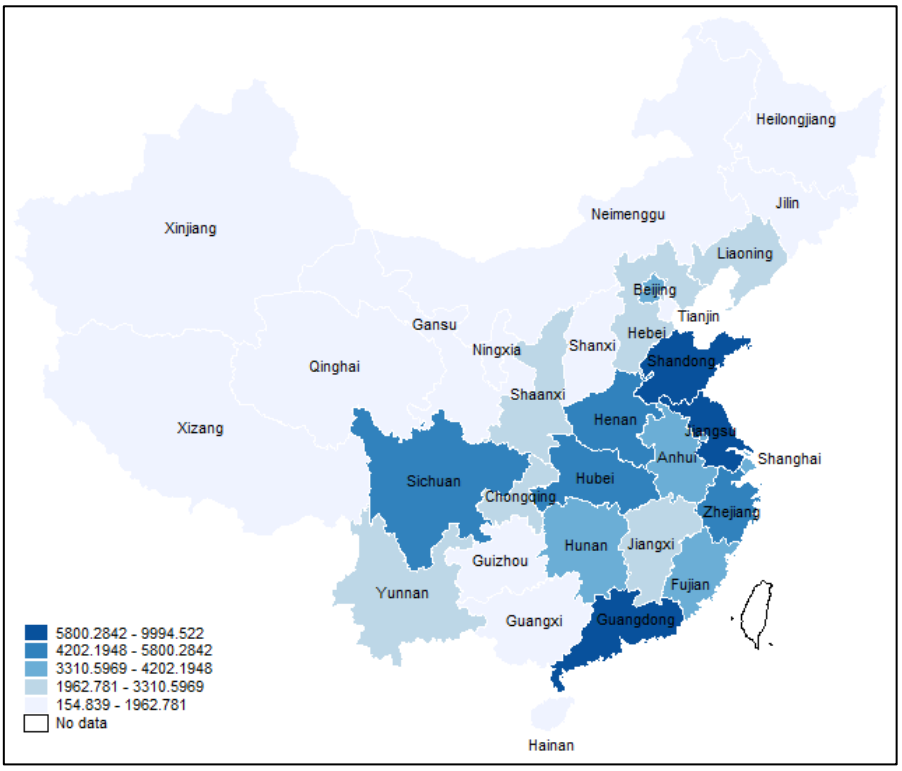

Figure 3-13 The provincial distribution of GDP (Billion, CNY) in China, 2018

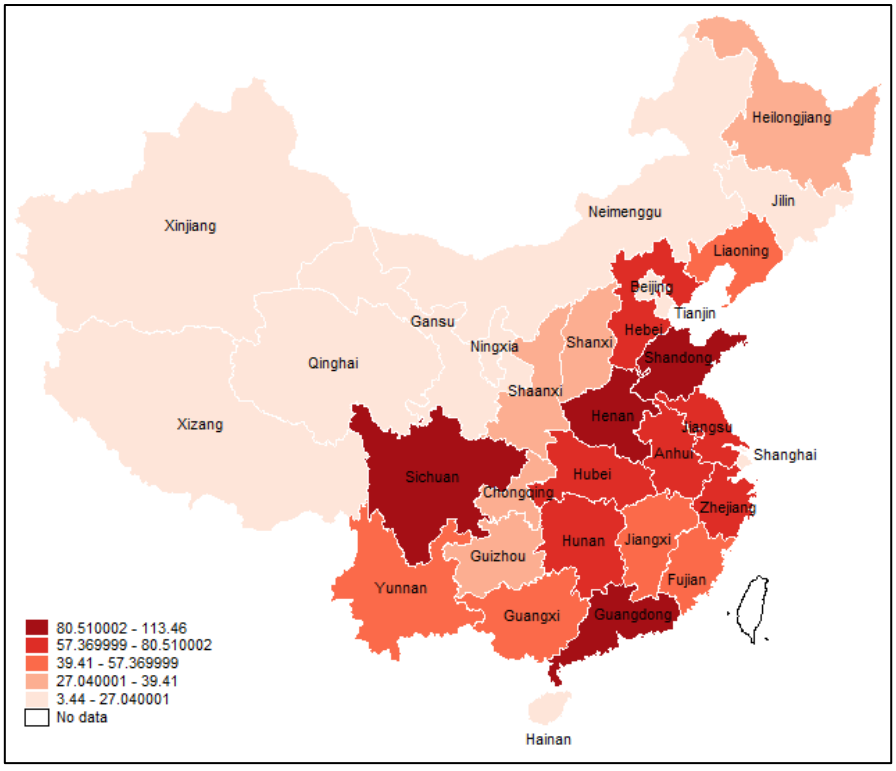

Figure 3-14 The provincial distribution of population (million) in China, 2018

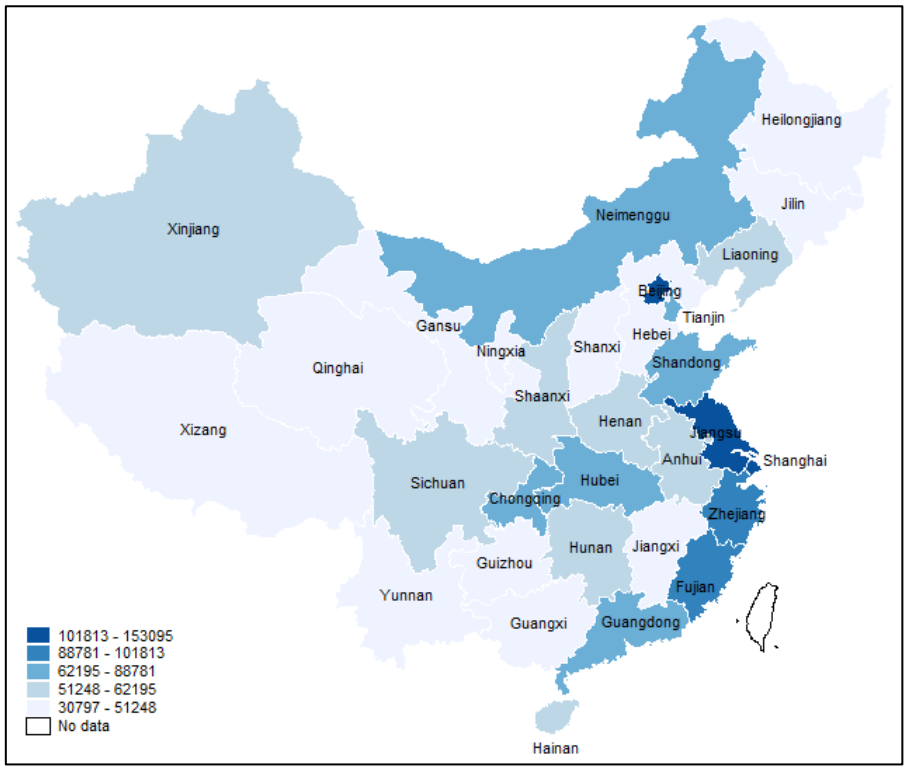

Figure 3-15 The provincial distribution of per capita GDP (CNY) in China, 2018

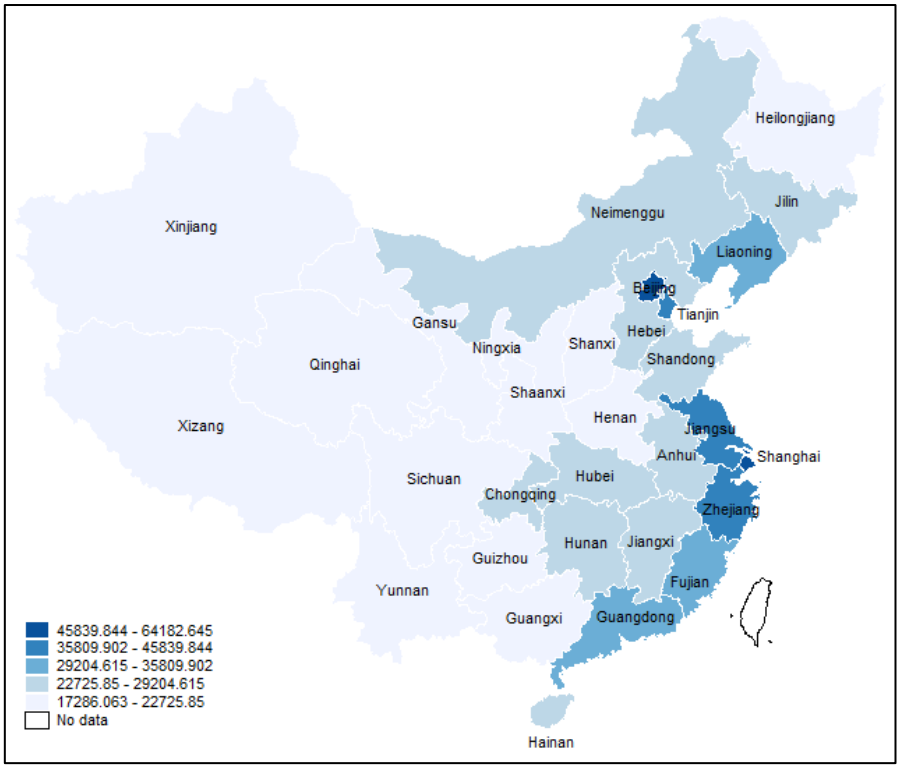

Figure 3-16 The provincial distribution of per capita disposable income (CNY) in China, 2018

Considering the per capita expenditure, the top 3 highest consumption expenditure regions were Shanghai (43 thousand CNY), Beijing (40 thousand CNY), and Tianjin (30 thousand CNY), while the top 3 lowest expenditure regions were Xizang (12 thousand CNY), Guizhou (14 thousand CNY), and Yunnan (14 thousand CNY) (Figure 3-17). More specifically, the top 3 highest healthcare expenditure regions were also Shanghai (3,275 CNY), Beijing (3,070 CNY), and Tianjin (2,677 CNY), while the top 3 lowest healthcare expenditure regions were Xizang (460 CNY), Jiangxi (1,000 CNY), and Guizhou (1,083 CNY) (Figure 3-18).

However, although the medical expenditure in the south was higher, if the ratio of per capita medical expenditure to disposable income is used to express the medical burden of residents, it can be seen that the medical burden in the northern region (especially northeast) and western region (8% to 9%) was significantly higher than that in the southeast coastal region (2% to 4%) (Figure 3-19).

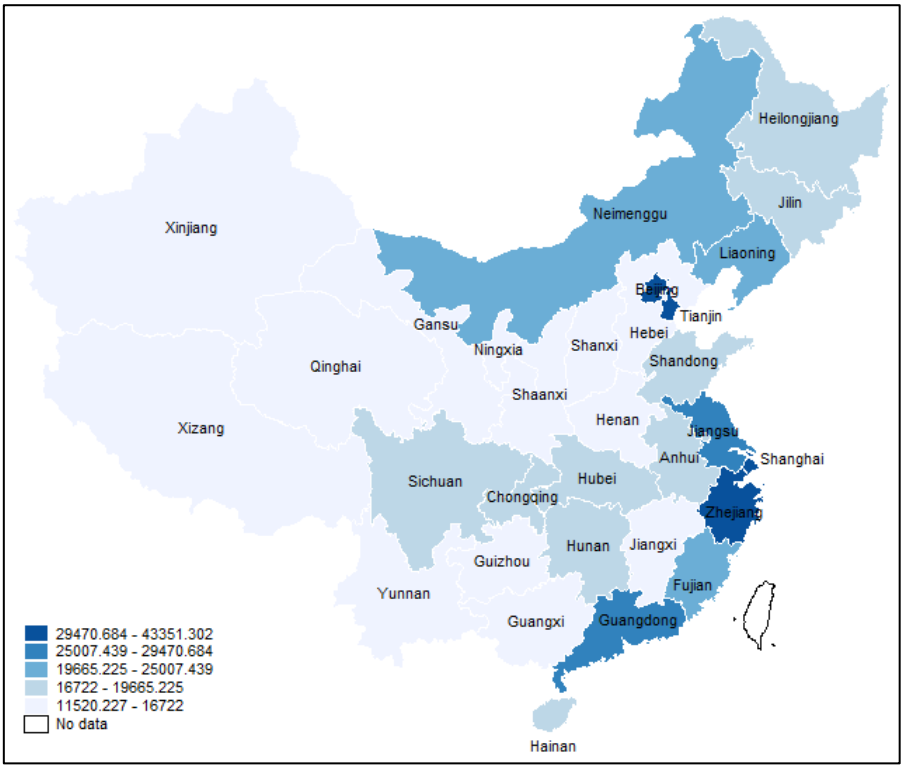

Figure 3-17 The provincial distribution of per capita consumption expenditure (CNY) in China, 2018

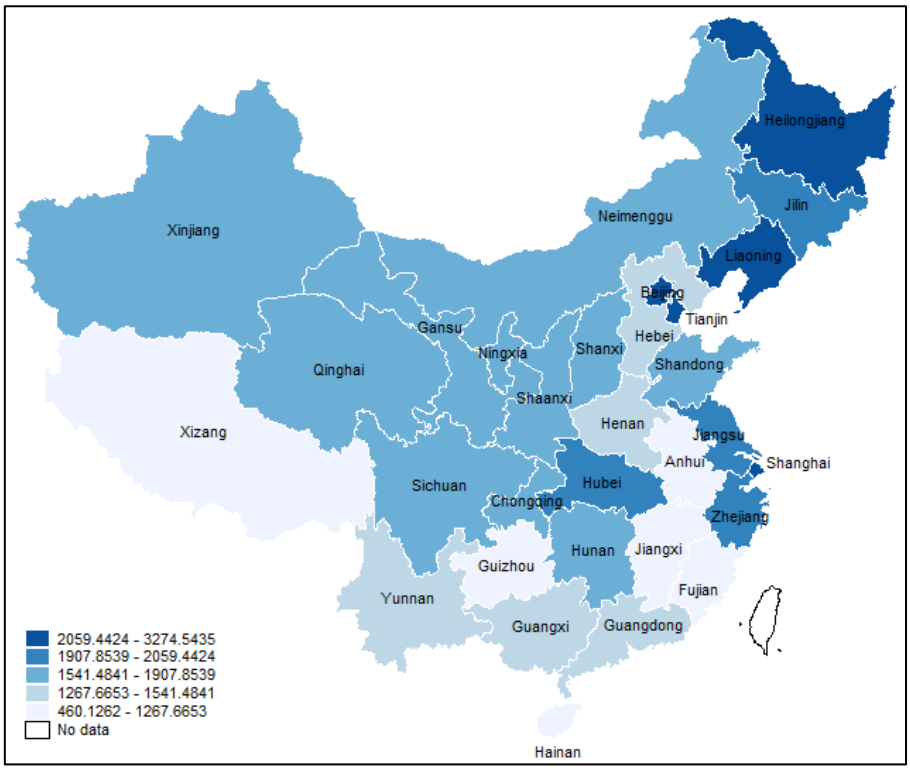

Figure 3-18 The provincial distribution of per capita healthcare expenditure (CNY) in China, 2018

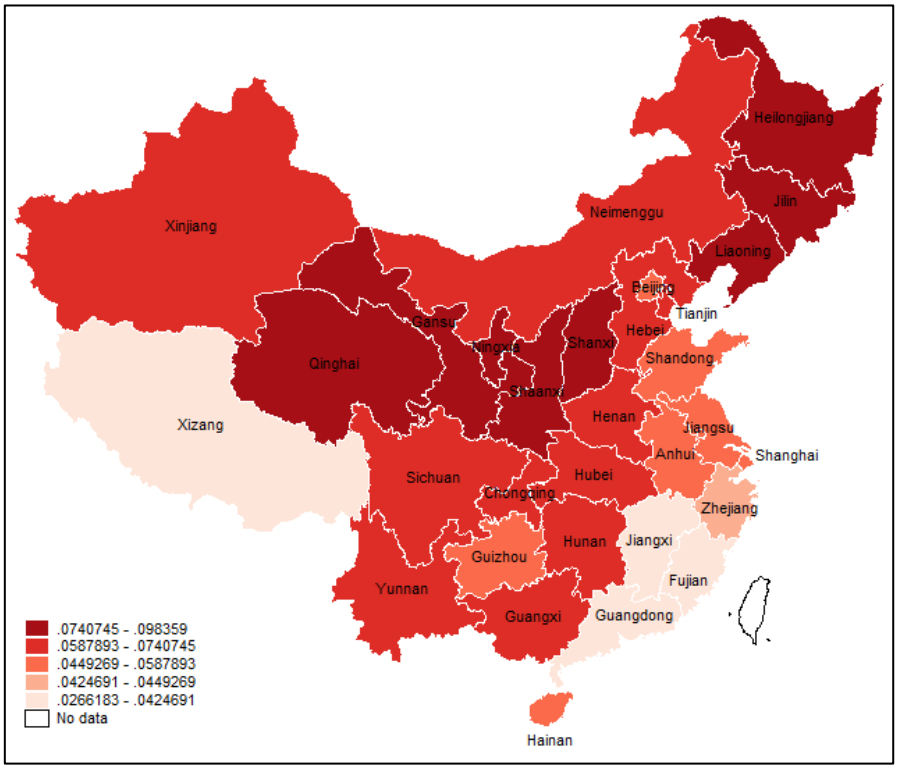

Figure 3-19 The provincial distribution of medical burden in China, 2018

3.5.2 Supply of Healthcare Resources

In 2018, there was a notable disparity in the distribution of healthcare institutions, with south-eastern regions experiencing scarcity while north-western regions exhibited abundance (Figure 3-20). Shanghai had the most limited healthcare institution, with one serving 4,580 inhabitants. Tianjin and Anhui followed with ratios of 2,744 and 2,537 inhabitants, respectively. In contrast, Xizang had the highest abundance, with one healthcare institution for every 504 inhabitants. For high-quality health resources, tertiary hospitals were scarcest in the middle regions, more available in the eastern regions, and most plentiful in the western regions (Figure 3-21). Hebei, Shandong, and Anhui emerged as the top three regions with scarce access to tertiary hospitals, accommodating 1,064 thousand, 1,022 thousand, and 930 thousand inhabitants per hospital, respectively. An interesting thing is that Beijing was the most abundant region, with only 211 thousand inhabitants per tertiary hospital.

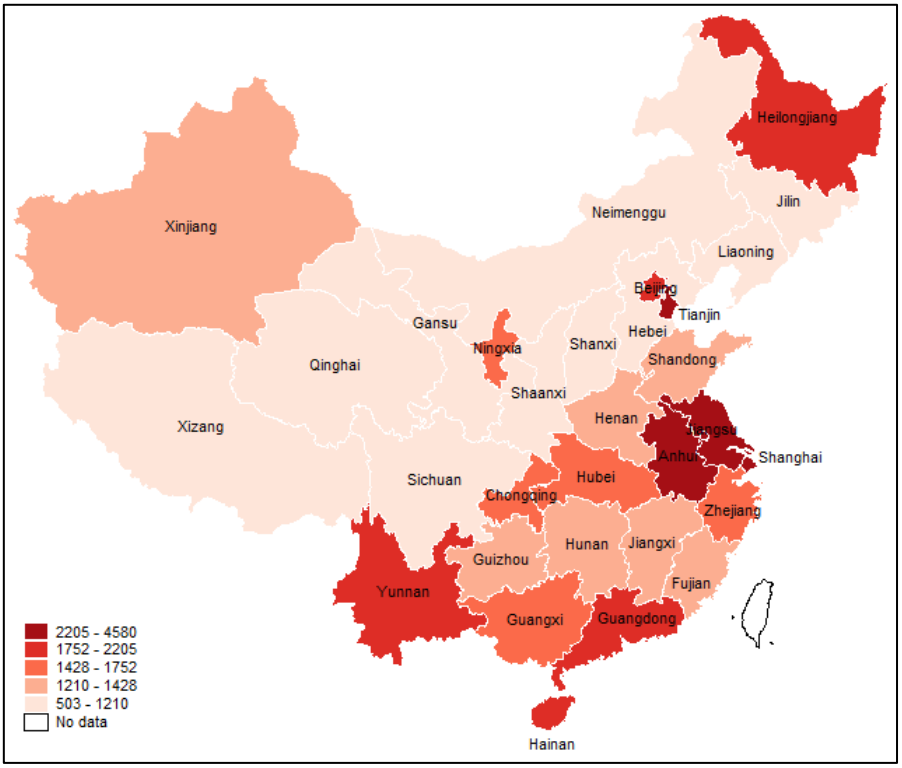

Figure 3-20 How many inhabitants share one healthcare institution in China, 2018

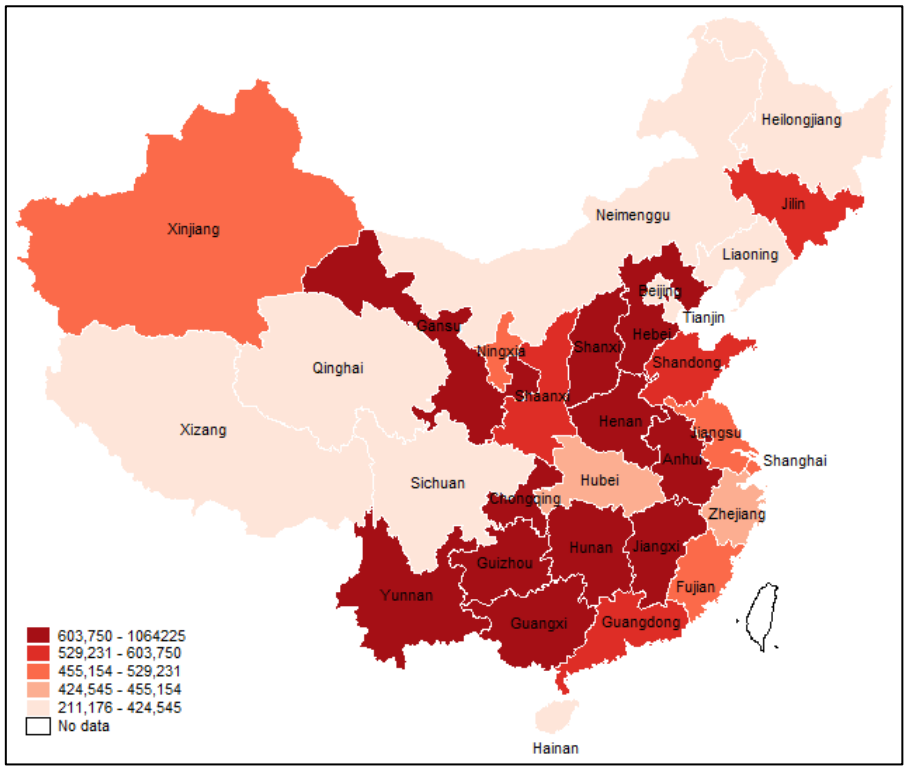

Figure 3-21 How many inhabitants share one tertiary hospital in China, 2018

For total medical personnel and sickbeds, in 2018, the east and northeast had a relative abundance of doctors, while the east and some parts of the central region had more nurses. Beijing led in medical personnel density with 11.9 per 1,000 inhabitants, followed by Zhejiang and Shaanxi with 8.5 each. Conversely, Jiangxi (5.3), Anhui (5.3), and Xizang (5.5) ranked as the top three regions with scarce medical personnel. More detailed, for licensed doctors and assistants, Beijing (4.6), Zhejiang (3.3), and Shanghai (3) were the top three abundant regions, whereas Jiangxi (1.9), Anhui (2), and Yunnan (2.1) faced scarcity. Regarding nurses, Beijing (5), Shanghai (3.6), and Shaanxi (3.6) were the most abundant, while Xizang (1.6), Hebei (2.3), and Gansu (2.4) were the scarcest. In terms of sickbeds, Liaoning (7.21), Xinjiang (7.19), and Sichuan (7.18) were abundant, whereas Tianjin (4.37), Guangdong (4.56), and Hainan (4.8) were scarce. (Figures 3-22/23/24)

International comparisons also underscore the regional imbalance of healthcare resources in China. Beijing's doctor ratio, at 4.6, only fell short of Norway's ratio (4.66) but exceeded those of other OECD countries (For example, Germany 4.25, UK 2.81, USA 2.61). However, no

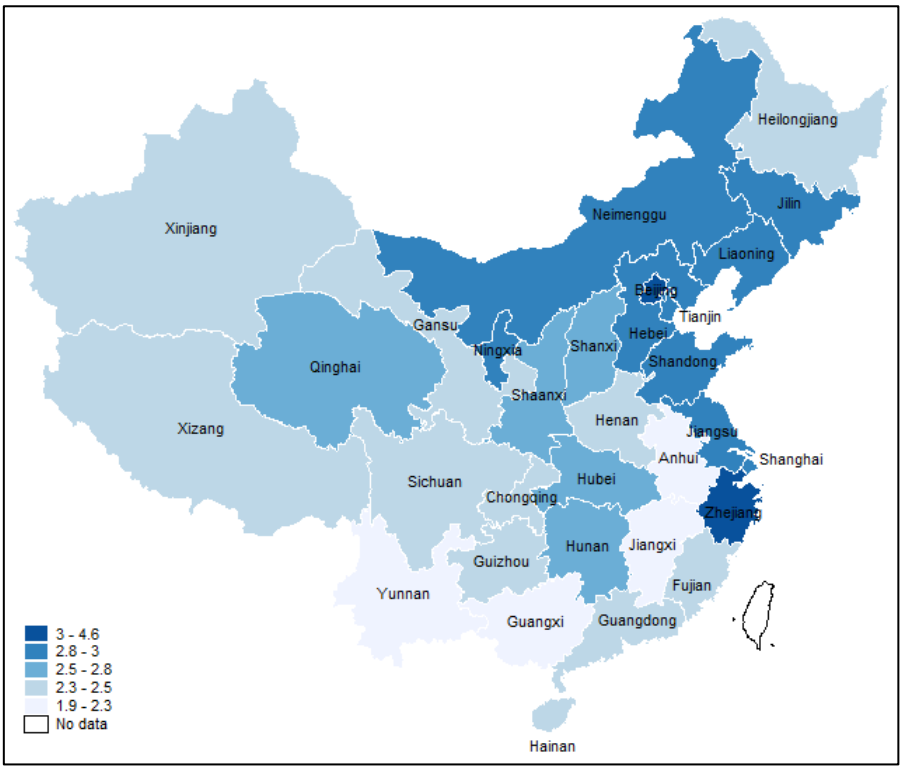

Figure 3-22 How many licenced doctors (assistant) per 1,000 inhabitants, 2018

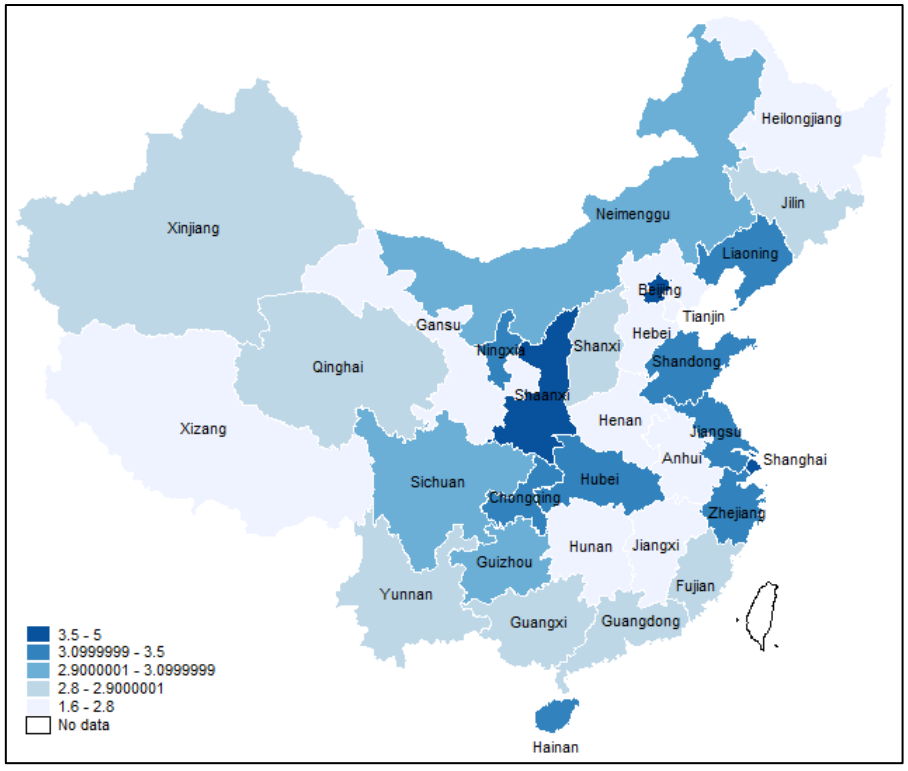

Figure 3-23 How many registered nurses per 1,000 inhabitants, 2018

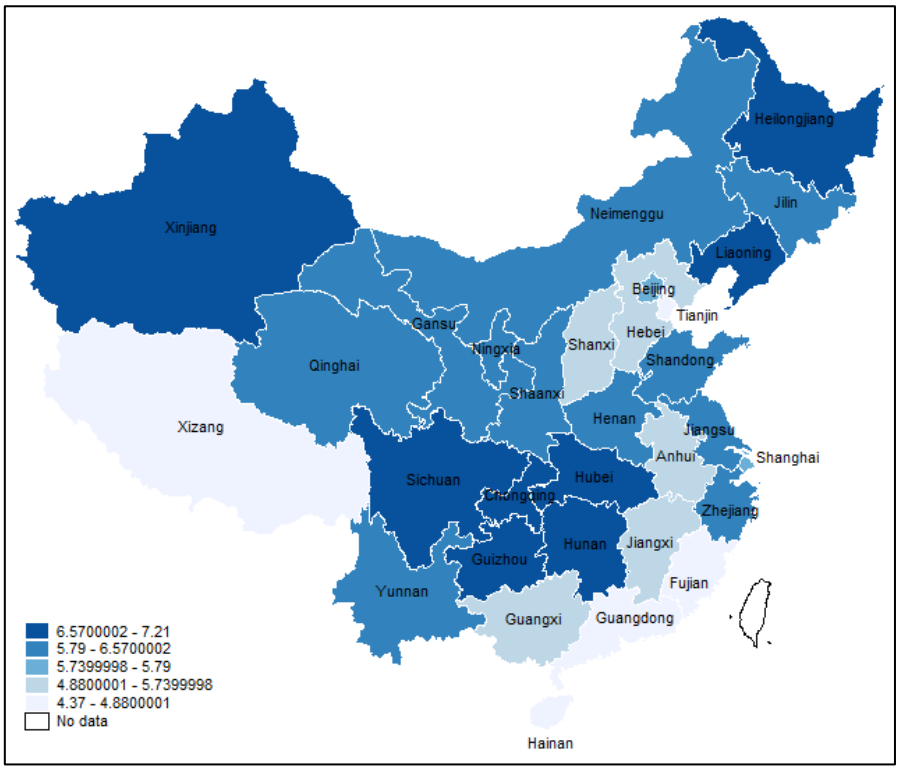

Figure 3-24 How many sickbeds per 1,000 inhabitants, 2018

other Chinese region reached the OECD countries' average (3.37). For nurses, despite Beijing having the highest nurse ratio at 5, it still fell below the OECD average of 8.55, with Hebei (2.3) and Xizang (1.6) even lower than Turkey (2.07), the country with the smallest nurse ratio in the OECD. The sickbed ratio in China performed well, with only Guangdong (4.56) and Tianjin (4.37) falling short of the OECD average (4.73).

3.5.3 Health Service Use and Fee

In 2018, healthcare institution visits were more frequent in the south-eastern coastal regions and less so in the northern regions. Conversely, hospitalisation rates were higher in the central and western regions, while the eastern regions experienced comparatively lower rates. Outpatient hospital visit costs were elevated in the north-eastern, eastern, and south-central regions, but lower in the central and western regions, with a similar distribution pattern of hospitalisation costs.

More specifically, the top three regions with the highest number of outpatient visits (per year) were Shanghai (11.15), Zhejiang (10.94), and Beijing (10.92), while the top three regions with the lowest number were Heilongjiang (2.96), Shanxi (3.49), and Hunan (3.9). The top three regions with highest annual hospitalisation rates were Chongqing (22.7%), Guizhou (22.6%), and Hunan (22.3%), while the top three lowest were Tibet (9%), Tianjin (10.4%), and Hainan (12.8%).

The top three highest outpatient costs (per visit) were in Beijing (545 CNY), Shanghai (379 CNY), and Tianjin (339 CNY), and the top three lowest were in Tibet (183 CNY), Gansu (188 CNY), and Henan (194 CNY). The top three highest hospitalisation costs were in Beijing (22,619 CNY per patient), Shanghai (18,390 CNY), and Tianjin (17,148 CNY), and the top three lowest were in Guizhou (5,731 CNY), Gansu (5,804 CNY), and Yunnan (6,423 CNY). (Figure 3-25/26/27/28)

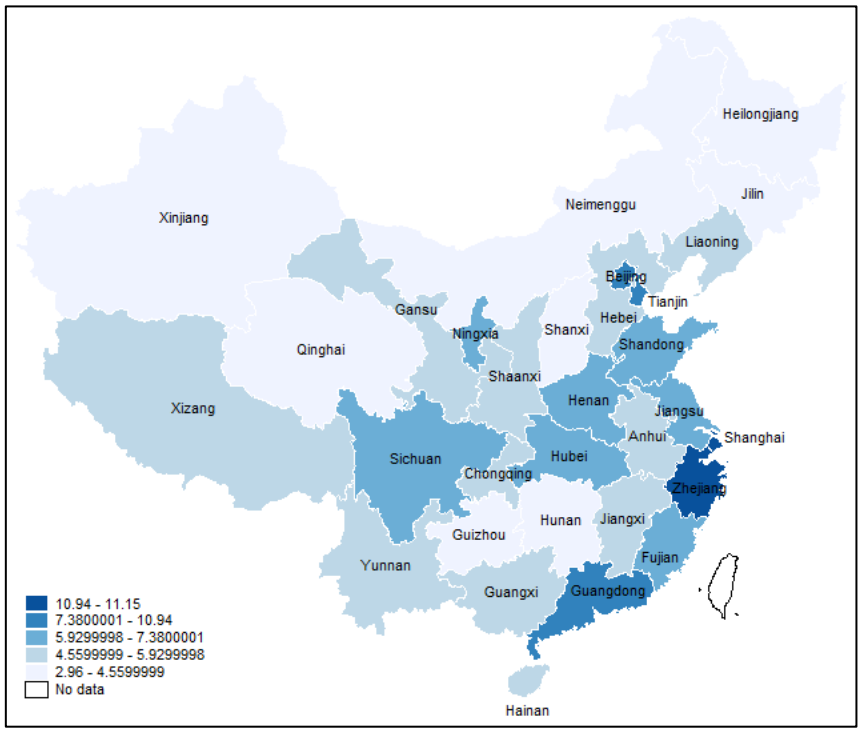

Figure 3-25 The provincial distribution of annual visits times in China, 2018

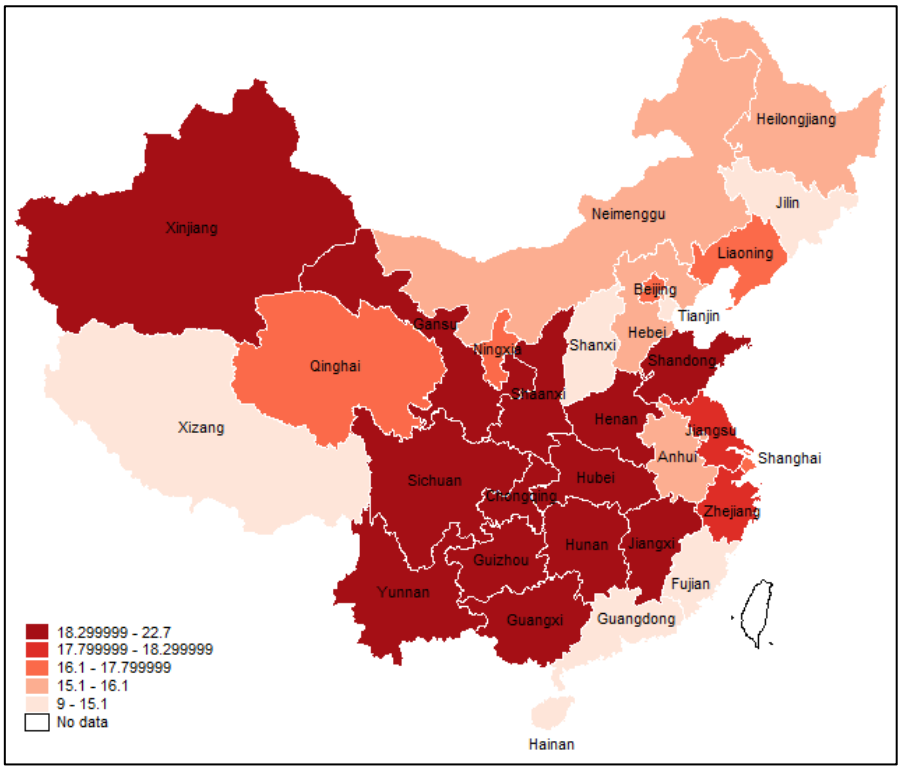

Figure 3-26 The provincial distribution of annual hospitalisation rate (%) in China, 2018

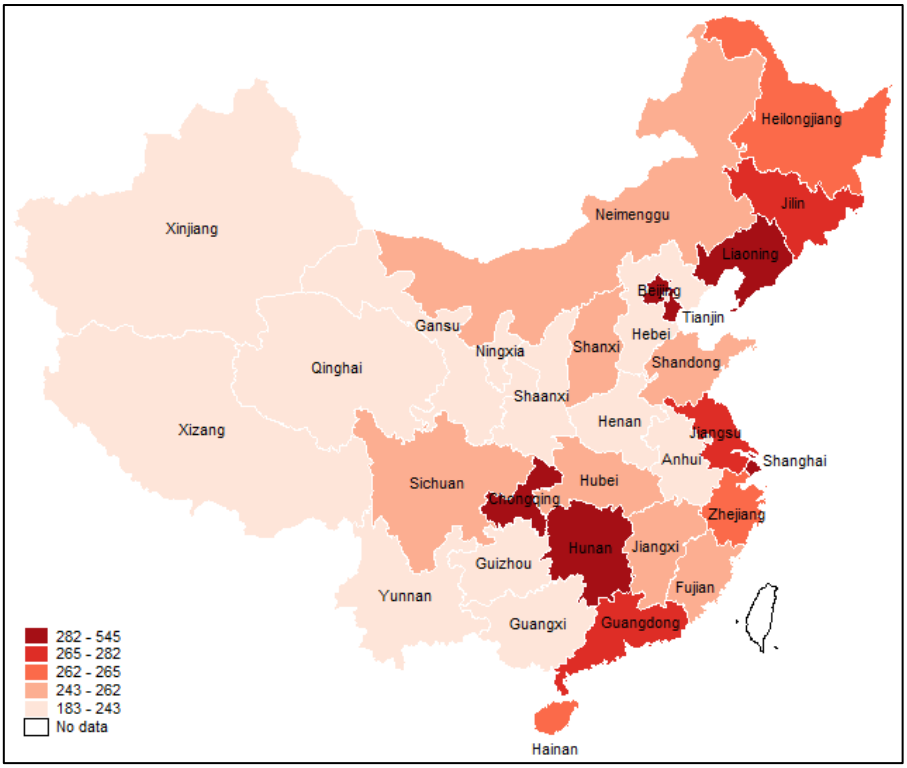

Figure 3-27 The provincial distribution of outpatient costs (CNY) in China, 2018

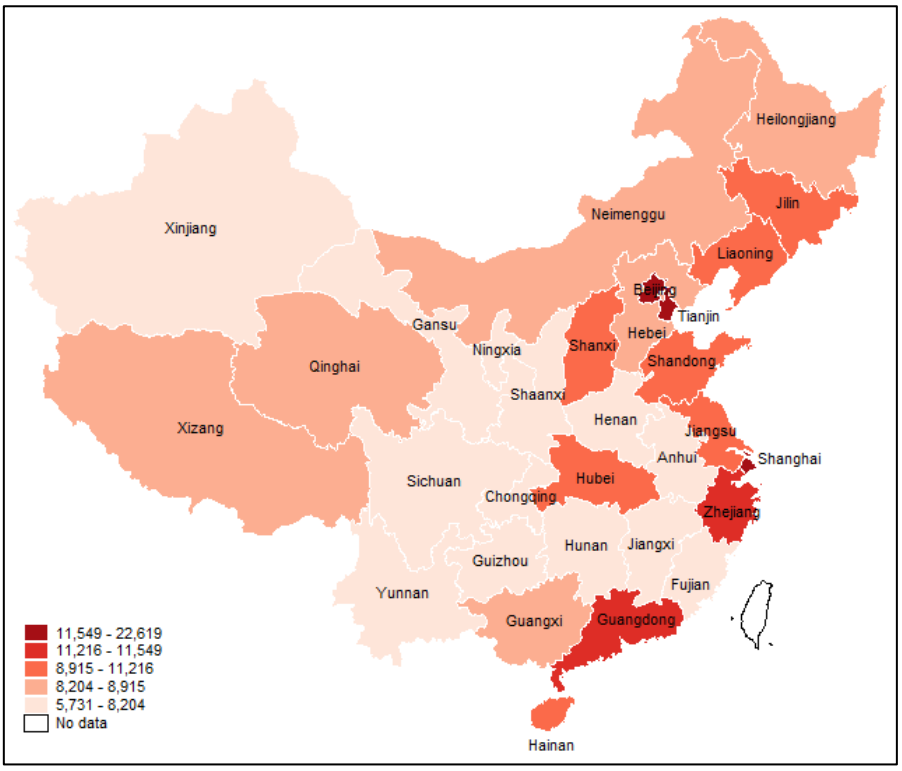

Figure 3-28 The provincial distribution of hospitalisation costs (CNY) in China, 2018

3.6 Discussion and Conclusion

This chapter uses various indicators to illustrate trends and regional differences in China. The country is experiencing rapid ageing and a decline in fertility, which heralds a future population crisis and an inevitable rise in medical expenses. Despite rapid economic development, China still lags behind developed countries. Meanwhile, although there is no proposal for "right" levels of health expenditure, increased spending on healthcare is linked to better health outcomes, particularly for developing countries. While the average medical expenses are affordable, the escalating medical burden for rural residents warrants vigilance.

From a medical resources supply perspective, the overall number of institutions remains stable. Notably, the large increase in hospitals and tertiary hospitals reflects improved access to quality medical resources. With the increase in the number of hospitals, the number of doctors, nurses and sickbeds also increased. The ratio of nurses to doctors exceeded 1:1 after 2013. The per capita data also shows that the accessibility of medical resources, especially high-quality

medical resources, increased. However, internationally, only the data on sickbeds has reached the level of developed countries, and the data on doctors and nurses per capita still lags far behind the level of developed countries. With the increase in medical resources supply, people's medical needs have been released. The rapid growth of outpatient and hospitalisation services has been accompanied by an increase in average medical expenses.

The provincial disparities in China are significant: the economic advantages of the eastern and southern coastal regions are clear, which is in contrast to the western and northern inland regions. Beijing, the most economically developed region in China, has a GDP per capita comparable to Greece, which ranks in the bottom tenth of OECD countries. Healthcare expenditure is higher in the eastern coastal regions, but given income levels, the burden of healthcare is heavier in the west. Due to the smaller population in the West, medical facilities are relatively plentiful, while the central region is extremely short of them. The number of healthcare workers is highly correlated with the economy, with the more economically developed regions having a higher density of healthcare workers. Economic conditions are positively correlated with outpatient visits and negatively correlated with hospitalisation visits, with the southeast coast having more outpatient visits and the central west having a higher rate of hospitalisation.

As can be seen from the above, there are dramatic changes in China's economy, population and health services over the last decade or so, as well as significant regional variations. These are likely to have had a significant impact on the implementation of health insurance integration. For example, economically developed regions have more adequate budgets, so they are more likely to implement integration earlier or provide better financial subsidies and reimbursement terms after integration.

In conclusion, this chapter provides four key ideas for later quantitative analysis:

1) Time effect must be considered, which covers factors that cannot be directly controlled by microdata, such as the increase of medical demand caused by ageing, the change in health service supply and price caused by economic development, etc;

2) It is necessary to separate the analysis of urban and rural areas because the rural economy and medical resources are far worse than those of cities
3) Disparities among cities or provinces must be considered, which relate to the changes in the supply and demand of health services within a region;
4) The implementation of integration is more likely to be planned than randomised.

The next chapter will review the theories related to health services and health insurance (such as the theory of supply and demand in economics). Based on these theories, the framework of empirical analysis can be initially built.

Chapter 4 The Review of Theories

In economics and the health field, there are many theories around health, such as the demand and supply of healthcare, insurance, equity, and equality. The inception of health economics in 1963 is marked by Kenneth J Arrow's influential paper, "Uncertainty and the Welfare Economics of Medical Care" (Arrow, 1963). In his paper, Arrow explained the unique characteristics of the medical care market, highlighting aspects such as the uncertainty of disease occurrence, the uncertainty of health services, information asymmetry, and the issue of moral hazard.

Since this thesis focuses on the impact of the integration of health insurance policy on the use of health services and expenses, this chapter selectively reviews five theories related to health service and health insurance: supply and demand theory, Grossman's health demand model, Andersen's healthcare behavioural model, adverse selection, and moral hazard. The first three theories explore factors influencing health service use and expenses, while the latter two discuss specific concerns related to health insurance. A comprehensive examination of these theories aids in delineating the variables suitable for empirical analysis and refining the research assumption.

The structure of this chapter unfolds as follows: Section 4.1 discusses the general theory of supply and demand, along with health service theory, while Section 4.2 addresses health insurance issues. The concluding section 4.3 offers a brief discussion of all reviewed theories and a summary of the entire chapter.

4.1 General Demand and Healthcare Demand

This section begins by briefly outlining the general demand theory and its influencing factors. Subsequently, it introduces two theoretical framework models closely related to health and the use of health services: the Grossman model and the Andersen model.

4.1.1 General Demand and Consumer Theory

In microeconomics, the "quantity demand" for a particular good is the amount that an individual is willing and able to purchase (Mankiw, 2020). This quantity is mainly influenced by two factors: individual needs and the ability to pay. More specifically, it is affected by price, income, preference, expectation, the number of buyers, etc. On the other hand, the "quantity supply" for a good represents the amount that the sellers are willing and able to sell. This quantity is affected by price, input, technique, etc (Mankiw, 2020). The interplay between supply and demand establishes the equilibrium price and quantity of a good. Therefore, from a fundamental economic perspective, an individual's health plays a crucial role in determining their health service needs, which subsequently influence healthcare use, alongside the individual's economic capacity, typically measured by income.

The elasticity of demand is an important concept in economics that measures the responsiveness of demand to changes in factors such as income or price (Mankiw, 2020). Goods with high price elasticity of demand experience greater changes in demand when prices change, compared to goods with low price elasticity of demand. Newhouse and Phelps (1974) were the first to measure the price and income elasticity of outpatient visits and hospitalisation days, indicating that health services are inelastic in price and income. Zheng *et al.* (2017) reviewed elasticity studies in China, indicating that common health services are inelastic, and elasticities are different in outpatient and hospitalisation services, rural and urban areas, different hospital levels, and different geographic locations (eastern/central/western).

According to the above, in this thesis, when analysing the impact of health insurance on health services, an individual's health status and income are control variables that must be included. And the analysis should also separately consider outpatient and hospitalisation services, as well as urban and rural areas.

In addition, as early as 1961, economist Schultz proposed that health is human capital (Schultz, 1961). According to human capital theory, human capital consists of health, knowledge, skills, and work experience. Mushkin (1962) first pointed out that health and education are equally

important in human capital. Nevertheless, unlike other forms of human capital, health is the most basic human capital and is the basis of all other forms of human capital. Without health, the investment efficiency of human capital will decrease and even lead to the loss of other human capital. Everyone acquires an initial health stock through inheritance, which depreciates with age but increases with health investment (Grossman, 1972).

Healthcare is the way to maintain or improve health, including the prevention, diagnosis, and treatment of disease, illness, injury, and other impairments. There are two main theoretical frameworks for people's demand/choice of healthcare. In economics, Grossman Health Demand Model first established the connection between health demand and healthcare demand, and analysed impact factors of healthcare demand; in health science, Andersen's Behavioural Model elaborated on societal and individual determinants of health service use. Numerous studies around the world, including China, are derived from these two frameworks.

4.1.2 The Grossman Health Demand Model

The Grossman Health Demand Model, originally proposed by Michael Grossman in 1972, forms the basic model of the health demand grounded in human capital theory (Grossman, 1972). Grossman posited that direct attainment of health is not possible; rather, it occurs indirectly through access to healthcare. Therefore, healthcare demand is the "derivative demand" of health demand. The model examined the impact of age, working hours, income, health service prices, and education on healthcare demand.

Over the past five decades, the Grossman model has found widespread application in economic empirical studies and has also undergone reformulation and extension by various researchers (Muurinen, 1982; Wagstaff, 1986, 1993; Nocera and Zweifel, 1998; Ried, 1998; Jacobson, 2000; Zweifel, 2012). For example, a defect of the Grossman model is that it does not involve the individual's choice of different healthcare institutions and services. Gertler and Mwabu addressed this deficiency by introducing factors such as quality of health service and travel time into the Grossman model (Dor, Gertler and van der Gaag, 1987; Mwabu, Ainsworth and Nyamete, 1993). In China, many empirical studies have examined the feasibility of the

Grossman model. In Urban areas, the distribution of residents' health status is relatively uniform; women's education has a positive impact on health, men's education level has no significant impact; age has a greater impact on men's health than women's, income or wages have no significant impact on health; overall, it is confirmed that women are more in line with the Grossman model prediction than men in China (Liu *et al.*, 2004; Zhao and Hou, 2005).

Therefore, when analysing the policy impact of health insurance on health services in this thesis, factors such as age, sex, income, and education must be considered as essential control variables, drawing from the Grossman Model and its derivative research.

4.1.3 The Andersen's Behavioural Model of Healthcare

Another widely adopted framework is Andersen's behavioural model of health service utilisation (Andersen and Newman, 1973). This model categorises individual determinants of health service utilisation into three components: predisposing, enabling, and illness level. Predisposing components influence patient preferences for health service use, even if they are not directly related to health services. These include demographics (such as age and sex), social structure (such as education and occupation), and attitudes and beliefs (which are about the patient's views of the disease and treatment; the more people trust the treatment, the more health service they will use). Enabling components indicate the ability of individuals to access health services. These include income, health insurance coverage, and the number of local healthcare institutions. The illness level reflects patients' judgments of their health problems, which is the most direct factor affecting the use of health services.

In contrast to the Grossman Model, Andersen's model considers both objective and subjective factors, analysing not only demand factors but also supply factors, such as the number of healthcare institutions available. Thus, this is a more practical model that not only analyses the demand of patients, but the actual choices patients face in real life. In China, a study based on the Andersen model showed that there were significant sex differences in the factors influencing healthcare costs for the rural elderly individuals. Male elderly medical expenses

depend more on predisposing components, while female elderly medical expenses depend more on enabling components (Song and Zuo, 2010).

Therefore, when analysing the impact of policies on health services, it is important to consider the number and level of health institutions[19] as control variable, alongside attitudes and beliefs.

4.2 Health Insurance

Health insurance serves as a mechanism to compensate for economic losses caused by disease. Essentially, it acts as a risk transfer mechanism. It functions by collecting premiums in advance and establishing health insurance funding pools. These funding pools are used to provide financial compensation when insured individuals incur health expenses due to illness in healthcare institutions.

The biggest difference among the health insurance market and other markets is the information asymmetry among healthcare institutions, insured individuals, and insurance institutions, which leads to market failure. The market failure of the health insurance market is mainly manifested in adverse selection and moral hazard (Arrow, 1963). Although adverse selection and moral hazard are common problems in private insurance research, they may be not considered as problems in this study. This is because the implementation of the integration policy is not determined by the individual, that is, the policy is "mandatory" for the individual. Therefore, these two issues are only briefly described.

4.2.1 Adverse Selection

Adverse selection arises due to information asymmetry between patients and insurance institutions. Individuals in poor health are more inclined to buy health insurance, while individuals in good health are less likely to do so (Arrow, 1963). If the number of high-risk

[19] In China, there are two rules for the level of healthcare institutions. One is the administrative level, such as the village clinic/town-level/city-level/provincial-level hospital. The best level (highest quality) is provincial. The other is the hospital level, in practice, it includes 3 grades, that is, the primary, the secondary, and the tertiary, and 3 subsidiary levels, A, B, C. The best level is Tertiary A.

patients in the insured population is higher than the number of low-risk patients, the reimbursement rate will be very high.

Adverse selection usually occurs in private insurance markets because the insurance purchase is optional, but not in a social insurance background because the insurance purchase is usually compulsory. In China, the adverse selection issue may occur because the resident medical insurances are not compulsory. However, the Chinese government addresses this issue through family enrolment, requiring entire families to enrol together (CHEN, 2016).

In this study, adverse selection is not considered a problem in analysing the policy impact of insurance integration. The reason is that China has almost achieved UHC, which means, almost everyone has insurance. Even if the integration affects people's willingness to participate, the number of people affected is very small.

4.2.2 Moral Hazard

Moral hazard is another important market failure in the health insurance market, which includes ex-ante and ex-post moral hazard (Arrow, 1963). Ex-ante moral hazard refers to individuals reducing their efforts on activities that can prevent diseases or reduce the occurrence of diseases. For example, after having insurance, individuals may reduce their daily physical exercise. Ex-post moral hazard refers to the excessive use of health services because insurance reduces the marginal cost of health services.

The design of any health insurance system must take into account the two basic functions: dispersing disease risk and preventing moral hazard (Arrow, 1963). Therefore, the optimal theory of insurance design is produced. It mainly discusses how to balance the social benefits caused by risk-sharing and the social welfare loss caused by moral hazard (Zeckhauser, 1970; Arrow, 1971, 1976; Cutler and Zeckhauser, 2000). Traditional health insurance design focuses on avoiding excessive health service utilisation, incorporating features like deductibles, co-payments, and ceilings. Deductibles ensure patients cover expenses up to a certain amount, with

insurance reimbursing amounts beyond that. Co-payment rate (reimbursement rate)[20] refers to the reimbursable expenses that are paid by the insurance institution and individual together according to the agreed proportion rather than fully reimbursed by the institution. Ceiling refers to the upper limit of the institution's reimbursement, above which the excess expenses will not be reimbursed. Previous studies proved that diseases with low health service price elasticity should have higher reimbursement rates, such as hospitalisation and emergency, so health insurance should give priority to compensating critical illness (Feldstein and Friedman, 1977; Besley, 1988; Feldman and Dowd, 1991; Manning and Marquis, 1996; Blomqvist, 1997).

Although a higher co-payment ratio can control resource waste caused by moral hazard, the inability of patients to choose the best services and insufficient awareness of benefits and risks may lead to the underuse of services, especially preventive health services which have a high price elasticity. Therefore, some scholars challenge the traditional optimal design. Nyman argues that not all moral hazards are welfare losses and that the extra consumption caused by income transfer is welfare neutral or welfare increase (Nyman and Maude-Griffin, 2001). In the context of Chinese health insurance, Huang and Gao (2010) illustrated that older people with insurance have a life expectancy of five years longer than those without insurance. The increased medical costs and healthcare use due to insurance are therefore effective demand rather than excessive demand since they help prolong life expectancy in China.

4.3 Discussion and Conclusion

This chapter reviewed five economics and healthcare theories relevant to this thesis research: supply and demand theory, Grossman health demand model, Andersen healthcare behavioural model, adverse selection, and moral hazard. Jointly, these theories help establish the analytical framework for this thesis, identifying some important variables, and shaping analytical assumptions.

[20] Co-payment rate and reimbursement rate are from different perspectives. The co-payment rate is the proportion of expenses that the patient ought to pay, and the reimbursement rate is the proportion of the expenses that the insurance institution ought to pay. The sum of the two is 100%.

The supply and demand theory in economics, when applied in the health field, highlights health status as the pivotal role influencing the demand for health services. Self-rated or self-reported health has been commonly used in previous health research in China to reflect individuals' health status (Ma, Zhao and Gu, 2016; Liu and Lin, 2018; Su *et al.*, 2019). However, a study (Mu, 2014) has raised concerns that individuals in different regions may use varying criteria when assessing their health, potentially leading to an underestimation of health differences across regions. Consequently, some previous studies (Huang and Zhang, 2017; Liu, 2017) have incorporated specific diseases, such as cancers and chronic diseases, into their analyses as influencing factors for health service use. Therefore, in the empirical analysis of this thesis, whenever feasible, both measures of self-rated health and specific diseases will be employed to comprehensively examine the impact of individual health status on health services and costs.

The elasticity, an important concept in supply and demand theory, indicates degrees of responsiveness to changes in price or income for different goods. Previously, Zhou *et al.* (2011) found that in rural China, the income elasticity of the probability of an outpatient visit is not significant, but the income elasticity of the probability of hospitalisation is statistically significant; Zheng *et al.* (2017) also highlighted that elasticities are different in outpatient/inpatient services, rural/urban areas, different hospital levels, and different geographic locations (eastern/central/western) in China. This heterogeneity means that in empirical research, outpatient and inpatient services, as well as urban and rural areas, should be analysed separately. Consequently, this thesis research, consistent with prior research, advocates for separate analyses of outpatient/inpatient services and urban/rural areas, a practice that will be highlighted in the subsequent empirical chapters of this thesis.

Two theoretical models of healthcare are reviewed subsequently. Grossman's model, rooted in basic economic theory, extends its analysis of health service use by incorporating demographic and social factors such as age, education, and social factors alongside income, price, and health status (Grossman, 1972). Although this model initially lacked consideration of the supply side of health services, such as the quality and quantity of health services, it was extended by subsequent researchers (Dor, Gertler and van der Gaag, 1987; Mwabu, Ainsworth and Nyamete,

1993). In addition, Andersen's model offers a more comprehensive analytical framework encompassing both supply-side and demand-side factors, as well as subjective and objective aspects of health services (Andersen and Newman, 1973). However, the measurement of subjective factors, such as patients' beliefs and attitudes, may pose challenges in practice. In the subsequent empirical research review chapter, I will discuss in detail how previous studies have applied these two theoretical models and will also refer to these studies to construct a robust empirical analysis framework in this thesis.

Two pertinent theories to health insurance are finally reviewed: adverse selection and moral hazard. The adverse selection highlights the voluntary purchase of insurance in a free market based on individual circumstances, leading to an imbalance in the health status of the insured population and failed risk sharing (Arrow, 1963). However, as insurance integration is mandatory for residents, theoretically, adverse selection is not considered to be an issue. Moral hazard concerns the potential for excessive medical treatment and resource waste due to insurance (Zeckhauser, 1970; Arrow, 1971, 1976; Cutler and Zeckhauser, 2000). However, some researchers argue that changes caused by moral hazard may be neutral or welfare-increasing and this has been shown to be the case in the Chinese context (Nyman and Maude-Griffin, 2001; Huang and Gao, 2010). Therefore, a key assumption in this thesis is made: "Insurance-induced increases in health service use and costs are efficient (welfare-increasing) rather than excessive (welfare-losing)".

In conclusion, through this chapter, three key points of this thesis research are drawn out:

1. Separate analyses of outpatient and inpatient services, as well as distinctions between urban and rural areas, are recommended, allowing for potential differences in their price and income elasticity of demand.
2. To better identify integration effects, it is advisable to include demographic information, socioeconomic background, and health resource abundance in empirical applications, according to Grossman and Anderson's models.
3. Adverse selection and moral hazard, typical concerns in insurance research, are not applicable in this thesis research due to the "mandatory" nature of integration. However, a

crucial assumption is made: the increase in healthcare use and costs resulting from integration is deemed "welfare-increasing" rather than "welfare-losing".

The next chapter will review statistical methods for causal identification and regression analysis, establishing a clear and detailed empirical analysis framework.

Chapter 5 The Review of Statistical Methods

This chapter builds upon the theoretical discussion in the previous chapter and further explores how to use statistical methods to quantify the impact of insurance integration on healthcare. Chapters 2 and 3 have explained that the implementation of insurance integration is not random; rather, it is influenced by factors such as local population, economy, finance, and medical resources. This lack of randomness may introduce potential selection bias issues. Therefore, the commonly used method, linear regression, may only uncover association, not causation, between integration and healthcare outcomes. Section 5.1 addresses this challenge by providing an in-depth examination of causal identification strategies. It introduces crucial statistical concepts, briefly discusses common methods, and thoroughly explores the chosen method for this thesis: Difference-in-Difference (DID).

In addition, many outcomes in healthcare research, such as binary outcomes (e.g., whether outpatient care occurred) and count outcomes (e.g., the number of hospital visits), are non-normally distributed. This also makes the linear regression model inappropriate. Thus, Section 5.2 explores various analytical models. It initiates by discussing the use of different parametric models based on the type of outcome variable and then illustrates the application of non-parametric models. Through all the above methodological reviews, the empirical analysis framework applicable in this thesis is clarified in Section 5.3, the conclusion.

5.1 Causal Identification Strategy

In general, policy evaluation serves three purposes: 1. To evaluate the impact of an experienced and documented policy on outcomes, sometimes also called treatment evaluation; 2. To forecast the impacts of an existing policy in other environments, such as different groups of people; 3. To forecast the impact of a never-historically experienced policy (Heckman and Vytlacil, 2007). This study focuses on the first scenario, that is, to evaluate the impact of integration insurance on healthcare use and expenses. The section initially outlines the methodology for estimating the causal effect of a policy and subsequently addresses challenges associated with relying

solely on observational data for the study. Various methods to mitigate these challenges are then discussed.

5.1.1 The Treatment Effect and Selection Bias

The theoretical basis of treatment evaluation is the Rubin counterfactual framework (Rubin, 1974). In this framework, an individual has two potential outcomes: one is the outcome if the individual receives treatment, and one is the outcome if the individual does not receive treatment. These two outcomes are termed the counterfactual outcomes of each other. Denote a dummy variable indicating whether individual *i* received treatment as A, and the outcome of interest as *y*. Then for individual *i*, *yi* has two potential outcomes, that is:

$$y_i = \begin{cases} y_{1i} \; if \; A_i = 1 \\ y_{0i} \; if \; A_i = 0 \end{cases}$$

Rubin defined that, the difference between the two potential outcomes is the treatment effect of the treatment on this individual, that is, $y_{1i} - y_{0i}$. Thus, at the population level, the Average Treatment Effect (ATE) is defined as the mean (unconditional expectation) difference between observed and counterfactual outcomes, that is:

$$ATE = E(y_{1i} - y_{0i})$$

However, in reality, it is impossible to simultaneously observe the treated and untreated outcomes of an individual. It is only possible to directly compare outcomes between treated and untreated groups. This comparison is also problematic because there may be inherent differences between the two groups. The direct comparison of the difference in means between the two groups can be expressed as:

$$E(y_{1i}|A_i = 1) - E(y_{0i}|A_i = 0)$$

This can also be rewritten as a two-part sum form:

$$[E(y_{1i}|A_i = 1) - E(y_{0i}|A_i = 1)] + [E(y_{0i}|A_i = 1) - E(y_{0i}|A_i = 0)]$$

The first part is the difference between the expectation of the true and counterfactual outcomes of treated individuals, known as the Average Treatment Effect on Treated (ATT):

$$ATT = E(y_{1i} - y_{0i}|A_i = 1) = [E(y_{1i}|A_i = 1) - E(y_{0i}|A_i = 1)]$$

The second part is the inherent difference between treated and untreated individuals, which is called "Selection Bias". Therefore, the direct comparison of means between the two groups only measures the association between treatment and outcome but cannot establish causation due to the inclusion of outcome differences unrelated to treatment.

To address the selection bias issue, one way is to randomly assign the treatment between the individuals, or collect data through a Randomised Controlled Trial (RCT). In an RCT, since the treatment A_i is independent of (y_{1i}, y_{0i}), the conditional expectation equals the unconditional expectation, that is:

$$ATT = E(y_{1i} - y_{0i}|A_i = 1) = E(y_{1i} - y_{0i}) = ATE$$

And the selection bias is equal to zero, because:

$$E(y_{0i}|A_i = 1) - E(y_{0i}|A_i = 0) = E(y_{0i}) - E(y_{0i}) = 0$$

Thus, in this case, the direct comparison of the outcomes between treated and untreated individuals is ATE, that is:

$$E(y_{1i}|A_i = 1) - E(y_{0i}|A_i = 0) = ATT + 0 = ATE$$

The above explanation, using mathematical expectation, clarifies the issue of selection bias and the feasibility of estimating ATE through RCT. The RCT-based ATE estimate can also be explained through a regression model, which also facilitates the exploration of other methods later. Consider the regression model:

$$y_i = \alpha + \beta A_i + \varepsilon \quad \textbf{[1]}$$

The predicted mean of treated individual is: $\hat{y}_{treated} = \hat{\alpha}_{ols} + \hat{\beta}_{ols}$

The predicted mean of untreated individual is: $\hat{y}_{untreated} = \hat{\alpha}_{ols}$

Therefore, the mean difference is: $\hat{y}_{treated} - \hat{y}_{untreated} = (\hat{\alpha}_{ols} + \hat{\beta}_{ols}) - \hat{\alpha}_{ols} = \hat{\beta}_{ols}$

The ordinary least squares (OLS) estimator $\hat{\beta}_{ols}$ is also called "Differences estimator". Angrist and Pischke (2009) have also emphasised that linear regression models require random assignment of treatment variables for causal interpretation.

However, RCT data is not available as the integration implementation is not randomised. Consequently, the analysis in this thesis will rely on secondary/observational databases. When

working with observational data, selection bias can be caused by many variables, which are called “confounders”. Some confounders are observable, and some are unobservable.

Observables are those variables that are already known to impact outcomes and treatment and can be directly involved in analyses. For example, age, which influences healthcare use and expenses, is likely to affect integration implementation. Governments in cities with more elderly people may be less willing to integrate than in cities with more young people because the increase in fiscal pressure caused by integration will be greater. Demographic variables like these are typically collected into databases and can be directly controlled in analyses.

However, some confounders may not be directly involved in analyses, either because they are not collected into the database or not measured. Unobservable confounders in a health insurance integration study may include variables like the number of local healthcare institutions, the price of local health services, or the economic level. While not usually collected, these variables are likely affecting both the implementation of the integration and the outcomes. Statistically, unobserved confounders are also called the “Omitted Variable”.

In observational studies, when individuals are selected for treatment solely based on observed confounding variables, linear regression models can be used to estimate the ATT (Jones and Rice, 2011). This estimation relies on the Conditional Independence Assumption (CIA), assuming that, after accounting for control variables, potential outcomes are independent of the treatment status. In simpler terms, it assumes treatment assignment is comparable to a random process. Furthermore, the estimation also assumes that the conditional mean of the counterfactual outcome can be modelled as a linear function of the observed confounding variables (Jones and Rice, 2011). However, in healthcare research, many outcome variables exhibit non-normal and skewed distributions, such as the number of visits and cost of visits, challenging the linearity assumption. Therefore, Propensity Score Matching (PSM) is recommended as it doesn’t need a parametric model for the outcome. In addition, if unobservable confounders are important, other approaches, such as Difference-in-Difference and instrumental variables, are needed.

5.1.2 Common Approaches

Common approaches[21] to deal with selection bias in observational data include the Propensity Score Matching (PSM), Instrumental Variable (IV), Regression Discontinuity (RD) design, Difference in Difference (DID), and Synthetic Control Method (SCM). These causal identification approaches are widely used in health research in China, some studies employed IV and FE (Wagstaff and Lindelow, 2008; Lei and Lin, 2009; Liu and Zhao, 2014; Ma, Gu and Sun, 2015; Ma, Zhao and Gu, 2016; Su *et al.*, 2019); some used PSM with DID (Wagstaff *et al.*, 2009; Zhou, Zhu, *et al.*, 2014; Cheng *et al.*, 2015; Liu, 2017).

Table 5-1 provides a summary of these approaches. PSM can be used to estimate the treatment effect if there are selections on observables only. Unobserved confounders are categorised into time-invariant and time-varying. DID can solve those time-invariant unobserved confounders; IV, SCM, and RD can solve both time-invariant and time-varying unobservables. Based on the data type used in this thesis (panel data and repeated cross-sectional data, with details provided in empirical chapters) and pre-analysis results (for example, the instrumental variable is proven unnecessary), the most appropriate method is DID. Thus, this section briefly introduces approaches other than DID, while the idea of DID estimation, validity testing of DID, and extended approaches of DID will be discussed in detail in subsection 5.1.3.

Propensity Score Matching (PSM)

The match method is based on matching the individuals in the control group with the individuals in the treatment group according to some rules to obtain two comparable (similar characteristics) groups. Matching methods in research typically involve exact and inexact matching. Exact matching refers to pairing individuals directly based on specific key control variables. This approach is practical when the number of matching variables is limited. However, if too many variables are used, although the accuracy of matching is higher, more unmatched individuals

[21] Although there are other more sophisticated methods which combine the estimation of treatment effect and the assignment of treatment, such as Heckman structural models, control function, and correction approach, they require a higher level of expertise and may be less practical in the health insurance integration policy evaluation. Therefore, this study does not involve them.

Table 5-1 Summary of policy evaluation methods

Methods	Ideas	Solved confounders	Assumptions/Requirements
PSM	Matching treatment and control group according to a propensity score	Observables	The Conditional Independence Assumption (CIA), Common Support Condition, and Stable Unit Treatment Value Assumption (SUTVA)
DID	Twice differential operations eliminate baseline differences and time effects	Time-invariant unobservables only	The Parallel Trend Assumption, the Stable Unit Treatment Value Assumption (SUTVA), and No Anticipation Assumption; sometimes, it also requires the linearity assumption
IV	Decomposing the part of treatment variable that is unrelated to error term by two-stage regression	Both time-invariant and time-varying unobservables	An instrumental variable should be related to treatment but not to error term; it is difficult to find a good IV
SCM	Using multiple groups to synthesize a more appropriate control group	Both time-invariant and time-varying unobservables	Requires more units in the control group than units in the treatment group; more suitable for macro indicators
RD	Looking for a grouping variable can determine whether to receive treatment	Both time-invariant and time-varying unobservables	It requires clear rules of grouping and the grouping variable has a cut-off point

Abbreviation: PSM-Propensity Score Matching; DID-Difference-in-Difference; IV-Instrumental Variable; SCM-Synthetic Control Method; RD-Regression Discontinuity

will be excluded and more sample size will be reduced. PSM, as an inexact matching approach, converts multiple control variables into a single propensity score (it is defined as the probability of an individual receiving the treatment given the control variables) for matching and thus can help reduce the problem of controlling a large number of variables. Furthermore, PSM is intuitive and interpretable in creating similar treatment and control groups, and thus appealing to non-technical people (Zanutto, 2006; Caliendo and Kopeinig, 2008; Stuart, 2010).

The simplest PSM is the 1:1 nearest neighbour matching. In this matching approach, propensity scores are calculated for all individuals based on their control variables. Each individual in the treatment group was matched with another individual in the control group who had the closest propensity score (Wooldridge, 2012). However, this approach may produce poor matches if the nearest neighbour is too dissimilar. To address this limitation, an alternative approach called calliper matching is introduced. Calliper matching sets an acceptable range of propensity scores, within which untreated individuals can be selected as matches for treated individuals, but there is no uniform standard as to what distance is appropriate (Caliendo and Kopeinig, 2008). Apart from neighbour matching, there are two additional options: kernel matching and local linear regression matching. These methods take into account the weights of the outcomes for all untreated individuals, using nonparametric approaches. Kernel matching employs kernel functions, while local linear regression uses nonparametric local linear regression to determine the weights (Vinha, 2006).

The PSM needs to meet the CIA as well, and in addition, it also needs to satisfy Common Support Assumption and Stable Unit Treatment Value Assumption (SUTVA) (Hirano, Imbens and Ridder, 2003). The former requires that there is sufficient overlap in the characteristics of individuals in the treatment and the control groups to find adequate matches, it balances the distribution of observed characteristics across the treatment and the control group. And the SUTVA implies that the potential outcome for each individual is unrelated to the treatment status of other individuals, that is, no spillover. If there is a spillover of the treatment effect, that is, those individuals who do not accept the treatment are also affected by the treatment,

then the selection bias between the two groups cannot be well eliminated, and the final result is not unbiased. The advantage of PSM over linear regression models with control variables is that it does not require a correctly specified parametric model for the outcome.

There are two reasons why PSM cannot be used as the main (single) analysis approach in this thesis: 1). It is more suitable for single cross-sectional data (Koch and Alaba, 2010; Nguyen, 2012; Makhloufi, Ventelou and Abu-Zaineh, 2015); 2). Most importantly, it cannot handle unobservable confounders. As described in Chapters 2 and 3, in this study, unobservable factors, which affect the implementation and outcomes at the same time, such as economic development and medical resources, cannot be ignored.

Instrumental Variable (IV)

The Instrumental Variable (IV) method is commonly used in the literature. This method tries to find a variable, which is only correlated with the outcome through its effect on the treatment, called an "instrumental variable (IV)". By regressing the treatment on the IV, a new variable, the "estimated/expected treatment", can be obtained as a prediction. As the IV is unrelated to confounders, the predicted treatment, based on the IV, is also not correlated with confounders. It means the predicted treatment is approximately randomised. Thus, regressing the outcome variable on the predicted treatment removes the bias. IV can solve the "Omitted Variable" problem, whether from time-varying or time-invariant unobserved variables.

Despite its usefulness, the biggest challenge of IV is how to find a strictly exogenous IV, often termed a "good IV" (Wagstaff and Lindelow, 2008; Wooldridge, 2012). A variable is exogenous means that the variable is uncorrelated with the error term in the regression. Exogeneity can only be demonstrated against knowledge in the literature or expert opinion and cannot be directly tested. However, when there are multiple instruments for one endogenous variable, different IV estimates can be compared, which is called an overidentification test (Wooldridge, 2012).

The 2nd assumption is that IV is required to be related to endogenous (contrary to exogenous) treatment variables (Wooldridge, 2012). An IV is weakly correlated with an endogenous treatment variable, which is called a "weak instrument", leading to a precision loss in estimation. The 3rd assumption is that the IV estimation requires that the treatment effects are the same for everyone in the population or the treatment effects are different for everyone but the treatment assignment is not influenced by the unobserved individual effect (Wagstaff and Lindelow, 2008). When these assumptions are met, the ATE can be identified, otherwise, the Local ATE (LATE), which is the treatment effect of a particular subgroup affected by the instrument, is identified. Different instruments affect different individuals to enrol in treatment so give different LATE estimations. For example, in Wagstaff and Lindelow (2008) study of the impact of new rural cooperative medical care on financial risk, using government official status as an instrumental variable actually estimates ATE only for those subgroups whose treatment would be influenced by official status, not for the entire population.

Although finding a good IV is challenging, IV is common in policy evaluation studies with cross-sectional data. For example, Wirtz *et al.*, (2012) estimated the impact of health insurance on out-of-pocket costs using the proportion of the population belonging to Seguro Popular (one kind of insurance plan) in each state as IV, and the National Household Survey of Income and Expenditures cross-sectional data; Levine, Polimeni and Ramage (2016) estimated the impact of health insurance on medical costs, use of health services, and health outcomes, using cross-sectional survey data. Their IV is whether to use deep discount coupons to buy insurance. They also emphasised that the IV estimates the local average treatment effect. In addition, it is also seen as a supplement to the fixed-effect model when using panel data. For example, Wagstaff and Lindelow (2008) used three IVs in their study of the impact of new rural cooperative medical care on financial risk: whether the individual is a government official; whether he or she is head of the household; and whether the individual is a working member of the household. For some outcomes, IV gave similar results to FE, but FE was more reliable when insurance was exogenous.

However, there are two reasons why IV cannot be used as the main (single) analysis approach in this thesis: 1). Electronic Medical Record (EMR) data mainly contains patient disease information, so there is no variable that correlated with treatment but only correlated with the outcome through its effect on the treatment, which means it is impossible to find instrumental variables; 2). For China Health and Retirement Longitudinal Study (CHARLS) data, I tried this approach in the primary analysis and I will describe it in the empirical chapter later.

Synthetic Control Method (SCM) and Regression Discontinuity (RD)

There are also two other common methods of policy evaluation: the Synthetic Control Method (SCM) and Regression Discontinuity (RD).

The idea of SCM is to synthesise a virtual control group comparable to the treatment group by weighting multiple control groups, so it is more suitable to have a small number of treatment groups (or units) and a large number of control groups (or units). Moreover, it is more suitable the study with macro-level outcome variables, such as per capita Gross Domestic Product and average house price, rather than micro-level outcome variables. For example, Abadie, Diamond and Hainmueller (2010) investigated the impact of the tobacco control program on tobacco consumption. They used data from 38 states to synthesise a virtual control group to compare per capita tobacco use with California.

The idea of RD is that the allocation of treatment is determined by whether a certain continuous grouping variable exceeds a certain cut point/threshold. Since individuals on either side of the cut point are not considered to be systematically different (in either observable or unobservable characteristics), small differences in their values of the grouping variable are considered random. Therefore, the comparison of these two populations approximated a randomised trial, and thus local average treatment effects could be estimated (Imbens and Lemieux, 2008). This does not apply to insurance integration research, because the implementation of integration is determined by the local government in combination with political, economic and medical resource conditions, and there is no single indicator variable (no uniform and clear rules) that determines the implementation.

5.1.3 Difference-in-Difference Estimation

DID is the most widely used policy evaluation method in recent years worldwide (Huang, Zhang and Liu, 2022). There are many studies reviewing DID (Blundell and Dias, 2009; Imbens and Wooldridge, 2009; Lechner, 2010; Chen and Wu, 2015). DID first used in economics study was to evaluate the policy effect of post-schooling training on earning (Ashenfelter, 1978). In China, the earliest study using the DID method was to evaluate the effect of tax reform on farmers' income (Zhou and Chen, 2005). It was subsequently widely used in China's policy evaluation, such as the study on the impact of new rural cooperative medical care on household savings and consumption, the study on the impact of nutrition subsidies on adolescents' intelligence and physical strength, the study on the impact of housing purchase restriction policy on house price, and the study on the impact of state-owned enterprise privatization reform on enterprise revenues (Zhou, Wang and Li, 2011; Qi and Zhao, 2012; Zhang and Zheng, 2013; Zhang, Huang and Yan, 2016).

Classical Two-period DID

In DID method, the sample is divided into two groups, one is affected by the policy, as the treatment group; the other is not affected by the policy, as the control group. The differences between the two groups, and between before and after the implementation of the policy are calculated twice, eliminating the time-invariant characteristics difference between the two groups and the change of outcome value with time, in order to estimate the actual treatment effect (Lechner, 2011).

Based on the equation [1] in Section 5.1.1, a new dummy variable B is introduced, which distinguishes the states before and after the implementation of a policy. The value is 0 before the implementation and 1 after the implementation. That is:

$$B_t = \begin{cases} 1 \; if \; t \in the \; policy \; has \; been \; implemented \\ 0 \; if \; t \in the \; policy \; has \; not \; yet \; been \; implemented \end{cases}$$

The DID regression can be expressed by the following equation:

$$y_{it} = \alpha + \beta_1 A_i + \beta_2 B_t + \beta_3 A_i \times B_t + \varepsilon \qquad \textbf{[2]}$$

The mean of treated individual before the implementation is:

$\bar{y}_{treated\&before} = \hat{\alpha} + \hat{\beta}_1$

The mean of treated individual after the implementation is:

$\bar{y}_{treated\&after} = \hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$

The mean of untreated individual before the implementation is:

$\bar{y}_{untreated\&before} = \hat{\alpha}$

The mean of untreated individual after the implementation is:

$\bar{y}_{untreated\&after} = \hat{\alpha} + \hat{\beta}_2$

Calculating mean differences twice:

$$\left(\bar{y}_{treated\&after} - \bar{y}_{treated\&before}\right) - \left(\bar{y}_{untreated\&after} - \bar{y}_{untreated\&before}\right)$$
$$= \left[\left(\hat{\alpha} + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3\right) - \left(\hat{\alpha} + \hat{\beta}_1\right)\right] - \left[\left(\hat{\alpha} + \hat{\beta}_2\right) - (\hat{\alpha})\right]$$
$$= \left(\hat{\beta}_2 + \hat{\beta}_3\right) - \hat{\beta}_2 = \hat{\beta}_3$$

Thus, the $\hat{\beta}_3$ captures the actual average treatment effect on treat (ATT), which is called "Difference-in-Difference estimator".

To identify treatment effects, DID requires either panel data, that is, data from the same individuals at different time points, for example, Gotsadze *et al.*, (2015) used the DID approach to assess the impact of medical insurance for the poor on health service use among based on two years of identical household interview samples in Georgia; or repeated cross-sectional data, that is, samples from the same population at different time points, for example, Eduardo *et al.*, (2021) used the DID approach to assess the impact of the COVID-19 pandemic on low birth weight based on two-year electronic medical records in Argentina. Sometimes, DID is also used for single cross-sectional data by special study design. For example, Chen and Zhou, (2007) studied the effects of childhood famine exposure on adult height in China. They used the DID method to cross-section data in a single year, using age at the time of interview to distinguish whether or not experienced famine (equal to the time of birth was during the famine or after the famine ended), replacing the time variable representing before and after treatment in common DID studies.

For reliable results, DID also requires SUTVA to be satisfied. In addition, a most important assumption of DID is the Parallel Trend Assumption. It means that outcomes in the treatment

and control groups would have followed the same trend in the absence of treatment (Imbens and Wooldridge, 2009). If the outcomes of the two groups fail to satisfy this assumption, then the final estimate includes the difference in trend between the outcomes of the two groups, i.e. deviates from the true treatment effect. When using repeated cross-sectional data rather than panel data, in addition to the parallel trend, there is an implicit assumption that the characteristics of the sampled populations at different time points do not differ systematically (Lechner, 2011).

To avoid violating the assumption of the parallel trend, the use of DID models usually requires that the data originate from a "Natural /Quasi-Experiment", i.e. the treatment/policy is implemented randomly (Meyer, 1995; Lechner, 2010). However, in reality, policy implementation is usually planned/purposeful rather than random, which may (but not necessarily) lead to violating the parallel trend assumption. In this case, the selection of the control group (which population is more similar to the treatment group) is not consistent across studies because it is dependent on the particular dataset used and study objectives. For example, Wagstaff *et al.* (2009) discussed two different control group settings in their NRCMS study. One control group is individuals in a county but not enrolled in NRCMS, call "*the non-participant*", the other one is individuals in a county which did not implement NRCMS, called "*the non-exposed*". Since NRCMS is voluntary, adverse selection is a problem which needs to be considered. Some residents choose to join the NRCMS because they are sick or have just had an illness; some residents may not participate because they believe they do not need health services for the foreseeable future. Considering the adverse selection is a time-varying unobserved confounder, although it is unclear which estimation is actually closer to the truth, using the non-participant in their study is likely to have a greater bias than using the non-exposed. Similarly, in the research on the policy impact of the carbon emissions trading system on corporate R&D (Liu and Zhang, 2017), the treatment group is clear "pilot high carbon emission industries in pilot cities". The authors proposed two control groups: "non-pilot industries in pilot cities" and "pilot industries in non-pilot cities". However, using either one may violate the parallel trend assumption. If the former, the R&D investment of enterprises in the pilot industry relative to other industries may change over time (even if the carbon emissions

trading policy is not launched); if it is the latter, even in the absence of carbon emissions trading pilots, there may be different investment change trends due to differences in the actual conditions of the two types of regions, because the selection of pilot regions is not random. This dilemma is the same in the analysis of my thesis. In the end, unlike Wagstaff's decision, Liu and Zhang used both control groups, using the DDD approach (this will be discussed later) to minimise bias.

Multi-period DID

When the data are more than two periods, DID can be expanded into a two-way fixed effect model (TWFE). A series of time/year dummy variables $\boldsymbol{\mu_t}(=\sum \boldsymbol{year})$ replace one before-after dummy variable, which captures the year effect that does not change with the individual and is called "time fixed effect"; a series of individual dummy variables $\boldsymbol{\lambda_i}$ $(=\sum \boldsymbol{individual})$ replace one policy grouping dummy variable, which captures individual effects that do not change over time (such as unobserved individuals characteristics) and is called "individual fixed effect". Therefore, the TWFE model can be expressed by the following equation:

$$y_{it} = \alpha + \beta_3 A_i \times B_t + \lambda_i + \mu_t + \varepsilon \qquad \textbf{[3]}$$

Where the β_3 still represents the actual treatment effect.

However, when using repeated cross-sectional data, since the individuals in each period are not the same, only time fixed effects can be used and the grouping variable has to be retained, as shown in the following equation:

$$y_{it} = \alpha + \beta_1 A_i + \beta_3 A_i \times B_t + \mu_t + \varepsilon \qquad \textbf{[4]}$$

In reality, many policies are rolled out gradually rather than all at once. Thus, an advantage of the TWFE model is that it allows different individuals to receive treatment at different times, known as "staggered-DID" (Beck, Levine and Levkov, 2010; Chernozhukov *et al.*, 2013; Borusyak, Jaravel and Spiess, 2021). As shown in the following equation:

$$y_{it} = \alpha + \beta treated_{it} + \lambda_i + \mu_t + \varepsilon \qquad \textbf{[5]}$$

A single dummy variable $treated_{it}$ replaced the interaction term in equation [3] because different individuals receiving treatment at different times. That is:

$$treated_{it} = \begin{cases} 1 \; if \; t \in individual \; i \; received \; treatment \; at \; time \; t \\ 0 \; if \; t \in individual \; i \; has \; not \; received \; treatment \; at \; time \; t \end{cases}$$

When treatments occur at different time points, an additional implicit assumption is there are no effects of being treated in the future on current outcomes. This is called the "No Anticipation" assumption (Crepon *et al.*, 2010; Borusyak, Jaravel and Spiess, 2021). It requires that the individuals' outcomes have not changed before treatment. If this assumption is violated, beforehand changes in individuals' outcomes can significantly weaken the estimated treatment effect after the onset of treatment.

A second advantage of TWFE is that the dynamic effects of policies can be explored when data are available at multiple periods after treatment. The equation is as follows:

$$y_{it} = \alpha + \beta A_i \times B_t + \sum \gamma_t \, A_i \times B_t \times lag_t + \lambda_i + \mu_t + \varepsilon \qquad \textbf{[6]}$$

where the $\boldsymbol{lag_t}$ are a series of dummy variables indicating which year after the implementation, so $\boldsymbol{\gamma_t}$ indicates the effect difference between two years. However, Goodman-Bacon (2021) pointed out that the reliability of staggered DID estimation results relies on the assumption that the treatment effect does not change over time. Therefore, dynamic effect analysis is not applicable to staggered DID.

The last and most important advantage is about the parallel trend assumption. Since it is not possible to see trends in outcomes of individuals without treatment in the treatment group after the treatment was given, one possible approach is to examine trends in the treatment and control groups before the treatment was given. It requires that data are available at multiple periods before treatment. If there are no significant changes in the difference in outcome between the two groups before treatment, the parallel trend is likely to be satisfied.

The standard two-period DID model cannot be empirically tested for this assumption as there are only two periods of data. However, referring to La Ferrara, Chong and Duryea (2012), Wang and Wu (2019), and Chen *et al.*(2020), when multiple periods of pre-treatment data are used, this can be tested using the following equation (based on the equation [3]):

$$y_{it} = \alpha + \sum \beta_t A_i \times time_t + \lambda_i + \mu_t + \varepsilon \quad [7]$$

If none of the coefficients β_t before the policy is significant, it means that there is no significant change in the difference between the treatment group and the control group before the implementation of the policy, then the parallel trend assumption is highly likely satisfied. In addition, if the time of treatment varies, the test is slightly more complicated. In this instance, the comparison is not made at two "absolute" time points, but rather at two "relative" time points. The first step is to subtract the time of policy implementation from the time of data collection, and then put the resulting new variable "time gap" into the equation [7] to replace $time_t$, and then estimate the equation. Likewise, if coefficients before treatment are not significant, the assumption is likely to be satisfied.

Another test is called the "placebo test". This refers to another DID estimation using a "fake" treatment group, which is not affected by the policy (Gertler *et al.*, 2016). Referring to La Ferrara, Chong and Duryea (2012) and Li, Lu and Wang (2016), the purpose of the placebo test is to exclude the influence of other factors, such as policies. If the placebo results are still significant, then the raw results are unreliable.

Extension DID

Two extension DID approaches can solve the non-parallel trend problem: Propensity Score Matching with Difference-in-Difference (PSM-DID) and Difference-in-Difference-in-Difference (DDD).

If PSM and DID are used in tandem, PSM solves observed confounding, while the DID solves the time-invariant unobserved confounding. This combined approach is appropriate if non-parallel trends are attributed to observable factors. Conversely, if the non-parallel trend stem from unobservable factors, PSM-DID is not useful. The combination of PSM and DID is increasingly used, such as the study on the effects of Germany employment transitions on health, the study on the impact of patient education on medication adherence, and the study on the

influence of the mixed-ownership reform on the total factor productivity of state-owned enterprise (Gebel and Voßemer, 2014; Sari and Osman, 2015; Liu, Zhang and Lan, 2016).

DDD is an alternative method to solve the non-parallel trend problem. As mentioned earlier, there are two choices for the control group: the non-participants and the non-exposed. The DDD method uses both control groups and produces difference calculations thrice to eliminate the trend difference between the treatment and the control group. Taking the earliest DDD empirical research as an example (Gruber, 1994), the author analysed the impact of mandatory maternity insurance in the United States on the wages of 20-40 years old married women at the end of the 20th century. He used married women aged 20-40, single men aged 20-40 and people over 40 years old before and after the policy in both experimental and non-experimental states. First, he used the data of the 20-40 married women before and after the policy in the experimental state and non-experimental state to conduct a DID estimation, then used the data of the control group before and after the experimental state and non-experimental state control group population policy to conduct another DID estimation, and finally subtracted the two DID results to get the DDD estimation result.

Similarly process in the analysis of my thesis, in the context of insurance integration, all individuals can be divided into four groups belonging to policy area group (T)/non-policy area group (C) and resident group (R)/non-resident group(E). First, DID estimation is performed using resident groups in policy areas and non-policy areas. This DID estimator includes the true policy effect and also the bias caused by the non-parallel trend. The second step, DID estimation is performed using non-resident groups in policy areas and non-policy areas. Since the policy does not work for non-residents, this DID result is the bias caused by the non-parallel trend. Subtracting the two DID estimates yields the true policy effect. In short, the sole purpose of introducing the second DID is to remove the bias in the first DID.

The DDD method can also be estimated using a regression model. Based on equation [2], another group dummy variable C is introduced. It is used to distinguish whether an individual is a resident or a non-resident. The DDD regression can be expressed by the following equation:

$$y_{it} = \alpha + \beta_1 A_i + \beta_2 B_t + +\beta_3 C_i + \beta_4 A_i \cdot B_t + \beta_5 C_i \cdot B_t + \beta_6 A_i \cdot C_i + \beta_7 A_i \cdot B_t \cdot C_i + \varepsilon \quad [8]$$

There are 8 different expectations depending on the condition of the three dummy variables. After three different calculations, the β_7, the coefficient of the triple interaction term, captures the actual treatment effect.

The advantage of DDD is that all available data can be exploited, and most importantly, it does not need to satisfy two parallel trends separately in the two DIDs because it can eliminate the non-parallel trend bias, only needs to satisfy the Ratio Parallel Trend, which means that the non-parallel trend caused bias of both DID estimators is the same (Olden and Moen, 2022). In Gruber (1994)'s study, the Ratio Parallel Trend means the difference in average untreated outcomes between the treatment group (married women 20-40) and the never receiving treatment group (men 20-40 plus all people over 40 years old) changes in the same way in the treated and in the untreated states. In the context of insurance integration, it requires that the relative outcomes of the treatment and control groups in the policy area follow the same trend as the relative outcomes of the treatment and control groups in the non-policy area. This assumption can only be explored in its rationality in the application through literature and cannot be statistically tested.

The DDD method, proven to reduce bias in policy effect estimation, has witnessed a substantial increase in usage, from fewer than 100 studies in 2007 to 928 in 2017 (Lechner, 2011; Berck and Villas-Boas, 2016; Olden and Moen, 2022). Well-known studies include the study on the impact of California overtime penalty on work overtime, the study on the impact of Australian summertime on energy use, and the study on the impact of Added-value tax reform on labour share (Hamermesh and Trejo, 2000; Kellogg and Wolff, 2008; Xie, Mu and LI, 2017).

According to all the above discussions, Table 5-2 provides a summary of DID and its extension methods. Classical DID is only suitable for two-period data; when multi-period DID is used for panel data and repeated cross-sectional data, the model settings are slightly different, but the treatment time is required to be constant; multi-period staggered DID allows the treatment time to change but cannot estimate the dynamic treatment effect; PSM-DID and DDD are used for

situations where parallel trends are not met. Based on the characteristics of the data used, in the quantitative analysis of this thesis, the DID method will be the preferred method, and the PSM-DID and DDD will be supplemented methods.

In addition, as a globally popular causal identification method, DID has developed rapidly in recent years. After Goodman-Bacon (2021) discussed the heterogeneity issue of two-way fixed effects, an important development direction of DID is the heterogeneity of policy effects. For example, Callaway and Sant'Anna (2021) proposed aggregation methods, using the "csdid" command in Stata, which enables the examination of treatment effects across different dimensions and provides a summary of the overall treatment effect; Borusyak, Jaravel and Spiess (2021) proposed an efficient estimator when the heterogeneity of treatment effects is not restricted ("did_imputation" in Stata); Chaisemartin, D'Haultfoeuille and Guyonvarch (2023) proposed a DID design with multiple groups and periods, which can calculate dynamic policy effects ("did_multiplegt" in Stata). However, as these methods were developed during the closing period of my study, they were not used in this thesis but should be considered for use in future research.

Table 5-2 Summary of DID and its extension methods

Methods	Equation	Requirements/Applicability
Classic DID	[2]	2 period, panel or repeated cross-sectional data
Multi-period DID (Two-way Fixed Effect)	[3]	Multi-period, panel data, treatment time constant
Multi-period DID (Time Fixed Effect)	[4]	Multi-period, repeated cross-sectional data, treatment time constant
Multi-period staggered-DID	[5]	Multi-period, panel or repeated cross-sectional data, treatment time varying
Dynamic effect DID	[6]	Multi-period post treatment data, panel or repeated cross-sectional data, treatment time constant
PSM-DID	-	Non-parallel trends stem from observable confounding
DDD	[8]	Non-parallel trends stem from unobservable confounding, introducing a new control group.

Note: equation [1] is the linear regression used in RCT, and equation [7] is the regression used to test parallel trend assumption, not for DID estimation.

5.2 Analytical Models

In addition to causal identification strategies, analytical models also play a crucial role in policy evaluation. This section first discusses parametric regression models for different outcome variables, then considers non-parametric models.

5.2.1 Parametric Regression Models

The most commonly used regression model is the general linear regression model, which is based on the Ordinary Least Square (OLS) estimation technique. This method aims to find the best-fitting model by minimising the sum of the square residuals across all observations. To obtain a Best Linear Unbiased Estimator (BLUE)[22], certain key assumptions must be satisfied. These assumptions include a linear correlation between the independent variable and the dependent variable, constant variance of the error term (homoscedasticity), and a normally distributed error term (Wooldridge, 2012). In the general linear model, a statistical indicator to assess the goodness of fit of the model is R-squared. R-squared indicates the proportion of variation of the dependent variable that can be explained by the independent variables included in the regression model (Wooldridge, 2012).

However, in many economics and healthcare studies, including my study, the outcome variables often fail to meet some assumptions for the general linear regression model. Using the general linear regression model under such circumstances may lead to inconsistent parameter estimates or inaccurate predictions. Thus, more complex models are needed. Below is a brief discussion of different parametric regression models based on three types of variables: binary, count, and cost.

Binary Data

A binary variable refers to a variable that has only two possible values, 0 and 1. An example of a binary outcome is “the occurrence of hospitalisation”. Although binary variables can be

[22] The “Best” means the estimation has the lowest variance, and the “Unbiased” means the expectation/mean value of the estimation is equal to the true value.

analysed using the general linear model, known as the linear probability model (LPM), it presents limitations. One is that the LPM may produce negative predicted values, which are not realistic (Wooldridge, 2012). Logit and Probit regression models are developed to overcome the shortcomings of the LPM. They strictly restrict the predicted value to be between 0 and 1.

Regression results are typically interpreted based on predicted values and coefficients. Although the LPM is not recommended for predictive purpose analyses, it does offer more intuitive and interpretable coefficients. This is because the coefficient of the LPM directly mirrors the marginal effect. In contrast, the coefficients of probit and logit models do not reflect marginal effects. To derive an easily interpretable marginal effect value in these models, additional calculations are necessary.

In the probit and logit models, a single R-squared cannot be calculated with the same characteristics in the LPM. Instead, a pseudo-R-square is calculated. It is important to note that the pseudo-R-squared cannot be interpreted in the same way as an R-squared in OLS (Aldrich and Nelson, 1984; Ai and Norton, 2003; Hoetker, 2007; Wooldridge, 2012). There is no general test method to select between logit and probit models because they are very similar (Aldrich and Nelson, 1984). My study refers to use the Akaike information criterion (AIC) and the Bayesian Information Criterion (BIC) for model selection between logit and probit. Both AIC and BIC measure the goodness of fit by likelihood function while taking into account the complexity of the model. They penalise the model for having too many variables, so a lower value of AIC and BIC indicates a better-fitting model (Burnham and Anderson, 2004; Kuha, 2004; Aho, Derryberry and Peterson, 2014).

Count Data

Count data refer to outcome variables that have only non-negative integer values and are usually highly skewed, such as the commonly used indicator “hospital visit times”. The most common model for analysing count data is Poisson regression. It assumes the (conditional) variance equals the (conditional) mean, that is, no overdispersion. However, if the variance is much larger than the mean, the negative binomial model is an alternative model to analyse count data.

The likelihood ratio (LR) test can test the overdispersion so it can be used to select which one is better between the two models (Mullahy, 1997; Hardin, Schmiediche and Carroll, 2003; Wooldridge, 2012).

Cost Data

Due to the cost data that records zero values and the heavy right-hand tail, Generalised Linear Model (GLM) with gamma distribution and log link function is commonly recommended. Buntin and Zaslavsky (2004) compared eight different estimations of cost data, such as the two-part model, log-transformed OLS, and GLM, and suggested that GLM is the first choice. The GLM is a flexible extension of general linear regression. It allows the dependent variable to have different parametric distributions rather than only the normal distribution. It contains some common models, such as Poisson regression, logit regression, and so on.

The above parametric regression models combined with DID are widely used in policy evaluation. For example, in a study of the impact of the Italian innovation policy on business innovation, the Poisson DID model was used for the number of innovations, and the Logit DID model was used for whether a company has innovation (Bronzini and Piselli, 2016).

5.2.2 Nonparametric Regression Model

The models discussed so far primarily focus on changes in the mean value of outcomes. However, sometimes, policymakers and scholars are also interested in understanding the changes in the distribution of outcomes. For example, labour economists may be concerned about the impact of a policy on wages not only in terms of the average change for the entire population but also the changes in wages for individuals at different income levels. In this case, quantile regression is a useful method. It aims to estimate coefficients at the conditional median or other quantiles of the dependent variable rather than at the mean value.

Therefore, another extension of the DID method is Quantile Difference-in-Difference (QDID). Similar to quantile regression, QDID allows the estimation of treatment effects at different quantiles. It is a way to measure the heterogeneity of treatment effects across populations when

no variable is available for subgroup analysis. Many studies have discussed the estimation of this method (Athey and Imbens, 2006; Bonhomme and Sauder, 2011; Fan and Yu, 2012; Callaway and Li, 2017). Empirical studies using this method include the study on the effect of Norwegian childcare reform, and the impact of corporate donation tax credits on corporate donations (Havnes and Mogstad, 2015; Peng and Fan, 2016).

In this thesis, the application of the quantile model is restricted to city-level analysis and is not suitable for national-level analysis. The reason is that in the city-level analysis, there are uniform treatment rules and consistent health service/drug pricing for patients with the same disease in the same hospital. Consequently, different quantile costs can be indicative of varying disease severities. Then, treatment effects at different quantiles can thus be interpreted as differential treatment effects of policies on different health status subgroups.

Furthermore, differences in treatment effects at different quantiles can be linked to insurance policy design (deductibles, ceilings, and reimbursement rates). For example, an increase in the reimbursement rate may increase the total cost of patients with moderate severity because they are able to afford more amount or more expensive drugs or other treatments due to a reduction in out-of-pocket costs. However, this effect may not extend to mild or very severe patients. The cost of mild patients may not exceed the deductible, and the cost of very severe patients may surpass the reimbursement limit.

It is important to note that the above discussion is no longer reliable at the national level. Subgroups of patients in the same cost quantile cannot be assumed to have the same disease severity due to the mixing of patients from different regions, where the medical pricing and treatment protocols may differ. Consequently, the results of quantile analysis will be difficult to be interpreted well and realistically.

5.3 Conclusion

This chapter has reviewed various casual identification methods and regression models. The primary methodology problem in this thesis is to solve the potential selection bias or omitted

variable problems in observational data analysis. Therefore, special identification strategies are needed to identify causality rather than association. Through this chapter, three important things are clarified:

1. DID and its extension approaches are considered to be the preferred identification strategy for this study according to the datasets used.
2. Causal identification approaches should be combined with non-linear models, such as Logit regression, Poisson regression.
3. Quantile regression analysis is useful in city-level analysis to have an in-depth understanding of the mechanism of integration.

The next chapter will review previous empirical studies of insurance integration in China and internationally. These studies will contribute to the setting of detailed research objectives and the improvement of the empirical analysis framework.

Chapter 6 The Review of Empirical Studies

This chapter presents a full literature review of previous empirical studies on health insurance integration, mainly focusing on studies in China. The process of literature searching is described first, then Section 6.2 summarises and discusses the different results in thematic order: studies on people's attitudes, studies on inequality, non-policy evaluation studies on health service use and expenditure, and policy evaluation studies on health service use and expenditure. Given the limited number of studies on insurance integration in China, Section 6.3 also provides a concise review of findings from studies conducted in other countries. At the end of this chapter, research gaps are identified, and the study aims and detailed objectives are set.

6.1 Search Strategy

The initial plan was to conduct a full systematic literature search. However, practical considerations led to the abandonment of this plan. This is because the terminology is variable in sociology and economics articles, and the great variation in terms means that the search is not reproducible. For example, research related to health expenses may be expressed as "medical costs", "health expenses", "patient's financial risk", "patient's financial protection", "patient's economic burden", and so on. Consequently, an initial broad search was conducted using approximate keywords, followed by manual selection of studies related to the effect of China's social health insurance based on their titles and abstracts. Searches are performed using the Google Scholar search engine[23] (via Google LLC) to identify English literature only and the China National Knowledge Infrastructure (CNKI, via Tongfang Knowledge Network Technology Co., Ltd)[24] to identify Chinese literature only.

[23] The reason for using the Google Scholar search engine instead of other databases is that it can search the associated research based on rough keywords, although there may be potential research omissions.

[24] CNKI is a comprehensive China integrated knowledge resources system, including journals, doctoral dissertations, masters' theses, proceedings, newspapers, yearbooks, statistical yearbooks, eBooks, patents, standards and so on.

The objective of the literature search is to identify studies that evaluate the policy effect of health insurance integration on health service use and health expenses[25]. However, only 9 studies specifically focused on the policy effect. To broaden the scope, related studies on integrated health insurance, not limited to policy evaluation, were included in the review, such as those examining individuals' attitudes towards integration policy. These studies may indirectly reflect whether people benefit from reforms. Only articles published in peer-reviewed journals are included, grey literature, such as government studies, posters, conference papers, and students' dissertations and theses, are not included due to lack of peer review.

For the integrated health insurance study search, keywords include "(China OR Chinese) AND ((integrating OR integrated OR integration OR consolidating OR consolidated OR consolidation) AND (health insurance OR medical insurance) OR urban and rural resident basic medical insurance OR URRBMI)". The publication date is limited from January 2009 to June 2019. Quantitative empirical studies were included, while other study types were excluded. In addition, given the study's focus on the demand side (the insured/patients), macro-level quantitative research from the supply side (the insurer/government) was excluded.

6.2 Studies on Insurance Integration in China

The literature search involved iterative direct searches and consideration of studies in reference lists from topic-related literature. Therefore, the total number of papers searched is not available. As of July 2019, a total of 35 studies were identified, of which 29 are in Chinese and 6 in English. The distribution of research articles across years is as follows: 1 in 2011 (Chinese), 2 in 2013 (both in Chinese), 3 in 2014 (2 in Chinese, 1 in English), 2 in 2015 (both in Chinese), 3 in 2016 (all in Chinese), 7 in 2017 (all in Chinese), 14 in 2018 (11 in Chinese, 3 in English), and 3 in 2019 (1 in Chinese, 2 in English).

[25] Due to the limited number of studies on integrated insurance, another broader literature review of Chinese social health insurance has also been conducted. Compared to the integration review, the review of social health insurance is cruder. More details can be found in Appendix D.

These 35 studies are categorised based on research objectives into three groups: 1). Research on attitudes and willingness to participate in integrated insurance; 2). Research on the effect of integration, such as health status, medical expenses, and healthcare use; 3). Research on the health and healthcare equity of urban and rural residents. Further categorisation is applied to healthcare use and expenses research, distinguishing between policy and non-policy evaluations.

Although people's attitudes may not directly relate to health expenses or healthcare use, they can reflect people's self-perception of benefit changes. If people think their benefits will be damaged, they will oppose the integration or refuse to join the new integrated health insurance. Among the 35 studies, 9 are about the residents' attitudes and willingness to participate, 19 are about patients' financial burden and patient visits (9 of which are policy evaluations), and 7 are about urban-rural inequality.

6.2.1 Studies on People's Attitudes

Studies about patient attitudes can be further divided into three subgroups: attitudes toward integration before its implementation, willingness to participate and choices after integration, and satisfaction after integration. Because the research on attitudes is rich and has extensive content, the review concentrates on results directly relevant to my research topic, omitting less pertinent findings. For example, payment choices and benefits package choices are not included in the review. Table 6-1 provides a summary of attitude-related research.

The earliest empirical quantitative study (Qiu, Zhai and Hao, 2011) explored the attitudes and willingness of people in four cities regarding health insurance integration through a sampling questionnaire survey. Results revealed that 70.9% of interviewees were willing to integrate health insurances, with rural residents showing the strongest support. However, 17.4% of interviewees were unwilling. Main reasons for opposition included differing premium payment and benefits packages for urban and rural residents, insufficient management capacity, and distinct medical needs for urban and rural residents. Economic conditions and education levels were identified as mainly significant influencers of residents' attitudes through factor analysis

and logistic regression. Wealthier individuals were more supportive, while higher education levels correlated with greater opposition. A subsequent study (Wang *et al.*, 2014) conducted a similar questionnaire survey in four other cities, yielding similar results. 69.3% supported integration, with NRCMS interviewees exhibiting an 80.5% support rate. People from less developed regions were more likely to oppose the integration, which the authors suggest may be due to dissatisfaction with their existing health insurance scheme. Although the author did not explain this further, this explanation can be said to be reasonable. In general, the level of government management and welfare in underdeveloped areas is lower. Residents' dissatisfaction with the existing health insurance scheme leads to doubts about government capacity, which is also consistent with the insufficient management capacity in the second place of all the objections raised in the study.

Before and after integration, the individual's economic condition showed the same effect on the individual attitude, but the level of education showed the opposite effect. Before the integration, highly educated individuals were more opposed (Qiu, Zhai and Hao, 2011; Wang *et al.*, 2014). Two reasons can be summarised from the studies. On the one hand, they believed that integration would drag down their existing benefits, which was mentioned in both two studies. This is understandable. In the case of a constant supply of medical resources or just a small increase, the improvements in benefits for rural residents will inevitably reduce the actual benefits that urban residents (relatively higher educated) can enjoy. On the other hand, highly educated residents may be more aware of integration challenges, such as inefficient government management (distrust for the government). However, if the integration had already been implemented, highly educated residents would be more inclined to participate in integrated insurance (Sun, Zhao and Yin, 2017). While the study did not elaborate further, one possible reason is that people with higher education know more about the policy content of health insurance, more clearly the importance of health insurance, and have a higher awareness of risk prevention than those with low education. This is consistent with another result in Sun, Zhao and Yin (2017) study, indicating a positive association between knowledge of health insurance policy content and the willingness to participate in integrated insurance.

Table 6-1 List of literature on people attitudes related to insurance integration in China

Literature	Data Sources	Regions	Population	Methodology	Results
Qiu, Zhai and Hao (2011)	Sampling survey/ questionnaire	4 cities in 4 provinces: Taicang City in Jiangsu Province, Chengdu City in Sichuan Province, Dongguan City in Guangdong Province, and Xi'an City in Shaanxi Province	People in UEBMI, URBMI, and NRCMS	Descriptive statistics, factor analysis, logistic regression	Most support integration; the more developed the economy, the more agreed to the integration; the higher the level of education, the more inclined to oppose.
Wang *et al.* *(*2014)	Sampling survey/ questionnaire	4 cities in 4 provinces: Foshan City in Guangdong Province, Changshu City in Jiangsu Province, Shenyang City in Liaoning Province, and Changchun City in Jilin Province	People in UEBMI, URBMI, and NRCMS	Descriptive statistics, statistical tests, logistic regression	Most support integration; the more developed the economy, the more agreed to the integration; the higher the level of education, the more inclined to oppose.
Sun *et al.* (2014)	Sampling survey/ questionnaire	2 cities in 1 province: Shuizuishan City and Guyuan City in Ningxia Autonomous	People in URRBMI; only outpatient and hospitalisation	Descriptive statistics, factor analysis, logistic regression	Adverse selection in multiple benefit package setting; satisfaction is not high; reimbursement scope and ceiling are the main factors affecting overall satisfaction.
Chen *et al.* (2015)	Sampling survey/ questionnaire	1 county in 1 province: Gaoqing County in Shandong Province	People in URRBMI	Descriptive statistics, statistical tests, logistic regression	Adverse selection in multiple benefit package setting: people with poor health tended to choose a benefit package with a higher reimbursement rate.

Literature	Data Sources	Regions	Population	Methodology	Results
Shang et al. (2016)	Sampling survey/ questionnaire in 2015	2 cities in 1 province: Weihai City and Zibo City in Shandong Province	People in URRBMI	Descriptive statistics, statistical tests, logistic regression	Adverse selection in multiple benefit package setting: people with better health tended to choose a low-level benefit package.
Sun, Zhao and Yin (2017)	Sampling survey/ questionnaire in 2015	5 cities in 1 province: Weihai City, Dongying City, Zibo City, Linyi City, and Jining City in Shandong Province	People in URRBMI	Descriptive statistics, statistical tests, logistic regression	Income, education, and knowledge of insurance policies positively related to the willingness to participate. An important reason for dissatisfaction was the cancellation of outpatient reimbursement after integration.
Yan et al. (2018)	Sampling survey/ questionnaire, survey time unknown	Anonymous	People in URRBMI	Descriptive statistics, statistical tests	Urban residents more satisfied than rural residents; highest satisfaction was with the insurance institutions and lowest satisfaction with health insurance policies.
Shan et al. (2018)	Sampling survey/ questionnaire, 2014 to 2015	Chongqing City, Ningbo City and Heilongjiang Province (city unspecified)	Population unspecified	Descriptive statistics, statistical tests, logistic regression	47.6% of the interviewees were dissatisfied with the integration; many interviewees argued that integration reforms failed to significantly reduce the large gap between urban and rural residents.

Literature	Data Sources	Regions	Population	Methodology	Results
Gao (2018)	Sampling survey/ questionnaire, survey time unknown	1 city in 1 province: Nanjing City in Jiangsu Province	Population unspecified	Descriptive statistics, statistical tests, logistic regression	Overall satisfaction is not high; only 17.78% of people expressed satisfaction with the integration management; the increase in premiums caused some residents to abandon their insurance; the main purpose of patients going to the clinic is to prescribe drugs rather than to see a doctor, the real visit ("consulting") is still concentrated in high-level hospitals.

As for satisfaction with the integration, research for different regions and populations showed different results. Among the five studies, three showed overall dissatisfaction, while two showed overall satisfaction.

Sun *et al.* (2014) measured satisfaction in Ningxia using a scale of 1 (very satisfied) to 5 (very dissatisfied). The findings revealed relatively low overall satisfaction, with 3.19 for outpatients and 2.73 for hospitalisations. Gao (2018) conducted a study in Jiangsu and found that only 17.78% of interviewees expressed satisfaction with the integration work of medical insurance management. In addition, the increase in premiums caused some residents to abandon their insurance. Shan *et al.* (2018) found a dissatisfaction rate of 47.6% in interviewees from Chongqing City, Ningbo City, and Heilongjiang Province (city unspecified). Compared with the two studies before integration (Qiu, Zhai and Hao, 2011; Wang *et al.*, 2014), people's attitudes were consistent. Dissatisfaction with the management system and the actual benefits occupied first and third place. Notably, equity (equal access to healthcare) was the most crucial reason for interviewees to support integration, and failure to reduce inequality ranked second among interviewees' dissatisfaction.

In contrast, two studies reported overall good satisfaction levels but also highlighted specific areas of dissatisfaction. Sun, Zhao and Yin (2017) found that 73.7% of the residents were satisfied in five cities in Shandong Province. However, dissatisfaction arose from the cancellation of outpatient reimbursement after integration in Weihai City[26]. Yan *et al.* (2018) explored the satisfaction of residents with health insurance policies, insurance institutions, and designated medical institutions after integration. Overall, urban residents were more satisfied than rural residents (3.78 vs 3.24, 1-very dissatisfied, 5-very satisfied). However, satisfaction with the insurance policy was the lowest among the three aspects of satisfaction studied. In terms of satisfaction with various aspects of the insurance policy itself, satisfaction with the

[26] Similarly, a government study (Changxue *et al.*, 2016) also mentioned that some health services that can be reimbursed in NRCMS could not be reimbursed after integration, which has led to rural residents' opposition to integration.

scope of reimbursement was the lowest. The interviewees indicated that this was because some commonly used drugs and some valuable drugs could not be reimbursed.

As can be seen from the above research, integration is what people expect, but people have also realised the potential damage caused by integration. The low satisfaction after integration reflects that the actual effect of the integration may be worse than expected. The main problem is that some services and drugs are no longer being reimbursed. Therefore, although the adjustment of the drug reimbursement catalogue is not the focus of this study, it is worth discussing in conjunction with subsequent empirical analysis of the effect of integration, because the shrinking of the reimbursement catalogue may lead to a less significant effect of integration.

6.2.2 Studies on Inequality

Six out of seven studies on healthcare inequality used cross-sectional data, with one using panel data to assess differences between integrated and unintegrated regions, or inequalities within integrated regions. Four studies used the Concentration Index (CI) method to measure inequality, while three studies used policy evaluation methods. The summary of these seven studies is presented in Table 6-2.

Yang et al. (2018), using the CI, identified pro-rich inequities in hospitalisation reimbursement across NRCMI, URBMI, and URRBMI. Pro-rich inequality means that richer insureds benefit more from insurance than poorer insureds. The situation was the most severe in NRCMS (highest CI: 0.2473) and the mildest in URRBMI (lowest CI: 0.1032). Li and Tang (2018) found fair outpatient costs, pro-poor inequalities in outpatient visits and hospitalisation rates, along pro-rich inequality in hospitalisation costs. Wang et al. (2019) found inequalities in both NRCMS and URRBMI for rural residents, with NRCMS exhibiting a more serious situation. Among them, outpatient visits had pro-rich inequality, but hospitalisation visits had pro-poor inequality. Chang et al. (2019) found pro-rich inequality in hospitalisation. Despite claims that integration mitigated inequality, these studies were unable to definitively reflect the causal impact of integration due to limitations in their study design and regional/population differences.

Table 6-2 List of inequality-related literature on insurance integration in China

Literature	Data Sources	Regions	Population	Methodology	Results
Ma, Song and Gu (2016)	China Health and Retirement Longitudinal Study, 2008 and 2012	2 provinces: Gansu Province and Zhejiang Province	People in URBMI, NRCMS and URRBMI	DID, fixed-effect model	Integration mitigated 55.2% inequality in outpatient visit.
Ma, Gu and Sun (2017)	Sampling survey/ questionnaire, survey time unknown	3 areas in 1 province: Yixin, Taicang, and Xixia in Jiangsu Province	Urban and rural residents	Oaxaca-Blinder decomposition, fairness gap, PSM	Compared with unintegrated areas, the integrated area significantly alleviated the inequality of access to health care services and health status.
Ma, Gu and Song (2018)	China Migrants Dynamic Survey in 2014	31 provinces	Migrants in URBMI, NRCMS and URRBMI	Heckman model, PSM-DID	Integration has significant impact on mitigating inequality.
Yang *et al.* (2018)	the Fifth National Health Services Survey in 2013	8 areas in Anhui province	People in URBMI, NRCMS and URRBMI	Linear regression, concentration index	The CI of NRCMS was greatest, while that of URRBMI was lowest. Integration mitigated pro-rich inequity.
Li and Tang (2018)	Sampling survey/ questionnaire in 2016	1 city in 1 province: Shizuishan City in Ningxia Autonomous	People in URRBMI	Statistical test, concentration index	Inequality between different income groups was slight.

Literature	Data Sources	Regions	Population	Methodology	Results
Chang et al. (2019)	the Fifth National Health Services Survey in 2013	all pilot areas	People in URRBMI	Concentration index	A pro-rich inequity in hospitalisation utilization.
Wang *et al.* (2019)	China Health and Retirement Longitudinal Study, 2015	Pilot areas, not described in detail	People in NRCMS and URRBMI; rural resident only	Descriptive statistics, statistical tests, logistic regression, concentration index	A pro-rich inequity in outpatient utilization but a pro-poor inequity in hospitalisation utilization; for both the outpatient and hospitalisation care, the inequity degree of URRBMI was larger than that of NCMS.

In particular, Li and Tang (2018) and Chang et al. (2019), only estimated the CI in the URRBMI population. For Yang et al. (2018) and Wang et al. (2019), although they compared the CI of integrated and unintegrated regions, a simple two-group comparison was difficult to prove whether the differences were due to integration or other factors. In addition, as in those medical costs and health services research, regional and population differences may lead to opposite results. For hospitalisation inequality, Yang et al. (2018) and Wang et al. (2019) thought that it was pro-poor, but Chang et al. (2019) believe that it was pro-rich.

Ma Chao's research team used data from different sources to examine inequalities in different regions and populations (one is the China Health and Retirement Longitudinal Study, one is a self-designed questionnaire survey, and one is the China Migrants Dynamic Survey). All three studies demonstrated that integration mitigated healthcare inequalities (Ma, Song and Gu, 2016; Ma, Gu and Sun, 2017; Ma, Gu and Song, 2018).

The above research indicates inconsistent healthcare inequalities across different samples and service types. While most studies suggest integration alleviates inequality, certain study designs challenge their conclusions. If indeed inequality is alleviated, it indirectly supports the notion that integration policies have improved healthcare use and reduced patient burdens. Consequently, future empirical analyses should explore variations in integration effects among different subgroups and across different cities.

6.2.3 Non-Policy Evaluation Studies on Healthcare Use and Expenses

Table 6-3 provides a summary of impact-related research without the use of policy evaluation methods among the 10 identified studies. Four of these studies compared changes in integrated areas before and after implementation, while six only examined the post-integration period.

Among the post-integration studies, three focused solely on integrated areas after integration, one compared integrated and non-integrated areas after the integration, and two explored the performance of different insurance schemes post-integration, covering resident insurance,

employee insurance, and non-insured. Liu's research team, in their two studies, described outpatient and hospitalisation visits and costs in a county in Shaanxi Province after the integration (Liu, Wang, *et al.*, 2013; Liu, Xiong, *et al.*, 2013). The number of outpatient visits showed an upward trend, while the costs showed a downward trend; the average outpatient costs were lower than the national rural average, and the hospitalisation costs and length of stay in the hospital were also lower than the national average. Hu *et al.*, (2018) and Ma and Xu (2018) compared different insurance schemes' performance across provinces. Patients in URRBMI had a lower hospital visit rate than those in UEBMI and URBMI; the length of stay in the hospital was shorter than all other patients; the total outpatient costs were lower than those in UEBMI but higher than those in URBMI and NRCMS; the out-of-pocket outpatient costs were higher than all other patients; the total costs and out-of-pocket costs of hospitalisation are lower than all other patients.

Unlike other studies, Wang and Liu (2015) employed macro data from statistical yearbooks to compare differences between integrated and non-integrated regions after integration and found that patients in the integrated areas were diverted, with fewer visits to higher-level healthcare institutions, such as city hospitals, while with more visits to lower-level institutions, such as village clinics. However, another newer study (Gao, 2018) pointed out that the main purpose of patients going to the clinic was to obtain drug prescriptions rather than to see a doctor, and that the real visits ("consultation") were still concentrated in high-level hospitals[27]. Although Wang and Liu (2015) and Gao (2018) focused on different regions and populations, Gao's explanation may also apply to Wang and Liu's study. The process of going from receiving drugs once, but a larger quantity at a high-level healthcare institution to receiving drugs many times, but at a smaller quantity each time at a low-level healthcare institution cannot be called an increase in health service use. Therefore, drug prescribing and counselling visits could be the focus of future research.

[27] Although Gao's study involved the changes in health service use, the main objective of this study was to measure patients' satisfaction, so it was classified as an attitude study rather than health service research.

Table 6-3 List of impact-related non-policy evaluation literature on insurance integration in China

Literature	Data Sources	Regions	Population	Methodology	Results
Liu et al. (2013)	Hospital EMR database, 2010 and 2011	An unspecified county in Shaanxi Province	11 hospitals, people in URRBMI	Descriptive statistics	The median cost of outpatient visits in 2010 and 2011 was 47.54 CNY and 46 CNY, respectively. More than 50% of outpatient expenses were less than 50 yuan, and about 90% of outpatient expenses were less than 100 yuan. The average outpatient cost was on a downward trend.
Liu, Wang, et al. (2013)	Hospital EMR database, 2011	An unspecified county in Shaanxi Province	Rural resident in URRBMI; hospitalisation	Descriptive statistics	The average hospitalisation cost was 2,067 CNY (the median is 1,188 CNY); the average length of stay in hospital was 10 days; the proportion of medicines is high, exceeding 30%.
Xu and Zhang (2014)	Insurance database, 2010-2013	Tianjin Municipality	People in URRBMI	Descriptive statistics, unspecified regression model	The actual average outpatient reimbursement rate of adults and students/children were far lower than the theoretical reimbursement rate; the actual reimbursement rate of the three different levels of benefit packages was positively correlated with the level of premiums.

Literature	Data Sources	Regions	Population	Methodology	Results
Wang and Liu (2015)	Province Health Statistics Yearbook, 2012 and 2013	58 areas in 3 provinces: Guangdong Province, Anhui Province, and Sichuan Province	All residents	Oaxaca-Blinder decomposition, linear regression	Integration has facilitated the diversion of patients' visits: the number of hospitalisation, outpatient, and emergency visit decreased in urban hospitals, while the number of visit in township clinics increased.
Xie, Wang and Zhu (2017)	Hospital EMR database, 2013 to 2016	1 city in 1 province: Guangzhou City in Guangdong Province	Hospitalisation children/students in one hospital	Descriptive statistics, statistical tests, log-linear regression	After the integration, the number of hospitalisations increased but the length of stay in hospital decreased; in log-linear model, integration significantly increased the total cost and out-of-pocket expenses.
Xie, Hu and Chen (2017)	Sampling survey/questionnaire in 2011 and 2012	5 counties in 1 province: Haiyuan, Yanchi, Tongxin, Pengyang and Xiji in Ningxia Autonomous	People in NRCMS and URRBMI	Descriptive statistics, statistical tests	Interviewees' outpatient and hospitalisation visits increased, but self-reported health declined. There were no significant changes in the choice of medical institutions at different levels. Clinic visits increased and hospital visits decreased. The patient's satisfaction with the treatment increased significantly.

Literature	Data Sources	Regions	Population	Methodology	Results
Jiang, Luo and Zeng (2018)	Local Medical Insurance Authority and Government Open Information, 2015 to 2017	An anonymous city in Jiangxi Province	All insured	Descriptive statistics, analysis of variance	The average hospitalisation cost decreased; the average special outpatient cost increased due to the expansion of the special chronic disease directory for outpatients.
Dai et al. (2018)	Hospital EMR database, 2015 to 2018	1 city in 1 province: Lu'an City in Anhui Province	All insured; hospitalisation	Descriptive statistics	After the integration, the number of hospitalisations declined, but the per capita cost increased.
Hu et al. (2018)	Sampling survey/questionnaire in 2006	3 cities in 1 province: Guyuan City, Yinchuan City, and Shizuishan City in Ningxia Autonomous	People in UEBMI, URRBMI, and uninsured	Descriptive statistics, statistical tests, logistic regression	For residents with UEBMI, URRBMI, the outpatient and hospitalisation visits are higher than those without any insurance.
Ma and Xu (2018)	Pilot Surveillance of Health Service Utilization Behaviour of Residents in China in 2006	8 areas in 4 provinces: Gusu District and Jinhu County in Jiangsu Province, Qingshan District and Macheng City in Hubei Province, Liwan District and Yingde City in Guangdong Province, Qingyang District and Yanting County in Sichuan Province	Hypertensive Patients in UEBMI, URBMI, NRCMS and URRBMI	Descriptive statistics, statistical tests	Patients with integrated insurance's visits rate were lower than those in UEBMI and in URBMI; the LOS in hospital as lower than all other patients; the total outpatient cost as lower than those in UEBMI but higher than those in URBMI and NRCMS; the out-of-pocket outpatient costs were higher than all the other patients; the total costs and out-of-pocket costs of hospitalisation were lower than all other patients.

Xu and Zhang's study (2014) was very noteworthy, exploring the disparity between actual and theoretical reimbursement rates (RR) in URRBMI. They found that the actual average RR of adults and students/children were far lower than the theoretical RR in both hospitalisation and outpatient. The theoretical RR of hospitalisation was 56% on average, while the actual adults and students/children RR were 48% and 42% on average; the theoretical RR of outpatient was 30% to 40%, but the actual adults and students/children RR were only 12.03% and 7.83%. In addition, the actual RR of the three different levels of benefit packages was positively correlated with the level of premiums. Their study may hint at a problem for integrated insurance, that is, even if the integration improves the reimbursement rate, actual patients' reimbursement might not see corresponding improvements. In practical terms, this gap arises from constraints imposed by reimbursement catalogues, which dictate eligible services and medicines, and the design of insurance, including deductibles and ceilings. Consequently, only a fraction of the patient's incurred expenses becomes reimbursable. In addition, due to the hospital budget constraints[28], the doctor may be inclined to prescript those un-reimbursable drugs to patients, resulting in a scenario where the actual reimbursement for the patient's treatment falls short of the theoretical reimbursement.

Of the four studies comparing changes in integrated areas before and after integration, conflicting results were observed. It may be due to differences in regions and populations. For hospitalisations, Xie, Wang and Zhu (2017) and Xie, Hu and Chen (2017) found an increase in hospitalisations for children/students and rural residents, while Dai *et al.* (2018) found a decrease in hospitalisations for all insured. This inconsistency in overall population and subgroup results may be due to greater decrease in hospitalisation among adults and urban residents than the increase in hospitalisation among children/students and rural residents; it may also be due to differences in regions, making the results incomparable. Only one study (Xie, Hu and Chen, 2017) involved outpatient visits and found that they increased. In addition, Xie,

[28] The payment system for health insurance in China is global budgeting, which means the amount of reimbursement for a hospital is fixed in a year. If the reimbursement amount of a department exceeds the budget, the excess will be deducted from the salary of all doctors in the department (punish to the doctor).

Hu and Chen (2017) also found the situation of patient diversion (more clinic visits and fewer hospital visits). For costs, Xie, Wang and Zhu (2017) and Dai *et al.* (2018) found that total costs and out-of-pocket costs increased for hospitalisation visits. However, Jiang, Luo and Zeng (2018) found outpatient costs increased, but hospitalisation costs decreased. Dai *et al.* (2018) and Jiang, Luo and Zeng (2018) had the same population and methods, so the inconsistency of results may come from data source differences, regional differences, and policy differences.

Simple post-policy analyses, whether examining integrated insurance independently or comparing different insurance schemes, did not focus on policy effects and therefore cannot be viewed as policy evaluations. Similarly, comparison studies, whether comparing periods before and after policy implementation or integrated and non-integrated areas may be susceptible to biases stemming from temporal changes and regional variations. Therefore, greater attention should be directed towards studies employing rigorous policy evaluation methods to ensure a more accurate understanding of the policy effects.

6.2.4 Policy Evaluation Studies on Healthcare Use and Expenses

Of the nine studies using policy evaluation methods, eight used DID or PSM-DID design, while three used the fixed-effect model (or as a robustness test for DID), and one used the regression discontinuity design. Since these studies are closely linked to my research, the study design, such as causal identification strategy and variable selection, will be discussed in detail. Combining the theory and methodology review in previous chapters, as well as empirical research review in this section, the empirical analysis framework of this thesis will be constructed.

Data sources

The data sources for these nine studies vary. Six of them used data from the China Health and Retirement Longitudinal Study (CHARLS) spanning different years for national-level analyses: 2008 vs 2012 (Ma, Zhao and Gu, 2016), 2011 vs 2013 ((Li, 2017; Liu, 2017; Liu and Lin, 2018), and 2011 vs 2015 (Ma and Li, 2018; Su *et al.*, 2019). Data from CHARLS is popularly used in the health field, for example, in studies on health behaviours, health insurance, and health

inequality. While CHARLS provides nationally representative data for individuals aged 45 and above, its lack of representation for younger individuals may lead to an overestimation of policy effects on healthcare use and expenses. On average, older people are obviously in worse health and therefore have more medical needs and higher expenses. In addition, although the survey team has taken steps, recall bias is inevitable in any survey, such as income and the frequency of hospital visits. The CHARLS database will be also used for the national-level empirical analysis in this thesis and will be discussed in more detail in Chapter 8.

For three other studies, Huang and Zhang (2017) used 2014 and 2015 insurance claims data from Nansha District, Guangzhou City, Guangdong Province to estimate the impact of integration on medical visits and costs, so it is a city-level study; Chang, Su and Zhou (2018) used the 2012 and 2014 China Labour-force Dynamic Survey (CLDS) to estimate the impact of pilot integration on health status and also medical visits; Liu *et al.* (2018) used Electronic Medical Records (EMR) data from hospitals in Feng County, Shaanxi Province, from 2009 to 2011, so it is also a city-level study.

The CLDS database targets the working population aged 15-64, so it has a broader representation than CHARLS. However, since the CLDS database is not freely available to the public, it cannot be used in this thesis. The data from insurance claims and medical records are much more accurate than the survey data. However, they also have two disadvantages. First, their population is limited. The former is only for people with insurance and does not include self-payers (but this may not be an issue in insurance studies), and the latter is only for patients who visit hospitals (so it is not possible to analyse changes in visit rates). Another is that it does not include patient socioeconomic information, so available control variables are limited. This has the potential to lead to a larger deviation of the estimated value from the true value.

Due to the various populations covered by different databases, the policy evaluation results may be inconsistent. In addition, because the integration policy is implemented gradually, using national data from different years, the actual integration areas analysed are also different. These two factors may be the main reason for the inconsistent evaluation results of policy effects.

Causal identification strategy

From the research design of causal identification, eight studies adopted the DID method, and/or fixed effect model, and/or PSM-DID method because their data sources are all panel data or repeated cross-sectional data.

Ma, Zhao and Gu (2016) focused on rural residents, and used the standard two-period DID method with the non-exposed comparator, that is, rural residents in integrated areas versus rural residents in unintegrated areas. The authors argued that the parallel trend assumption might be violated due to the non-random nature of policy implementation. Since there are only two data periods, it is not possible to directly test for parallel trends. Consequently, PSM-DID and individual fixed effect models were used to test the robustness of the results.

Li (2017) also mentioned the non-randomness of policy implementation, so they directly used PSM-DID (standard two-period DID method with the non-exposed comparator) as the main analysis rather than DID and did not conduct any robustness analysis. This may lead readers to question the reliability of the results.

Under the terms of Guangzhou's integration policy, Huang and Zhang (2017) thought that urban residents would be less affected while rural residents would be more affected. Therefore, urban residents were used as the control group and rural residents as the treatment group, and then a standard two-period DID was used for estimation without robustness tests.

Liu (2017) used the PSM with the non-exposed comparator as the primary analysis, while the individual fixed-effects model and PSM-DID were used as robustness tests. The authors highlighted the unobservable omitted variable problem, that is, the non-random problem of policy integration highlighted in other studies. Therefore, an advantage of this study that is particularly worth pointing out is the use of Monte Carlo simulations to estimate omitted variable bias due to unobservable variables. The results of the Monte Carlo simulations were consistent in the direction of impact with the PSM results but significantly smaller, that is, PSM estimates were significantly affected by omitted variables.

Chang, Su and Zhou (2018) used standard two-period DID and the included population were residents with URBMI, NRCMS, or URRBMI. A flaw of this study is that it did not take into account the non-random nature of policy implementation as other studies have done, so no robustness testing or sensitivity analysis was performed.

Liu and Lin (2018) included residents with URBMI, NRCMS, or URRBMI. The author also mentioned the non-randomised policy implementation, so the primary analysis approach was PSM-DID (with non-exposed comparator). Different from other studies, in addition to the subgroup analysis, the author also introduced a quantile regression model, so that the heterogeneity of different populations affected by the policy could be discussed in more detail.

Ma and Li (2018) did not differ methodologically from previous studies, being PSM and standard 2-period DID.

Su et al. (2019) had the most rigorous design in all review integration studies. It rigorously selected people in NRCMS, URBMI, and URRBMI to explore the differences in hospital visits and expenses between the population in URRBMI and the population in NRCMS/URBMI in four pilot provincial regions: Zhejiang Province, Shangdong Province, Guangdong Province, and Chongqing Municipality. The PSM with standard 2-period DID method was used. Uniquely, the authors further included province and community fixed effects to reduce bias from unobservable confounding. Subsequent robustness tests were also performed using different matching methods and fixed-effects models.

In summary, many studies have emphasised the non-random implementation of policies and then selected the PSM as a supplement to the DID regression with control variables. As Stuart (2010) said, matching and regression with control variables have been shown to work best in combination, the idea called “double robustness” where regression is used to “clean up” small residual control variable imbalances between groups. However, insurance integration is a top-down reform, and its implementation is determined by local governments based on macro factors such as local economic, financial and medical resource conditions. When using

individual-level data (such as CHARLS data), these macro factors that determine whether to implement integration are unobservable confounders. As discussed in Chapter 5, PSM can only mitigate bias due to observable confounding but not unobservable confounding. Thus, if the single DID estimate violates the parallel trends assumption and is due to unobservable confounding, then PSM-DID may not provide more reliable results than single DID. This is a common deficiency in most studies. Only two studies (Liu, 2017; Su *et al.*, 2019) mentioned this issue and attempted to mitigate bias through Monte Carlo simulations and province/community fixed effects.

In addition, all DID models use the standard 2-period format and do not take into account the multi-stage/gradual implementation of policies in different regions. In previous studies, this was understandable due to the availability of data and the pilot nature of the policy. However, for newer studies, where more data are available, and more regions implement integrations, a staggered-DID model rather than a standard DID model is more appropriate.

Another study, Liu *et al.* (2018) claims to use the Regression Discontinuity (RD) method, but since there is no methodological description of RD and any RD results (neither regression tables nor images; only descriptive statistics results are provided), it is impossible to assess how good their application of the method is.

Control variables

In the studies using CHARLS data, similar control variables were selected, including demographic factors like age, sex, education, marriage, and various forms of income (1 used per capita annual household income, 4 used total annual household income, 1 used personal annual income, 1 used city average per capita annual income). In addition, involving economic factors, Ma, Zhao and Gu (2016) used the amount of medical expense reimbursement, Liu and Lin (2018) used the child's financial support and the amount of fixed assets, and Su et al. (2019) used the economic dealings with the child and pension.

Health-related variables were crucial, with four studies using self-reported health status, four using the number of chronic diseases, one using specific disease variables (e.g. disability, cancer, hypertension), and one considering cognition and depression. Daily life and behaviours were also included in all studies, covering aspects like smoking, drinking, daily activities, life satisfaction, and parent-child relationship. Uniquely, Ma, Zhao and Gu (2016) included the availability of running water and toilets. Moreover, five studies considered household registration and geographical differences, using variables like household registration type (agricultural/non-agricultural), geographic location (Eastern/Central/West), urban-rural, and community.

In contrast, the three studies not using CHARLS data employed a more limited set of control variables. Huang and Zhang (2017) used only sex, age, with chronic diseases or not, and household size; Chang, Su and Zhou (2018) used age, sex, education, income, smoking, drinking, drinking water source, sanitation, and distance from the nearest medical facility; Liu *et al.* (2018) used time, age, sex, and hospital level. As mentioned earlier, insurance claim and medical records data may have biased analysis results due to too few control variables, particularly lacking patient economic information.

In summary, the selection of control variables across these studies aligns with frameworks like the Grossman model and the Anderson model. Variables such as age, sex, income, education, marital status, and personal behaviour, are emphasised in the Grossman model, and are also referred to as predisposing component and enabling component in the Anderson model. Therefore, these frameworks will also guide the variable selection process in the subsequent empirical chapters (8 and 9) of this thesis.

Results

Finally, turn attention to their research results. Despite the use of CHARLS data in six studies, the results exhibit inconsistencies. Ma, Zhao and Gu (2016) observed a significant increase in outpatient visits (at the 5% level) and costs (only at the 10% level) for rural residents, with no effect on hospitalisation. Liu (2017) results showed that the integration did not affect health

service use, whether for all residents or for urban or rural subgroups. Ma and Li (2018) results showed that integrated insurance significantly reduced the probability of hospitalisation for residents but had no effect on hospitalisation costs and outpatient visits; the hospitalisation of agricultural hukou residents significantly reduced, but there was no impact on patients with non-agricultural hukou. Integrated insurance had different effects on people with different health statuses (good, fair, bad). Latest, Su et al. (2019) results showed that the integration had no impact on the whole population, but decreased rural residents' "need-but-not" hospitalisation visits (refers to the doctor asking the patient to be hospitalised, but the patient is not hospitalised), that is, increased hospitalisation. It also increased their hospitalisation expenses. These results are contrary to the results of Ma and Li (2018) study.

Moreover, Li (2017) results showed that the family medical burden (the proportion of annual household medical expenditure to total annual household income) was significantly reduced. However, Liu and Lin (2018) used same the data as Li (2017) used. The main results showed that the overall medical burden increased, while the burden of rural residents decreased, and urban residents were not significantly affected. Further, quantile regression results showed that different quantile groups, that is, groups with different medical burdens, are affected differently by policies, whether in rural or urban areas.

The inconsistency in these results may stem from a lack of uniformity in the analysis framework, despite using the same data source. Factors contributing to this inconsistency may include:

1). Since the policy was gradually implemented, variations in the years of data use represented different populations covered by different policies.

2). The choice of outcome variables was different, for example, "hospitalisation" and "need but not hospitalisation" were used as indicators in different studies.

3). There were differences in the division of urban and rural populations for comparison. For example, "rural residents" and "agricultural hukou" were used in different studies. Although the two indicators refer to a large overlap but are not exactly the same. "Rural residents" include those who have urban hukou but live in the countryside, while people with "agricultural hukou" include those who were born in rural areas but currently work in cities.

Table 6-4 List of impact-related policy evaluation literature on insurance integration in China

Literature	Data Sources	Methodology	Results	Control Variables	Pros and Cons
Ma, Zhao and Gu (2016)	CHARLS, 2008 and 2012	Standard DID, PSM-DID, individual fixed-effect model	Integration has significantly improved the outpatient visits and costs of rural residents, but the impact on hospitalisation is not significant.	per capita household income, medical expense reimbursement, sex, marital status, age, education, hypertension, smoking, chronic lung disease, major accident, disability, cancer, stomach problems, arthritis, rheumatism, running water, and toilets	Detailed robustness tests
Li (2017)	CHARLS, 2011 and 2013	PSM-DID	After integration, the family medical burden (=family annual medical expenditure / family annual income) significantly reduced.	urban/rural, geographic location (east, middle, west), sex, age, education, marital status, children's relationship, self-reported health, ADL(Activities of daily living) index, number of chronic diseases, family annual income	Lack of robustness tests
Huang and Zhang (2017)	Insurance database, 2014 and 2015	DID	Rural residents' expenditure on health services and the number of visits increased significantly.	sex, age, with chronic diseases or not, and household size	Socioeconomic background cannot be controlled due to database type
Liu (2017)	China Health and Retirement Longitudinal Study, 2011 and 2013	Descriptive statistics, PSM-DID, Monte Carlo simulation, fixed-effect model	Insurance integration had no effect on health service use.	age, sex, marriage, education, household registration, personal income, chronic diseases, smoking, self-reported health, city-level per capita income, community type	Using Monte Carlo simulation to investigate the unobservable confounding

Literature	Data Sources	Methodology	Results	Control Variables	Pros and Cons
Chang, Su and Zhou (2018)	China Labor-force Dynamic Survey, 2012 and 2014	Descriptive statistics, DID, probit regression, mediating effect model	Insurance integration reduced the behaviour of rural residents who do not see a doctor after illness, and improved the self-reported health of urban residents.	age, sex, education, income, smoking, drinking, drinking water source, sanitation, and distance from the nearest medical facility	Urban/rural disparity has been considered
Liu and Lin (2018)	China Health and Retirement Longitudinal Study, 2011 and 2013	Descriptive statistics, statistical tests, PSM-DID, Quantile-DID, probit regression	the overall medical burden increased, while the burden of rural residents decreased, and urban residents were not significantly affected. Further, quantile regression results showed that different quantile groups, that is, groups with different medical burdens, are affected differently by policies, whether in rural or urban areas.	sex, age, urban/rural areas, region, education, self-reported health, ADL, number of chronic diseases, life satisfaction, marital status, children's relationship status, family annual income, children's financial support, fixed assets	Quantile models are used
Ma and Li (2018)	China Health and Retirement Longitudinal Study, 2011 and 2015	Descriptive statistics, PSM-DID	Integrated insurance significantly reduced the likelihood of hospitalisation for residents, but had no effect on hospitalisation costs and outpatient visits; the hospitalisation of agricultural hukou residents significantly reduced, but there was no impact on patients with non-agricultural hukou. Integrated insurance had different effects on people with different health conditions.	sex, age, household registration, education, marriage, smoking, drinking, and family income	Hukou and health status subgroups are taken into account

Literature	Data Sources	Methodology	Results	Control Variables	Pros and Cons
Liu et al. (2018)	Hospital EMR database, 2009 to 2011	Descriptive statistics, statistical tests, regression discontinuity	The average outpatient cost decreased year by year, and the average compensation cost showed an upward trend. The outpatient clinics in township hospitals decreased, and clinic visits increased. The average hospitalisation cost and hospital stay in all hospitals was lower than the national average.	time, age, sex, and hospital level	Method and results related to RD are not presented
Su et al. (2019)	China Health and Retirement Longitudinal Study, 2011 and 2015	Descriptive statistics, statistical tests, PSM-DID, fixed-effect model, log-linear regression and logistic regression	Overall, integration did not significantly impact in the four pilot provinces; but had a significant effect on hospitalisation visit, hospitalisation and outpatient costs for rural residents	sex, age, hukou status, marital status, education level, family income, receive child, given child, live with child, pension and retirement, health status, chronic, smoke, drink, activities of daily living, cognition and depression	Using province/ community fixed effects to eliminate the unobservable confounding bias

In three studies not using CHARLS, Huang and Zhang (2017) results showed that rural residents' expenditure on healthcare (sum of outpatient and hospitalisation) and the number of visits increased significantly in Guangzhou city; Chang, Su and Zhou (2018) results showed that the integration decreased "Did not see a doctor after being sick" (equals to increase visits) of rural residents but had no significant effect of urban residents; Liu *et al.* (2018) results showed that the average outpatient cost decreased year by year. The outpatient visits in township hospitals decreased, and clinic visits increased. The average hospitalisation cost and stay in all hospitals was lower than the national average.

In summary, as can be seen from the above research, different studies have given diverse results about the effect of health insurance integration. However, due to differences in data sources, sample selection, and methods, it is difficult to compare the results of these studies, and challenging to judge which studies are more reliable. Despite this, there are still three points that should be paid attention to:

1). The potential negative effects of integration should not be ignored.

2). In line with theoretical expectations (Chapter 4), outpatient and hospitalisation responses to integration may differ or even oppose each other.

3). Different subgroups within the same population may respond differently or even oppositely to integration.

6.3 Additional Studies on Insurance Integration in Other Countries

Many countries around the world have implemented social insurance systems, and some of them have also undergone the reform of integrating different insurance schemes, such as South Korea, Turkey and Thailand. However, due to the wide variation in political and administrative systems in different countries, the details of insurance reform implementation are not comparable. Thus, this section only provides a brief overview of relevant research on reform in these countries.

In South Korea, Kwon (2009) reviewed the evolution of its health insurance system. Mandatory social health insurance has been in place in South Korea since 1977, and universal health

coverage was achieved in 1989. In order to improve equity in healthcare financing and to reduce the burden of patients, the integration of three main insurance schemes (one for government employees and teachers and their families; one for industrial workers and their families; and one for self-employed persons and workers in enterprises with fewer than five employees) began in 1998 and was completed in 2000. The authors mention that a single provider helps to improve efficiency, whereas, in other countries, such as Mongolia and China, the healthcare financing function is fragmented among the Ministry of Health, Ministry of Labour, Ministry of Finance, etc., resulting in a failure of collaboration among them[29]. As a result, the integration in South Korea significantly reduced administrative costs.

In Turkey, the Social Insurance and the General Health Insurance Law was ratified in 2006 and the integration of five insurance schemes was initiated in 2008 (Atun *et al.*, 2013). The unified general health insurance now offers a comprehensive package of benefits, including reimbursement for a range of preventive, diagnostic and therapeutic services. Unlike China's insurance integration after achieving universal medical coverage, Turkey's integration has greatly increased the coverage of insurance. As a result, patients had significantly lower out-of-pocket costs and a significantly lower catastrophic expenditure rate (this has been attributed to the expansion of the insurance scheme rather than the integration of the insurance scheme).

Thailand's health insurance system also underwent significant changes in the past due to fragmentation and duplication issues (Viroj *et al.*, 2007). In the 1990s, Thailand implemented multiple health insurance schemes, but by 2001, almost 29% of the population remained without coverage. In 2002, the National Health Security Act was approved, leading to the establishment of the National Health Security Office, which aimed at healthcare reform. Similar to Turkey, Thailand adopted a strategy that involved both integration and extension of insurance coverage. As a result, various insurance schemes were integrated into three schemes: the Civil Servant Medical Benefit Scheme, Social Health Insurance, and Universal Coverage Scheme.

[29] The Chinese government is also aware of the issues involved and therefore a single administration (National Healthcare Security Administration) was established in 2018 along with the integration of insurance schemes.

This reform successfully achieved Universal Health Coverage (UHC) in 2003. After the reform, the use of health services has significantly increase.

Lagomarsino *et al.* (2012) conducted a study examining healthcare reforms in nine developing countries. Among these countries, Kenya and Nigeria, in the early stages of reform, initially implemented health insurance schemes targeting the formal sector and civil servants before attempting to extend coverage to the poor and informal sectors. In contrast, India focused on expanding insurance schemes to cover the impoverished population. Ghana, the Philippines, and Vietnam, in the middle stage of reform, opted for a unified insurance scheme. The authors propose that these countries are moving towards integration risk pools. This shift may be driven by the advantages of administering larger risk pools, including reduced complexity, enhanced cross-subsidisation and risk-sharing capabilities, and improved long-term fiscal sustainability.

In summary, countries worldwide have been transitioning from insurance plans targeting specific groups to more comprehensive healthcare plans. However, the specific details of these reforms vary across different countries. For instance, some countries have initially implemented universal medical insurance and then proceeded to integrate various insurance plans, as seen in South Korea and Japan. Certain countries have chosen to integrate insurance plans while simultaneously expanding insurance coverage, such as Türkiye and Indonesia. Furthermore, some countries have ultimately adopted a single insurance plan, like South Korea and Vietnam, while others have maintained two or three insurance plans, as observed in Japan and Thailand. Additionally, there are notable differences in the governing bodies and insurance terms among these countries. These variations collectively limit the direct applicability of findings from other countries to my research. (Viroj *et al.*, 2007; Kwon, 2009; Ikegami *et al.*, 2011; Lagomarsino *et al.*, 2012; Atun *et al.*, 2013; Kim, 2017; Bazyar *et al.*, 2021)

6.4 Research Gap, Aim, and Objectives

The current empirical studies mainly focus on pilot cities or counties, with only a few extending to provinces, indicating a need for broader research. While the limitation is understandable due to data availability, there is merit in pursuing a higher-level analysis for future research.

Additionally, most studies used two-period data and methods to estimate the policy effects, even if data covered many regions. This research design assumed that the policy effects are the same in all regions. However, variations in policy content and implementation time may lead to heterogeneous effects, highlighting a crucial aspect for further exploration.

Moreover, existing research also showed not only the regions but also the sample selection and populations affected the estimation of policy effects. Most of the studies do not give a clear definition of the treatment and control groups and an explicit sample selection process. Even using the same data, the control group used urban/rural residents (including other insurance schemes), urban/rural "Hukou" (the place of living and Hukou are not always the same), or people enrolled in URBMI/NRCMS, can lead to different results. In addition, if a study wants to compare exposed and non-exposed populations or participate and non-participate populations, but the control group includes both non-exposed and non-participate populations, the estimates will be biased. Meanwhile, different populations, such as children/students, the elderly, all insured, and patients who visited hospitals only, also lead to different results. Therefore, conducting a study with clearly defined and uniform design—same data, definition, and methods—will enhance the reliability of results.

In alignment with current government policy objectives and drawing from existing research insights, my research aims to conduct a comprehensive evaluation of China's urban-rural health insurance integration and to explore its population heterogeneity. Specific research objectives are as follows:

Objective 1: To summarise the different regional integration policies and to discuss the potential benefits and problems of integration.

China's administrative structure can be officially (de jure) divided into three tiers from top to bottom: Provincial, Prefectural, and Township (Table 6-5). The provincial level, the highest tier, includes provinces, autonomous regions, and municipalities. They are directly managed by the central government. China has 23 provinces, 5 autonomous regions, 4 municipalities, and

Table 6-5 Administrative divisions of China

Three-Tier System	Five-Tier System	Region Name				
Provincial (1st)	Provincial (1st)	Autonomous region	Province	Municipality	Special administrative region	
Prefectural (2nd)	Prefectural (2nd)	Prefecture/ Autonomous prefecture/ Prefectural-level city/ Subprovincial-level city		New area/District/County	Regions	Municipality
	County (3rd)	District/ County-level city/ Autonomous County/ County				
Township (3rd)	Township (4th)	Subdistrict (street)/ Town/Township			District	Freguesia
	Basic Level (5th)	Community (urban area)/Village (rural area)				

*Note: *The table only lists common region names, not all of them; in total, there are 34 provincial-level regions, 334 prefectural-level regions, 2,862 county-level regions, 41,034 township-level regions, and 704,382 basic level autonomies in 2017.*

2 special administrative regions. Hong Kong, Macao, and Taiwan having distinct health security systems from other regions are excluded from this study. The prefectural level, the second tier, includes autonomous prefectures, counties, autonomous counties, and cities. Finally, the township level, includes townships, ethnic minority townships, and towns. In addition, there is a more practical five-tier system: Provincial-Prefectural-County-Township-Basic Level. The basic level is autonomous, and there is no formal government.

The current integration is at the city level (Prefectural), resulting in numerous integration policies across China. Analysing all these policies is impractical. Since a provincial capital city is a political centre, other cities are likely to follow the policies of the provincial capital city (taking the same or similar policies). Therefore, this study selects provincial capital cities for policy document analysis. The study adopts the perspective of the insured rather than the insurer. Thus, the policy content will include financing modes and benefits packages but will exclude management, funding pool, and drug lists.

The financing mode will include how to divide the non-employed population (potentially divided into children/adults/old people, or normal/poor), the amount of premiums of different groups of people, the amount of government subsidies, and so on. The benefit package will cover the scope and design of insurance. The scope refers to which healthcare institution can be reimbursed (in China, only the expenses in designated healthcare institutions can be reimbursed), and the design refers to deductible, co-payment rate, and ceiling. Details of social insurance reimbursement design in China will be elucidated in Chapter 7.

URBMI and URRBMI operate at the same administrative level, so the policy transfer is one-to-one. Therefore, it is easy to compare. Conversely, NRCMS operates at a lower level than URRBMI, resulting in a many-to-one transformation. In this study, only one NRCMS policy is selected for each city. Since databases do not contain all provincial capitals (Chapters 8 and 9 will describe the databases in detail), the document analysis results will be independent of the data analyses results.

Objectives 2&3: To estimate the effects of integration on health service use and costs, at the national level; and to explore the differences in policy effect among the different populations in socio-economic subgroups, at the national level;

The first part of the empirical analysis is to provide a nationwide estimate of the policy impact of integration on health service use and costs, which reflects whether overall integration has achieved the desired goal. The integration is mainly aimed at improving the benefits of rural residents, and existing literature show that the integration of the pilot area has different effects on urban and rural residents. Although the integration policy itself does not reduce the benefits of urban residents, their benefits may be reduced for two reasons: a) with fixed medical resources, the more rural residents use, the less urban residents use; b) some integration policies provide multi-level packages (different premiums and benefits). Some urban residents, especially those with limited financial means, may opt for a lower-level package, resulting in reduced benefits. Therefore, exploring the differential policies' impact on urban-rural regions and on different income groups are very important part of this study. There are other objectives of health insurance integration, such as controlling chronic diseases, that are important and could be the subject of future research.

Objectives 4&5: To assess the impact of different integration policies on health service use and costs at the city level, using ischaemic heart disease as an example; and to investigate the mechanisms of action of different integration policies at the city level and how they differ, using ischaemic heart disease as an example.

Due to the variations in integration policies across cities and the fact that existing studies have found that integration in some pilot areas has the opposite effect to what was expected, this study assumes that there are still differences in the policy effects, even with the formal implementation. Therefore, it is necessary to estimate the policy effect at a lower level (city-level) because the opposite effects may offset each other in a national-level estimation. Additionally, when data are sourced from the same institution, quantile regression analyses can explore the heterogeneous policy effects for patients with different disease severity.

The study aims to make the following contributions:

1). To provide a comprehensive review of the current stage of health insurance integration policy.

2). To provide a quantitative policy evaluation of full-implemented health insurance integration.

3). To provide a comparison of different integration designs and quantify their heterogeneous effects.

The results may offer suggestions for China's future healthcare reform and potentially inform reform efforts in other developing countries like Mexico. Furthermore, comparisons with insurance integration in other East Asian countries, such as Japan and South Korea[30], can be made to enrich the global understanding of health insurance integration policies.

[30] For example: Kim, Seongjo (2017) *Solidarity, labour, and institution: The politics of health insurance reform in Japan and South Korea.* PhD thesis, University of Sheffield.

Chapter 7 Comparative Qualitative Analysis of Integration Policies in Four Case Cities

7.1 Introduction

In China, the health security system encompasses various government-implemented systems that provide compensation to workers or citizens for income loss and medical expenses resulting from illnesses or other incidents such as childbirth or disability (Changxue *et al.*, 2016). A key component of this health security system is basic medical insurance, which is tailored to local conditions such as the economy, population health, and medical resources to meet the specific needs of residents. The integration of basic medical insurance is carried out at the city level, so this chapter focuses on the first objective of this thesis: to summarise the different regional integration policies and theoretically discuss the possible benefits and problems of integration.

Section 7.2 introduces insurance terms and reimbursement rules. It helps readers who are unfamiliar with reimbursement in China to understand how it works. Section 7.3 provides a summary of integration implementation timeline. The huge time span of integration may be one of the important reasons for the difference in insurance terms across cities. Section 7.4 discusses in detail the differences in insurance terms before and after the integration in case cities. From these cases, potential integration benefits and problems can be deduced. Finally, it is the conclusion of this chapter.

7.2 Insurance Reimbursement Design

In China, the reimbursement of insurance claims is influenced by various factors. Key factors include deductibles, reimbursement rates, and ceilings, all of which play a significant role in determining the amount of reimbursement provided. This chapter provides an analysis of these terms and examines premiums as well. The calculation of the actual reimbursement amount is

illustrated in Figure 7-1. Fees below the deductible are fully paid by the patient; fees between the deductible and the ceiling are reimbursed at a specified rate; fees above the ceiling are fully paid by the patient.

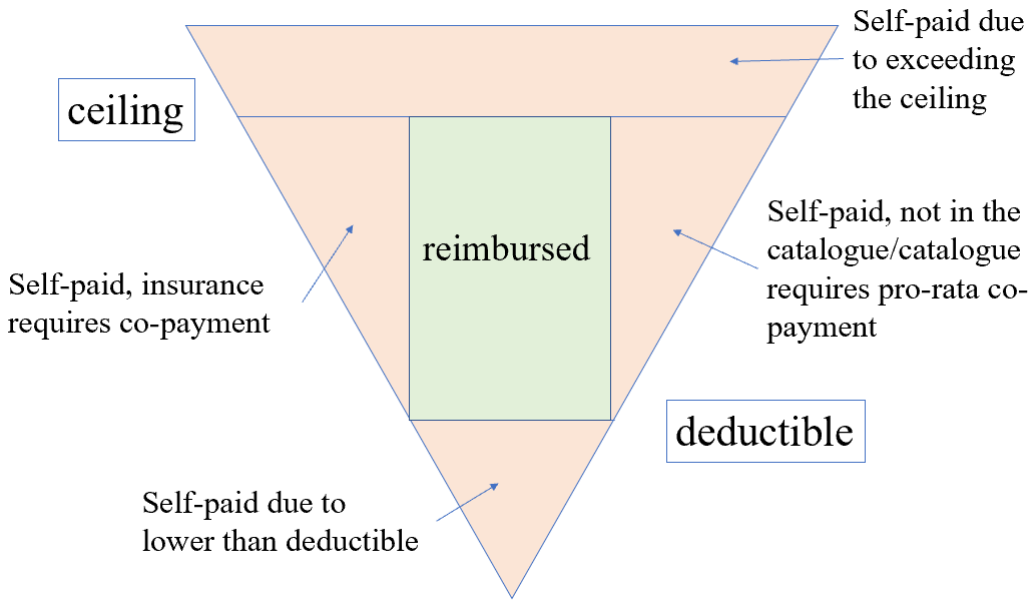

Figure 7-1 Calculation of actual reimbursement amount

Beyond the reimbursement terms (deductible, reimbursement rate, and ceiling), the actual reimbursement in medical insurance follows the principle of "two designated and three catalogues". The "two designated" are the institutions where health services and drugs can be reimbursed and include designated pharmacies and designated hospitals. The "three catalogues" cover the drugs and services available, including the basic medical insurance drug catalogue (there are nearly 200,000 types of medicines in China, and only 2,860 of them are reimbursable under the latest catalogue)[31]; the medical treatment items catalogue, such as examination fees and surgery fees; and the medical facilities catalogue, such as bed fees.

Each catalogue is subdivided into different sub-categories A, B and C. Category A is 100% reimbursed because they are clinically essential and widely used, which is uniformly stipulated by the whole country. Category B is partially reimbursed (usually 80%). It means that after a

[31] http://www.nhsa.gov.cn/art/2021/12/3/art_14_7430.html

certain percentage of the cost is borne by the individual, the remainder will be reimbursed in proportion to the medical insurance term. This is adjusted by each province, city and district according to the local economic level, medical needs and medication habits. Category C is fully self-paid, on the other hand, is not covered by basic medical insurance and the individual has to pay 100% out-of-pocket. This includes some specific healthcare products, high-grade drugs, newly developed drugs, imported anti-cancer drugs, etc.

It can be seen from the above, in clinical practice, the actual insurance reimbursement faces limitations due to factors like the ceiling, deductible, and reimbursement categories. The proportion of actual reimbursement to the total cost is lower than the reimbursement rate in insurance terms, leading to dissatisfaction among insured residents. However, the adjustment of the reimbursement catalogue is not part of the reform of insurance integration, so it is not discussed in detail in this chapter but will be discussed in the final chapter as a suggestion for future research.

7.3 Time of Integration

The integration implementation time is collected through online sources, including public government documents, government notices, public agency notices, local news, etc. Figure 7-2 displays the number of cities that implemented urban and rural basic medical insurance. There are 338 (prefectural-level) cities[32] in China, of which 337 have completed implemented URRBMI, and 1 city, Shenzhen, does not require integration because there are no rural residents.

There is considerable variability in the timing of integration across cities. The earliest implementation of URRBMI was in 2007, which was the year of the formal implementation of URBMI. This suggests that some cities extended rural insurance (NRCMS) to cover urban residents without establishing separate urban insurance. Before 2016, the year of the central

[32] There are two types of a city in China: prefecture-level and county-level. The prefecture-level cities are under the administration of the province and have their own integration policies. The county-level cities are under the administration of the prefecture-level cities and comply with the integration policies of the prefecture-level cities.

government's announcement, 22.85% of cities (77 of 337) had implemented URRBMI, rising to 27.30% (92 of 337) by the end of 2016. Over half of the cities (55.19%, 186 of 337) completed integration in 2017 and 2018, while a smaller number (17.51%, 59 of 337) finalised integration in 2019 and 2020. (Figure 7-2)

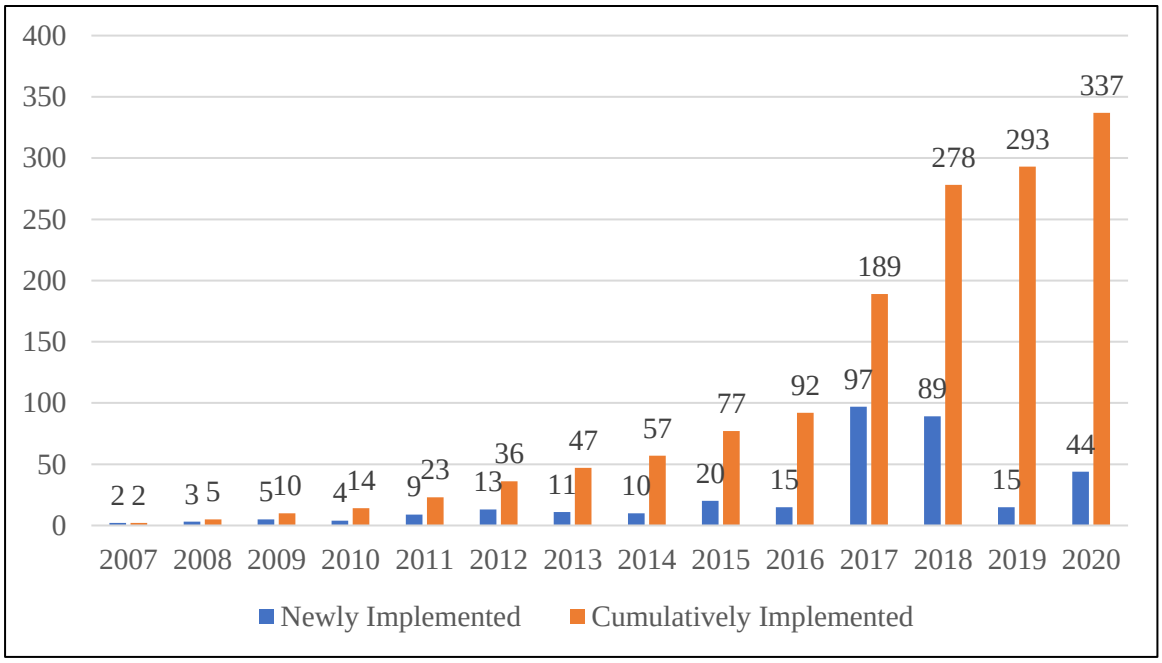

Figure 7-2 Time of implementation of urban and rural resident basic medical insurance

As a result, there are large differences in the way of integration across cities due to varied implementation times. An important feature is the multiple levels of premiums and benefits. Table 7-1 summarises the policy timing of 25 provincial capital cities and municipalities directly under the central government. The early implementation of integration and the more economically disadvantaged areas implemented a multi-level health insurance registration. As integration, in general, has increased the reimbursement benefits for participants and increased the expenditure of the insurance fund, cities that implemented integration early on have opted for multi-level health insurance registration, which is supposed to reduce the burden on the fund or mitigate the excessive growth of the fund expenditure. In contrast, most of the later implementation areas have practically fully unified integrated insurance. The two cities, Harbin and Lasa, offering multi-level registration, are both more economically disadvantaged areas.

Table 7-1 Summary of the policy time of provincial capital cities

Cities	Year	Optional Premiums and Benefits	Cities	Year	Optional Premiums and Benefits
Chongqing	2009	2 Levels	Zhengzhou	2017	×
Chengdu	2009	3 Levels	Beijing	2018	×
Tianjin	2010	3 Levels	Taiyuan	2018	×
Hangzhou	2011	2 Levels	Harbin	2018	2 Levels
Changsha	2011	×	Fuzhou	2018	×
Yinchuan	2012	3 Levels	Wuhan	2018	×
Kunming	2013	×	Nanning	2018	×
Jinan	2015	2 Levels	Lanzhou	2018	×
Guangzhou	2015	×	Nanjing	2019	×
Shanghai	2016	×	Hefei	2019	×
Nanchang	2016	×	Haikou	2020	×
Wulumuqi	2016	×	Lasa	2020	3 Levels
Shijiazhuang	2017	×			

7.4 Policy Cases

The specific details of integration policies are derived from publicly available government documents[33]. This study adopts the perspective of the insured individuals rather than the insurers, focusing on financing modes and benefits packages, but not including management and reimbursement catalogues. The financing mode includes the division of the non-employed population (children, adults, and the elderly), premium amounts for different groups, government subsidies, etc. The benefits package covers the scope and design of insurance. The scope refers to which healthcare institutions can be reimbursed, and the design refers to deductible, reimbursement rate, and ceiling.

Due to the similarity of policies in some areas, it is unnecessary and impractical to analyse all policies across all cities to get an overview of the differences. Therefore, this study provides detailed case studies of integration policies in four cities: Beijing, Tianjin, Chengdu, and Shanghai. They are located in different regions of China: Beijing and Tianjin in the north, Chengdu in the west, and Shanghai in the east. They have different economic development models and levels: Beijing is the capital of China, Tianjin is one of the trade centres of China,

[33] Web links to some of the policy documents can be viewed in Appendix E. However, many of the original documents are continually being removed from government websites and are no longer available.

and Shanghai is the finance centre of China. And they have different policy implementation times: Chengdu in 2009, which was one of the earliest implemented cities, Tianjin in 2010, and Shanghai and Beijing in 2016 and 2018 respectively. Thus, they can serve as representatives of the variations across cities.

The vast majority of insurance reimbursements are based on hospital type or level. In addition to the division of the tertiary hospital system mentioned in Chapter 3, many insurances, especially rural insurance, cover lower-level healthcare institutions. Generally, the quality of health services, from lowest to highest, are: village clinic, urban clinic (which is usually equal to community health centre), township hospital, county hospital, primary hospital, secondary hospital, and tertiary hospital. Three elements of the insurance terms are used for comparison: changes in premium, changes in outpatient reimbursement, and changes in hospitalisation reimbursement. In each subsection, the changes to the terms of case cities are described in detail in turn and a summary is given at the end.

7.4.1 Changes in Premium

In the literature review in Chapter 6, it is mentioned that change in insurance premiums is one of the factors that resist insurance integration and may also cause residents to abandon insurance enrolment. Therefore, changes in premiums before and after integration are discussed first. The premium consists of two parts, one is paid by the residents themselves, and the other part is subsidised by the government.

Before integration, the design of premiums between urban and rural areas in the four cities was various (Table 7-2). Premiums in urban areas usually vary by age (child/adult/elderly) and economic status (standard/poverty and disabled), but not in rural areas.

In Chengdu, the total premiums for rural residents were lower than those for urban adults but higher than those for urban children and students. Government subsidies for rural residents were higher, resulting in lower self-paid insurance premiums than standard urban residents. More specifically, the standard self-paid premium of rural residents (40 CNY) and urban adults

(275CNY) accounted for 0.62% and 1.62% of per capita net/disposable income (in 2008, the per capita net/disposable income of rural residents in Chengdu was 6,481 CNY, and that of urban residents was 16,943 CNY[34]).

Table 7-2 Premiums in case cities before integration (CNY)

City	Year	Area	Population		Premium		
					Total	Self-Paid	Government Subsidy
Chengdu	2008	Urban	Children/Student		80	35	45
			Adult	Standard	320	275	45
				Poor		245	75
				Poor elderly		215	105
				Disabled		96	224
		Rural	Standard		200	40	160
			Poor/Disabled			Free	200
Tianjin	2009	Urban	Children Student	Standard	100	60	40
				Poor/Disable		Free	100
			Adult	Standard	560	330	230
				Age> 70		120	440
				Poor/Disable		Free	560
		Rural	All Residents		110	30	80
Shanghai	2015	Urban	Adults	18-60	1,700	680	1,020
				60-69	3,300	500	2,800
				>70		340	2,960
			Children/Student		750	90	660
		Rural	Standard		1,800	270-400	1,400-1,530
			Poor/Disabled			Free	1,800
Beijing	2017	Urban	Children/Student		1,160	160	1,000
			Standard Adult		1,660	660	
			Elderly: male>60, female>50		1,360	360	
			Poor/Disable		1,360	360	
		Rural	All Residents		1200	160	1040

[34] http://gk.chengdu.gov.cn/govInfoPub/detail.action?id=59362&tn=2

In Tianjin, the situation was partially similar, with rural total premiums lower than urban adults but higher than urban children and students. However, although Tianjin provided more subsidies to urban adults, their self-paid premiums were still much higher than those for rural residents. The standard self-paid premium of rural residents (30 CNY) and urban adults (330 CNY) accounted for 0.28 % and 1.54% of per capita net/disposable income (in 2009, the per capita net/disposable income of rural residents in Tianjin was 10,675 CNY, and that of urban residents was 21,430 CNY[35]). Another difference was that poor and disabled residents in urban Tianjin could enrol in insurance free of charge but not in rural areas, while the situation in Chengdu was the opposite.

In Shanghai, total insurance premiums in rural areas exceeded those for normal urban adults and children but were lower than those for the elderly. However, self-paid premiums for rural residents were generally lower than for urban residents. Despite this, the premium burden of rural residents might be higher than urban residents. The standard self-paid premium of rural residents (230 to 400 CNY) accounted for 1.16% to 1.72% of per capita disposable income, while this ratio is 1.28% for urban adults (in 2015, the per capita disposable income of rural residents in Shanghai was 23,205 CNY, and that of urban residents was 52,962 CNY[36]). In Beijing, rural residents experienced lower total premiums, self-paid premiums, and premium burdens compared to urban adults. The burden of standard premium accounts for 0.66% to rural residents and 1.06% to urban residents (the per capita disposable income of rural residents in Beijing in 2017 was 24,240 CNY, and that of urban residents was 62,406 CNY[37]).

In summary, except for Shanghai, the burden of urban residents, that is, the proportion of self-paid insurance premiums to disposable income, was higher than that of rural residents. In addition, the more economically developed the city, the lighter the burden on urban residents. One reason is the more developed the city's economy, the better the government's financial situation, and the more insurance subsidies it gives residents.

---

35 https://www.tj.gov.cn/sq/tjgb/202005/t20200520_2468068.html

36 https://tjj.sh.gov.cn/tjgb/20160228/0014-287258.html

37 http://www.beijing.gov.cn/zhengce/zhengcefagui/201905/t20190522_60991.html

After integration, the policies of the four cities still showed differences, as outlined in Table 7-3.

Table 7-3 Premiums in case cities after integration (CNY)

City	Year	Population		Premium		
				Total	Self-Paid	Government Subsidy
Chengdu*	2009	Children/Student	Standard	120	40	80
			Poor/Disable		Free	120
		Adult	A	100	20	80
			B	200	120	80
			C	300	220	80
Tianjin	2010	Children/ Student	Standard	100	50	50
			Poor/Disable		Free	100
		Adults	Standard-A	560	330	230
			Standard-B	350	160	190
			Standard-C	220	60	160
			Poor/Disable	220	Free	220
Shanghai	2016	Adults	18-60	2,500	680	1,820
			60-69	3,800	500	3,300
			>70		340	3,460
		Children/Student		900	100	800
		Poor/Disabled		By age	Free	By age
Beijing	2018	Children/Student		1,640	180	1,460
		Standard Adult			300	1,340
		Elderly male>60, female>50			180	1,460
		Poor/Disable			Free	1,640

Note: *disabled adults register for insurance free of charge, but do not specify which level it is.

In Chengdu, the total premium for children and students witnessed a 50% increase, rising from 80 to 120 CNY. However, due to the substantial increase in government subsidies, the self-paid premium experienced a marginal increase of only 5 CNY (from 35 to 40 CNY). For adults, three premium levels were introduced, each accompanied by a government subsidy of 80 CNY: A–a total premium of 100 CNY, B–a total premium of 200 CNY, and C–a total premium of 300 CNY. Notably, even at the highest level, the self-paid premium (220 CNY) was lower than the standard adult premium in the city pre-integration (275 CNY). However, only the lowest level (20 CNY) was lower than the rural premium before integration (40 CNY).

In the aftermath of integration in Tianjin, the total premium for urban children and students remained unchanged, while the self-paid premium saw a reduction of 10 CNY (from 60 to 50 CNY). Similar to Chengdu, Tianjin also featured three premium levels in addition to poor and disabled people. The total premiums ranged from 220 to 560 CNY, while the self-paid part ranged from 60 to 330 CNY. Compared with before the integration, the maximum self-paid premium had not changed, but urban residents could choose lower premiums. However, even at the cheapest level, the self-paid premiums for rural residents doubled (from 30 to 60 CNY). In summary, the new scheme, with no alteration in the cap of total premiums, reduced both urban residents' and the government's spending but elevated spending by rural residents.

Shanghai, being the most economically developed city in China, witnessed the most substantial increase in government subsidies during integration. After integration, for urban children and students, the self-paid premium only rose by 10 CNY, while the government subsidy surged by 140 CNY. The self-paid premium for urban residents remained unchanged, while the government subsidy increased by up to 78% (from 1,020 to 1,820 CNY). However, for non-elderly adults in rural areas, despite the increased government subsidies, the self-paid premiums still increased greatly, ranging from 70% (400 to 680 CNY) to 152% (270 to 680 CNY).

Additionally, Beijing, being the capital of China, adopting relatively conservative stance in policy implementation, experienced an 87.5% increase in self-paid premiums for rural adults (160 to 300 CNY) post-integration, while urban adults witnessed a 54.5% reduction (660 to 300 CNY). Despite an overall increase in government subsidies for all residents, urban residents benefit far more than rural residents.

Overall, there are differences in the changes in premiums across the four cities (Table 7-4). Both early pilot cities (Chengdu and Tianjin) chose to change their contribution models to multiple optional levels. Therefore, for residents, some would pay more, and some would pay fewer premiums after integration. This feature may be due to local financial burden considerations. After the integration of these two cities, the subsidies for some groups of people, such as the elderly, are actually greatly reduced. However, from previous research, multi-level

insurance appears that may introduce adverse selection, whereby healthy people choose the cheaper level of premium and unhealthy people choose the more expensive level. Conversely, Shanghai and Beijing treated all adults equally, with Beijing reducing premiums for all urban residents, potentially fostering easier integration implementation with less resistance from residents. However, for rural residents, premiums have increased in all four cities. This is consistent with previous research, which is detrimental to the introduction of integration because it increases the burden on rural residents and may lead residents to drop out of insurance. The disparities in government subsidies reflect the varying financial affluence levels among cities, with Beijing and Shanghai uniformly increasing subsidies for all residents in contrast to Chengdu and Tianjin.

Table 7-4 Summary of premium changes

Regions			West	North	East	Capital
Cities			Chengdu	Tianjin	Shanghai	Beijing
Year			2009	2010	2016	2018
Self-Paid	Compare to URBMI	Adult	↓	-/↓	-	↓
		Elderly	↓	↑/↓	-	↓
		Disable/Poor	↑/↓	-	-	↓
		Student/Child	↑/↓	↓	↑	↑
	Compare to NRCMS		↑/↓	↑	↑	↑
Subsidy	Compare to URBMI	Adult	↑	-/↓	↑	↑
		Elderly	↑/↓	↓	↑	↑
		Disable/Poor	↑/↓	↓	-	↑
		Student/Child	↑	↑	↑	↑
	Compare to NRCMS		↓	↑	↑	↑

Note: ↑ increase; ↓ decrease; - no change;

7.4.2 Changes in Outpatient Reimbursement

The outpatient reimbursement terms before the integration were also different in 4 cities (Table 7-5). In Chengdu, outpatient services were not reimbursed in either rural or urban areas, so residents had to pay all the fees by themselves. In Tianjin and Shanghai, rural residents enjoyed better reimbursement benefits than urban residents, while it is difficult to say which was better in Beijing. In Tianjin, pre-integration, the urban insured had no reimbursement while the rural insured were subsidised by 20 CNY per year. In Shanghai, the reimbursement rates decreased

with the increase of the hospital level in both urban and rural areas. Despite rural residents having a deductible of 300 CNY, they were fully reimbursed for fees exceeding this amount, while urban residents faced a standard deductible of 1,000 CNY. The ceiling for urban residents was only 35% of that for rural residents (1,750 CNY vs 5,000 CNY). In addition, despite the limited and lower-grade reimbursable medical institutions for rural residents, their reimbursement rates (ranging from 50% to 80%) were significantly higher than those for urban residents (ranging from 50% to 65%). In Beijing, rural residents had lower deductibles, a higher ceiling, and lower reimbursement rates than urban residents.

Table 7-5 Outpatient reimbursement in case cities before integration (CNY)

City	Area	Institution Level	Deductible	Reimbursement Rate	Ceiling
Tianjin	Urban	All	None		
	Rural	Township County Outside County (designated) Outside County (non-designated)	Subsidy 20 per person per year		
Shanghai	Urban	primary	age 18-59:1,000 age>=60/disabled:300	65%	350 per visit; 1,750 per year
		secondary		55%	
		tertiary		50%	
	Rural	Clinic	300	300+80%	5,000
		Community		300+70%	
		District		300+60%	
		City		300+50%	
Beijing	Urban	All	650	50%	2,000
	Rural	Primary	100	50%	3,000
		Secondary/Tertiary (TCM Hospitals)	550	40%	
		Secondary/Tertiary (Other Hospitals)		35%	

Note: TCM-Traditional Chinese Medicine

Table 7-6 summarises outpatient reimbursement terms after integration. In Chengdu, individuals transitioned from having no outpatient benefits to receiving a subsidy of 16 CNY

per person per year. However, this subsidy was notably limited compared to the per capita health expenditure in 2009 for both rural (304 CNY) and urban residents (940 CNY)[38].

Table 7-6 Outpatient reimbursement in case cities after integration (CNY)

City	Institution Level	Deductible	Reimbursement Rate	Ceiling
Chengdu	Subsidy 16 per person per year			
Tianjin	All	800	Children/Student: 30% Adult: A- 40% B-35% C-30%	3,000
Shanghai	Clinic	0	80%	unpublished
	Primary	disabled/age>=60/student/children: 300 18-60 years old:500	70%	
	Secondary		60%	
	Tertiary		50%	
Beijing	Primary	100	55%	4,000
	Secondary/Tertiary	550	50%	

In Tianjin, after integration, benefits were better than before. Students and children, after exceeding 800 CNY in payments, were eligible for a 30% reimbursement rate, with a 3,000 CNY ceiling. Adults enjoyed the same deductible and ceiling as children, with reimbursement rates varying based on premium levels: A-40%, B-35% and C-30%. It should be noted that the insurance level (A/B/C) and drug category (A/B/C) are independent. For example, a patient enrolled in A-level insurance using category B drugs would have a reimbursement rate of 32% (40% multiplied by 80%). Compared to rural residents, urban residents benefited slightly more.

In Shanghai, after integration, the ceiling of outpatient reimbursement was not published. The deductible for urban residents decreased from 1,000 CNY to 500 CNY, while for rural residents, it increased from 300 CNY to 500 CNY. Reimbursement rates increased for both urban and rural residents, but urban residents might benefit more. In Beijing, deductibles decreased for urban residents but remained unchanged for rural residents. Reimbursement rates increased for both urban and rural residents, with a greater increase in rural areas. Ceilings were raised more for urban areas. Overall, the integration favoured urban residents in Beijing, aligning with the trends observed in Tianjin and Shanghai.

[38] Chengdu Statistical Yearbook 2010

Overall, outpatient reimbursement in the four cities has been significantly improved, which may lead to an increase in outpatient visits. However, the impact of this enhancement varies between urban and rural residents due to disparities in reimbursement terms prior to integration, with urban residents standing to gain more post-integration. In addition, the presence of deductibles and ceilings may introduce variations in the influence of policies on different health subgroups. For example, patients with very good or very poor health may be less affected if their expenses fall below the deductible or exceed the ceiling. On the other hand, those with moderate health status and medical expenses within the reimbursement range might be more affected by these policies.

7.4.3 Changes in Hospitalisation Reimbursement

Tables 7-7 and 7-8 summarise the hospitalisation reimbursement items across four cities.

In Chengdu, rural residents exhibited higher reimbursement rates for hospitalisation compared to urban residents, albeit with lower reimbursement ceilings. In Tianjin, reimbursement rates for rural residents were comparable to those of urban residents, ranging from 40% to 60% in township hospitals and from 0% to 50% in county hospitals. However, the reimbursement rates for rural residents dropped sharply to between 10% and 20% for medical treatment outside the county, which were much poorer than those for urban residents. Consequently, the overall benefit for rural residents in Tianjin lagged behind that of urban residents.

In Shanghai, the insurance reimbursement framework exhibited relative simplicity. After paying a deductible of 50-300 CNY according to the level of the hospital, urban residents enjoyed 60%-80% of the reimbursement, not exceeding 9,000 CNY per month. There was no deductible for rural residents, and all medical expenses would be reimbursed from 50% to 80% according to the level of the hospital, no more than 12,000 CNY per month.

Table 7-7 Hospitalisation reimbursement in Chengdu and Tianjin before integration (CNY)

City	Area	Institution Level	Deductible	Reimbursement Rate				Ceiling
Chengdu	Urban	All	Student/Children: 20% of total costs	Student/Children:	80% of total costs: <1000CNY, 65%;1001-5000, 70%;5001-10000,80%,>10000,90%			80,000
		Community	Adjusted according to employee insurance	Adults	65%			67,772
		Primary			60%			
		Secondary			55%			
		Tertiary			50%			
	Rural	Township	50	Standard Adults	90%	Poor/Disabled	95%	50,000
		County	200		65%		70%	
		Higher than county	500		50%		55%	
Tianjin	Urban	Primary	0	Adult/Children/ Student	60%	Elderly	65%	Children/Student: 180,000 Adults/Elderly: 110,000
		Secondary	300		55%		60%	
		Tertiary	500		50%		50%	
	Rural	Township	0	0-200/40%, 200-2000/45%, 2000-5000/50%, 5000-10000/55%, 10000+/60%				Unpublished
		County	0	0-200/0%, 200-2000/35%, 2000-5000/40%, 5000-10000/45%, 10000+/50%				
		Outside County (designated)	1000	20%				
		Outside County (non-designated)	1000	10%				

Beijing's reimbursement design for urban residents was more straightforward than Shanghai's, uniformly offering a 70% reimbursement rate across all institutions, with an annual ceiling of 180,000 CNY. However, rural residents' reimbursements, akin to those in Tianjin's rural areas, were determined according to the actual medical expenses and the hospital level. Primary institutions could reimburse 75%, secondary institutions could reimburse 65% to 80%, and tertiary institutions could reimburse 55% to 67%. The total annual reimbursement could not exceed 180,000 CNY.

Table 7-8 Hospitalisation reimbursement in Shanghai and Beijing before integration (CNY)

City	Area	Institution Level	Deductible	Reimbursement Rate	Ceiling
Shanghai	Urban	primary	50	age <60: 80% age>=60/disabled:90%	300 per day; 9,000 per month
		secondary	100	age <60: 70% age>=60/disabled:80%	
		tertiary	300	age <60: 60% age>=60/disabled:70%	
	Rural	Community District City	0	80% 75% 50%	12,000 per month
Beijing	Urban	All	Student/Children: 650 Adults: 1300	70%	180,000
	Rural	Primary	300	75%	180,000
		Secondary	1,000-20,000:65%; 20,000-50,000: 70%; above 50,000: 80%		
		Tertiary	1,000-20,000:55%; 20,000-50,000: 60%; above 50,000: 67%		

In summary, before the integration, the difference in hospitalisation reimbursement between urban and rural areas was not consistent in different cities. Chengdu urban residents had a higher ceiling but lower reimbursement rates compared to rural residents. In Tianjin, there was little difference in reimbursement rates for low-level hospitals between rural and urban residents, but a significant gap emerged in high-level hospitals. Shanghai's rural residents had a higher ceiling but lower reimbursement rates than urban residents. Beijing exhibited almost no difference in reimbursement between urban and rural residents.

These variations suggest that a national-level analysis may not reveal a significant impact of integration on hospitalisation. Two key reasons support this view. Firstly, since the purpose of integration is to reduce the difference in medical treatment between urban and rural areas by implementing the same insurance terms, the direction of efforts in different cities may be different, for example, the situation before the integration of Shanghai and Chengdu was just the opposite. Therefore, if a similar situation is extended to the whole country, some cities need to increase the reimbursement of urban residents but do not or less improve the reimbursement of rural residents, while some cities do the opposite. As a result, policy effects at the national level are not significant. Secondly, if many cities, especially late-integrated cities, like Beijing, had little difference between urban and rural areas before the integration, then the integration would not have much impact on residents' medical visits and expenses other than simplifying the insurance management and reimbursement process.

After integration, all four cities witnessed a simplification of reimbursement rules, eliminating the previous division of reimbursement rates based on actual medical expenses (Table 7-9). Inferring the impact in Chengdu is complex due to the presence of multi-level insurance. In urban areas, residents opting for A-level insurance experienced lower reimbursement ceilings, accompanied by a drop in reimbursement rates. Conversely, those selecting B or C-level insurance enjoyed an increase in reimbursement rates by up to 20%. Rural residents faced a decrease in both ceiling and reimbursement rates with A-level insurance, but benefited from increased reimbursement rates with B or C-level insurance and increased ceilings with C-level insurance.

In Tianjin, rural residents benefited more than urban residents. Urban residents' benefits also varied based on insurance levels. A-level insurance maintained an unchanged ceiling (110,000 CNY) but saw a 5% increase in reimbursement rates across all hospitals. Opting for B or C-level insurance resulted in reduced ceilings by 20,000 CNY and 40,000 CNY, respectively. While B-level reimbursement rates remained unchanged, C-level rates experienced a 5% reduction across all hospitals. Rural residents, irrespective of their chosen insurance level, experienced significant benefits compared to the pre-integration scenario.

Table 7-9 Hospitalisation reimbursement in case cities after integration (CNY)

City	Institution	Deductible	Reimbursement Rate					Ceiling
Chengdu	Township	50	A	65%	B and Children/ Student	90%	C	90%	A-10,000; B-50,000; C-60,000; Children/Student:80,000
	Community and Primary	100		60%		80%		85%	
	Secondary	200		55%		65%		80%	
	Tertiary	500		35%		50%		65%	
Tianjin	Primary	0	A and Children/ Student	65%	B	60%	C	55%	A-110,000 B-90,000 C-70,000; Children/Student:180,000
	Secondary	300		60%		55%		50%	
	Tertiary	500		55%		50%		45%	
Shanghai	Primary	50	Standard	80%		Elderly/ Disabled		90%	Unlimited*
	Secondary	100		70%				80%	
	Tertiary	300		60%				70%	
Beijing	Primary	300 (adult 1st visit)/150		80%					250,000
	Secondary	800 (adult 1st visit)/400		78%					
	Tertiary	1,300(adult 1st visit)/650		75%					

Note: "Unlimited" See news reports, no official documents were found.

In Shanghai, after integration, the reimbursement terms only changed the ceiling for urban residents and increased the reimbursement rate for rural residents. In Beijing, both urban and rural residents witnessed a rise in the reimbursement ceiling to 250,000 CNY. However, evaluating whether rural or urban residents benefited more from integration was challenging due to the intricacies of reimbursement calculations specific to rural residents before integration.

Overall, contrary to more consistent changes in outpatient reimbursement terms, changes in hospitalisation reimbursement terms were not similar across cities. The integration effect in Chengdu may be weak, potentially due to internal offset effects resulting from multi-level insurances (the effect of a low level is opposite to that of a high level). Tianjin's terms suggested a significant improvement in reimbursement for rural residents, but the same may not hold for urban residents. Reimbursement in Shanghai would be improved for rural residents but might not be for urban residents. Beijing, for both urban and rural residents, primarily witnessed a substantial increase in the reimbursement ceiling, beneficial for those with high medical expenses but potentially having limited impact on others.

These results support previous speculation that the overall effect of integration on hospitalisation may be insignificant. This can be explained in three ways. First of all, because the adjustment of insurance terms involves deductibles and ceilings, people with different medical expenses in the same area are affected differently. For example, if the ceiling increased, patients with high expenses would benefit more while patients with low expenses would be not affected. Second, within a city, the impact on urban and rural residents may be the opposite, resulting in an insignificant integrated impact at the city level. Finally, differences in policies in different cities lead to inconsistent city-level effects, which in turn leads to insignificant effects at the national level.

7.5 Discussion and Conclusion

This chapter addresses the first research objective of this thesis by conducting a qualitative analysis of publicly available government policy documents. The focus is on understanding the

similarities, differences, potential benefits, and issues, associated with insurance integration policies across diverse cities. This chapter starts by describing the timeline of integration policy implementation across prefectural-level cities nationwide. The earliest introduction of urban and rural residents' health insurance was in 2007, followed by subsequent pilot implementation in some cities. In 2016, a state notice was issued, marking the commencement of nationwide promotion and implementation of integration. Finally, the integration work was completed nationwide in 2020. The examination of the insurance design in provincial capital cities reveals that areas, where integrations were implemented earlier or where economic conditions were poorer, have adopted multi-level health insurance registration.

The results of the integration timeline of prefecture-level cities and the insurance design of provincial capital cities jointly suggest that local fiscal conditions may be an important reason for the large differences in the timing of implementation of integration policies in different cities, as well as an important reason for the differences in integration terms and potential integration effects. However, as this is only a qualitative study, the extent to which fiscal and economic conditions are related to differences in the timing and terms of implementation of integration policies cannot be answered at present but could be determined through future quantitative research.

The in-depth analysis of four selected cities achieves the research objective of this chapter. In summary, premium structures and subsidies differ among the four cities, while changes in out-of-pocket premiums are consistent. Tianjin and Chengdu employed multi-level insurance registration, which may introduce new adverse selection, while Beijing and Shanghai did not. Tianjin and Chengdu reduced subsidies for the elderly, but Beijing and Shanghai increased subsidies for all residents, implying differing fiscal prosperity levels. Rural residents, overall, experienced significant increases in out-of-pocket premiums, prompting reconsideration of insurance registration. This reconsideration may stem from the dissatisfaction with rapidly rising premiums (too expensive) and the possibility that the total household premium could be comparable to medical expenses, rendering the insurance seemingly useless. In the final chapter of this thesis, I will discuss in more detail the impact of premium changes.

Regarding changes in outpatient reimbursement terms, there are notable similarities across cities: better deductible, reimbursement rate, or ceiling. These improved reimbursement terms may boost residents' outpatient visit rates and reduce out-of-pocket outpatient costs. However, pre-integration, the outpatient reimbursement terms for urban residents seem to be worse than those for rural residents. This suggests potential greater benefits for urban residents following the integration.

Conversely, although all four cities have simplified their hospitalisation reimbursement terms, the changes have almost no commonality. Due to the adoption of multi-level insurance, the terms in Chengdu and Tianjin vary from person to person. Rural residents in Tianjin, regardless of which level of insurance they choose, receive better benefits than before integration, but this is not the case in Chengdu. This also makes the changes in terms of these two cities not comparable to those in Shanghai and Beijing. Beijing has raised the reimbursement ceiling for all residents, but Shanghai has only raised the reimbursement ceiling for urban residents. Shanghai has uniformly increased the reimbursement rate for rural residents, but the situation in Beijing varies by person (by health status and actual expenses).

Disparities in changes to urban reimbursement terms, distinctions in reimbursement terms for urban versus rural residents, and variations among individuals with diverse health conditions collectively indicate that the overall policy effect of nationwide integration might be inconsequential. This is because policy effects in different subgroups may offset each other. This conclusion emphasises the importance of city-level analysis over country-level analysis in quantitative research, and also underscores the need for policy heterogeneity analysis among different groups, such as subgroup analysis or quantile analysis.

While providing valuable insights, two limitations merit consideration. Firstly, the representativeness of the selected cases for the national integration situation may be insufficient. Although the diversity in locations, policy implementation times, and positioning (economic, financial, political) among the four selected cities, they are relatively economically developed cities in China. Cities with relatively backward economies may exhibit different changes in

insurance reimbursement terms, probably due to their poorer local fiscal conditions. This is evident in the integration timeline and policies of provincial capital cities. These differences may affect the generalisability of the findings.

Secondly, the analysis predominantly focuses on changes in the demand side of health services, neglecting the management aspects and supply side. As a result, this analysis may not be able to comprehensively discuss changes in health service use and costs. Insurance management, like the reimbursement process, could indirectly impact health service use. For example, the two reimbursement scenarios, automatic reimbursement by the hospital during the visit and manual submission of claims by the patient after the visit, are likely to affect patients' decision-making, especially for patients with limited financial resources. However, changes in the reimbursement process are part of the healthcare informatisation reform and go beyond the scope of insurance integration reform, so they are not included in this study.

Moreover, the supply side, such as adjustments of the insurance reimbursable list, as described in Section 7.2, is an important factor in determining the health service use and costs. If there is a significant increase in reimbursable drugs, the medical costs for the entire population may also be significantly affected. Especially when analysing a specific disease population rather than the national population, adjustments to specific drug reimbursement may have a more dominant impact than adjustments to insurance reimbursement terms. However, although the dynamic adjustment of the reimbursement drug list is part of the social insurance reform, it is independent of the insurance integration reform and the two are not synchronised. Therefore, it was not included in this chapter's study. Considering its potential importance, the final chapter of this thesis will also briefly discuss the impact of adjusting the reimbursement list. In the future, studies combining different reforms in China will be helpful to provide a more holistic understanding of health service use and costs.

In conclusion, this chapter conducts a qualitative case analysis around the first research objective of this thesis, which is to summarise the different regional integration policies and to discuss the potential benefits and issues of integration. The main finding is that the changes in

outpatient reimbursement terms are relatively similar across cities during integration, but the changes in hospitalisation reimbursement terms are quite different. This result provides a possible explanation for the insignificant impact of insurance integration on hospitalisation in previous studies, and once again emphasises the need for separate analysis of outpatient and hospitalisation services, as well as consideration of potential heterogeneous policy impacts. However, the qualitative analysis in this chapter may be constrained by two limitations: cases may not adequately represent the entirety of China's integration scenario, and there is a lack of exploration into insurance administration and healthcare supply. The former may prevent the conclusions analysed in this chapter from being generalisable to other regions, while the latter, though outside the scope of insurance integration research, may affect the use and cost of health services.

Chapter 8 National Level Quantitative Analysis of the Impact of Integration on Health Service Use and Health Expenses

8.1 Introduction

This chapter focuses on the second and third objectives of this thesis, which aim to evaluate the impact of integration on health service use and health expenses at a national level and also to analyse policy effects heterogeneity, including the variations between urban-rural areas and between high-low income groups. This chapter attempts to answer two questions. Firstly, whether integration has improved the healthcare accessibility of health care and alleviated the economic burden related to healthcare for residents; and secondly, how different the integration policies are among different resident subgroups. The analysis specifically examines two types of services: outpatient visits and hospitalisations. Each type is studied concerning service use (whether or not to visit and length of stay in the hospital) and out-of-pocket costs.

The next section details the methodology of this chapter, covering data sources, population selection, outcome and control variables selection, as well as the statistical model setting and analysis process. Section 8.3 presents the main analysis results, followed by the results of robustness tests and supplementary analyses in Section 8.4. Finally, Section 8.5 provides a summary and overall conclusion of the empirical work undertaken in this chapter.

8.2 Methodology

This section consists of five parts. Firstly, the source and structure of the database are described, and its strengths and weaknesses are discussed. The population selection process and results are then described and reported. The third part describes the choice of outcome and control variables in detail. This is followed by a description of the detailed statistical methods, model

specification and coding issues. Finally, the statistical analysis framework and steps of the chapter are given.

8.2.1 Data Sources

This chapter relies on data sourced from the China Health and Retirement Longitudinal Study (CHARLS) database. The CHARLS is a large-scale interdisciplinary household study in China. It focuses on collecting high-quality nationally representative data of Chinese residents aged 45 and above, to serve the needs for scientific research in many areas on the middle-aged and elderly populations of China. Data from CHARLS is popularly used in the health field, for example, in studies on health behaviours, health insurance, and health inequality (Ma, Zhao and Gu, 2016; Li, 2017; Liu, 2017).

The CHARLS is carried out by Peking University in China[39]. The survey is usually conducted once every two years. The first wave took place in 2011, covering 10,000 households and 17,500 individuals in 150 counties and 450 villages. It has a panel structure[40]; rounds of 2011, 2013, 2015, and 2018 are currently publicly available[41]. The population of CHARLS comes from 28 provinces, which have variations in a wide-ranging set of socioeconomic factors and other related health, nutritional and demographic measures. This survey uses complex sampling methods to collect a sample that can represent the whole country[42].

---

[39] It is supported by the National Natural Science Foundation of China, the Behavioural and Social Research Division of the National Institute on Aging, and the World Bank.

[40] For more data structure information, see "Zhao, Yaohui, John Strauss, Xinxin Chen, Yafeng Wang, Jinquan Gong, Qinqin Meng, Gewei Wang, Huali Wang. (2020). China Health and Retirement Longitudinal Study Wave 4 User's Guide, National School of Development, Peking University".

[41] Data access via (English available): http://charls.pku.edu.cn/

[42] For more sampling information, see "Zhao, Yaohui, John Strauss, Gonghuan Yang, John Giles, Peifeng (Perry) Hu, Yisong Hu, Xiaoyan Lei, Man Liu, Albert Park, James P. Smith, Yafeng Wang. (2013). China Health and Retirement Longitudinal Study: 2011-2012 National Baseline User's Guide, National School of Development, Peking University." and "Zhao, Yaohui, Yisong Hu, James P Smith, John Strauss, Gonghuan Yang. (2014). Cohort Profile: The China Health and Retirement Longitudinal Study (CHARLS), International Journal of Epidemiology, 43 (1): 61–68."

There are many reasons for choosing this database. Firstly, it can reflect the middle-aged and older population (age>45) of the whole country (large coverage of population), which is necessary for national-level policy effect estimation. Secondly, the survey asks in detail about the interviewees' health service use and expenses, such as “In the last month have you visited a public hospital, private hospital, public health centre, clinic, or health worker’s or doctor’s practice, or been visited by a health worker or doctor for outpatient care?”, and “What was the medical cost for all the hospitalisations you received during the past year? (Only include fees paid to the hospital, including ward fees but excluding wages paid to a hired nurse, transportation costs, and accommodation costs for yourself or family members.)”, therefore, it includes key variables the study needed, and it also contains sufficient information that can be used for controlling or matching, such as demographics, health status and function, and biomarkers. Thirdly, it is a bi-annual survey, which means that the data is collected from the same individuals who are followed up every two years. Therefore, the data structure is suitable for some policy evaluation methods, such as panel data models and the DID method.

On the other hand, the drawback of using the CHARLS database is that it contains only data on middle-aged and elderly people. An absence of young people may lead to an overestimation of policy effects on health service use and health expenses (on average, older people are obviously in worse health and therefore have more medical needs and higher expenses). In addition, although the survey team has taken steps to reduce recall bias, such as asking interviewees to use receipts to prove their medical expenses, slight recall bias is inevitable, such as income and the frequency of hospital visits.

8.2.2 Sample Selection

In this chapter, the Difference-in-Difference (DID) and Difference-in-Differences-in-Differences (DDD) methods are employed to identify policy effects, with specific model settings detailed in subsection 8.2.4. A key aspect of this analysis involves selecting appropriate treatment and control groups, primarily based on the type of insurance held by individuals. Individuals who have or use urban and rural resident medical insurance (“城乡居民医疗保险”), or urban resident medical insurance (“城镇居民医疗保险”), or new rural cooperative

medical insurance ("新型农村合作医疗保险") are used in the DID estimation. In this population, the treatment group consisted of residents located in areas that implemented integrated insurance by the end of 2018, while the control group included residents in areas without integrated insurance at that time. In addition, a non-resident group is also formed, including individuals who have employee medical insurance ("城镇职工医疗保险") or do not have any of the above insurance types. However, the non-resident group is used solely as an additional control group for DDD estimation and is not included in the basic DID estimation. Individuals with Government Insurance Scheme[43] ("公费医疗") are excluded from the analysis. The reason is that their treatment is entirely free of charge, so it is not influenced by external factors other than their health condition. Thus, these individuals are incomparable to other individuals.

There are 76,690 observations in the database, of which 73,416 observations are included in the analysis, while 3,274 observations are excluded due to: 1) being younger than 45[44] or missing age information; 2) having government insurance; 3) Identification (ID) missing; 4) in Shenzhen city[45] (Figure 8-1). In the enrolled sample, 53,965 observations belong to the resident group (used in DID) and 19,451 belong to the non-resident group (used in DDD). Within the resident group, 48,146 observations are from areas that implemented integration before the end of 2018 (the treatment group in DID), and 5,819 are from other areas (the control group in DID). For the non-resident group, 17,255 observations are from areas that implemented integration before the end of 2018, while 2,196 are from other areas.

[43] As mentioned in Chapter 2, Section 2.2, the Government Insurance Scheme is a remnant of the old system and represents only a very small percentage of the population.

[44] Two reasons for being younger than 45 years: 1. Incorrect completion of the questionnaire; 2. Spouse of the sampled respondent is younger than 45 years (both the sampled respondent and the spouse were included in the interview, but no age limit was set for the spouse, see https://forum.charlsdata.com/viewtopic.php?t=8572)

[45] There are no rural residents in Shenzhen, so the reform of integrated insurance does not exist.

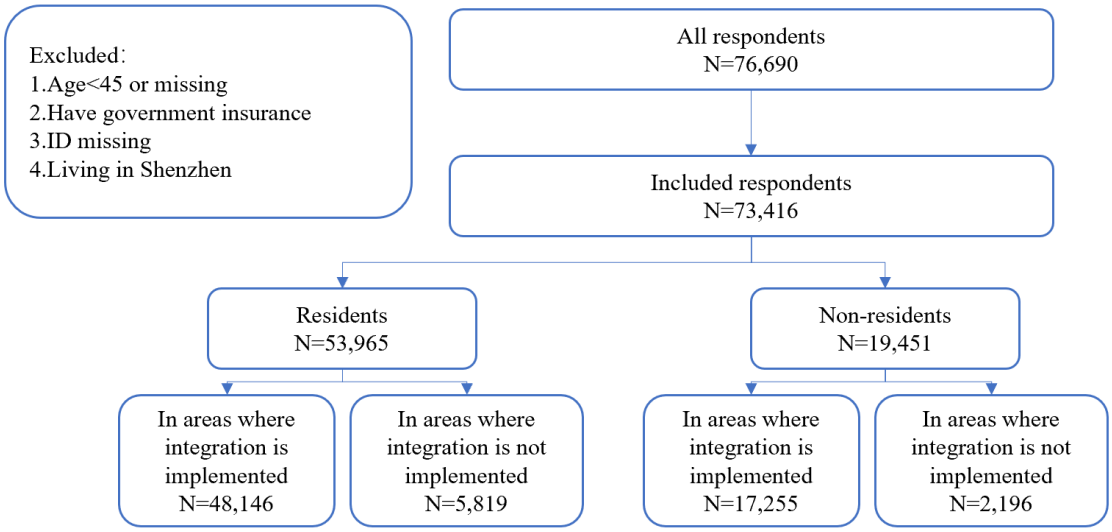

Figure 8-1 Sample selection

8.2.3 Variables Selection

This chapter involves the selection of variables, encompassing both outcome (dependent) and control (independent) variables, with a separate description for each.

Outcome (dependent) variables

As mentioned in Chapter 4 on economic theory, different types of services need to be analysed separately. Therefore, in this study, a total of 5 outcome variables are categorised into outpatient and hospitalisation services, as displayed in Table 8-1.

Table 8-1 Summary of outcome variables

Service types	Time	Variable	Variable type
Outpatient	Last month	Occurrence of outpatient visits	Binary
	Last time	Out-of-pocket costs for outpatient visit	Cost
Hospitalisation	Last year	Occurrence of hospitalisation	Binary
	Last time	Length of Stay (LOS) in hospital	Count
	Last time	Out-of-pocket costs for hospitalisation	Cost

In accordance with the research objectives and questions, this chapter uses three outcome variables to assess the accessibility of health services: occurrence of outpatient visits, occurrence of hospitalisation, and length of stay (LOS) in hospital; and uses two outcome

variables to measure the residents' healthcare-related financial burden: out-of-pocket (OOP) costs for outpatient visit and out-of-pocket costs for hospitalisation. The five original questions of these variables from the questionnaire are provided in Appendix F.

Control (independent) variables

The selection of suitable control variables from the database is guided by the theoretical literature review in Chapter 4 and the empirical research review in Chapter 6. These variables are grouped into five categories: demographics, health status, behaviours, affordability, and macro indicators. When necessary, variables are re-coded to ensure their appropriateness for the analysis. The original questionnaire questions are provided in Appendix F.

Demographic variables, termed predisposing components in Andersen's behavioural model, include sex, age, education level, and marriage/cohabitation. Sex is a dummy variable (1=male, 0=female). Age is calculated by subtracting the year of birth from the year of the interview. Education level is re-coded to rank from 1 (illiterate) to 5 (Doctoral degree). Marriage/cohabitation is a dummy variable (1=yes, 0=no) that combines the two questions of "whether married" and "whether cohabited". The purpose is to measure whether the respondent has a partner who can provide necessary life/medical help.

The second category is health status. The theoretical review in Chapter 3 shows that health status is the most critical factor that determines the use of health services and health expenses. The most direct variable is self-rated health status, which is a categorical variable from 1 (best) to 5 (worst). It has been widely used in previous empirical studies (Chapter 6). However, a previous study (Mu, 2014) shows that individuals in different provinces may employ varying criteria when evaluating their health status, so self-rated health in CHARLS may underestimate regional health differences. Therefore, three additional objective disease (count) variables are added to control for individuals' health status: "the number of disabilities", "the number of chronic diseases" and "the number of cancers". They can also be found in previous empirical studies (Ma, Zhao and Gu, 2016; Huang and Zhang, 2017; Liu and Lin, 2018). There are 5

types of disabilities, 14 types of chronic diseases, and 23 types of cancers, which are determined by the CHARLS project team (see the appendix for details).

Behavioural variables included "current smoking" (dummy variable, 1=yes, 0=no), "drinking alcohol more than once per month" (dummy variable, 1=yes, 0=no), and "kind of social activities in the past month" (count variable, 11 types determined by the CHARLS project team, see Appendix F for details). Some previous studies used the Activities of Daily Living (ADL) to measure the activity capacity of individuals. However, it is calculated that 30% of the individuals are missing this indicator. If ADL is included in the model, the sample size for the regression will be drastically reduced, which will reduce the regression accuracy. Therefore, in the end, ADL is not used as a control variable.

Fourthly, affordability, which is also called enabling component in Andersen's behavioural model. The most important variable is income, which determines whether patients are able to seek medical care. In this study, family income is used as a control variable, which is calculated by summing wages, bonuses, pensions, compensation, subsidies, assistance, donations, crops and forestry products (sold or consumed), livestock and aquatic life (sold or consumed), livestock products (sold or consumed), income from operating activities. In addition to basic health insurance, the reimbursement of other types of medical insurance also directly affects individuals seeking health services. Therefore, another insurance dummy variable, "whether there is commercial insurance", is also used as a control variable.

The last category is macro variables. Chapter 3 has already pointed out the huge differences between urban and rural areas and among different provinces in China's economy and healthcare, and the policy comparison in Chapter 7 also highlights the city-level differences in policy implementation. Therefore, one macro control variable is the urban-rural dummy variable (1=urban, 0=rural). The other is the city fixed effect, which is a series of city dummy variables. They reflect unobservable macro factors, such as the quantity of medical resources and the price of health services.

8.2.4 Casual Identification Strategy

As mentioned in Chapter 5, the challenge of using observational data to identify causality is in eliminating selection bias. Previous empirical studies have commonly used the standard two-period DID, sometimes combined with Propensity Score Matching (PSM). However, as discussed in Chapter 6, an important limitation of the standard two-period DID is the inability to conduct statistical tests to verify whether the data could satisfy an important assumption of DID: parallel trends. Moreover, the PSM-DID may not provide more reliable results because non-parallel trends due to unobserved confounding cannot be effectively controlled.

Drawing from the existing literature, I also explored the use of the proportion of resident insurance in the population as an Instrumental Variable (IV). It is considered a good IV in regression without control variables. However, when cities are included in the regression as control variables, the test result indicated that the IV is not necessary[46].

Consequently, building upon the statistical methods of previous studies, this chapter primarily adopts the multi-period staggered-DID method as the main analytical approach. Furthermore, pre-integration parallel trend tests and placebo tests are conducted to assess the robustness of the DID estimation results. In cases where the pre-integration parallel trend is not satisfied, a supplementary analysis using the multi-period staggered-DDD method will be performed.

Referring to Imbens and Wooldridge (2009), Li, Lu and Wang (2016), Chen *et al.*(2020), Lu, Dong and Ye (2021), Lin, Shen and Sun (2022), and combined with the research background of this chapter, a standard multi-period staggered-DID (two-way fixed effect) model is as follows:

$$\boldsymbol{y_{it} = \alpha + \beta treated_{it} + \theta' X_{it} + \lambda_i + \mu_t + \varepsilon_{it}}$$

Where the $\boldsymbol{\lambda_i}$ is the "individual fixed effect", that is, a series of individual dummy variables; the $\boldsymbol{\mu_t}$ is the "time fixed effect", that is, a series of year dummy variables; $\boldsymbol{treated_{it}}$ is a dummy variable that takes the value of 1 if the integration policy has been implemented for

[46] Wald test of exogeneity: chi2(1) = 0.64, Prob > chi2 = 0.4233

individual i at time t, and takes the value of 0 in other cases. The coefficient of $\boldsymbol{treated_{it}}$ represents the policy effect.

However, there are some issues when applying the above models in the preliminary analysis. Firstly, the panel fixed effect model is not suitable for some model specifications such as the Generalised Linear Model (GLM) and Probit Model, resulting in a great sample loss[47]. Secondly, the individual fixed effects display multicollinearity with variables such as age, sex, urban-rural area, and cities[48]. As discussed in Chapter 5, GLM is better than the general linear model when used for the regression of cost data, and the urban-rural area and city effects are important control variables discussed in subsection 8.2.3. Thus, the above model ends up not being used, and the following model is employed:

$$\boldsymbol{y_{it} = \alpha + \beta treated_{it} + \theta' X_{it} + \nu_i + \mu_t + \varepsilon_{it}}$$

Where the $\boldsymbol{treated_{it}}$ and $\boldsymbol{\mu_t}$ are the same as before. $\boldsymbol{\nu_i}$ is the city fixed effect, that is, a series of city dummy variables, replacing the "individual fixed effect".[49]

The robustness test of the results includes two parts, one is the pre-integration parallel trend test, and the other is the placebo test.

Drawing on the works of La Ferrara, Chong and Duryea (2012), Wang and Wu (2019), and Chen *et al.*(2020), the idea of the pre-integration parallel trend test is to estimate the outcome

[47] In Stata software, the two-way fixed effect model has two equivalent estimation commands, taking the general linear model as an example: 1. "reg y i.did i.id i.year"; 2. "xtreg i.did i.year , fe". However, due to the microscopic nature of this study, the number of individual IDs is very large. If individual fixed effects are used, the first command cannot be estimated, and only the second command can be used, that is, the "xt-" series of commands. However, the "xt" command does not work with GLM and probit, and when using "xtlogit", a large number of samples (59% of observations lost when estimating policy effect on outpatient visits) are lost due to "all positive or all negative outcomes".

[48] The appendix provides a comparison of results for different model specifications, using outpatient visits as an example.

[49] In the pre-analysis, the model using "implemented/not implemented" grouping dummy variables, the model using individual fixed effects, and the model using city fixed effects have the same direction estimation results, so it is feasible to use urban fixed effects instead of individual fixed effects to replace the "implemented/not implemented" grouping dummy variables.

difference between the treatment group and the control group in each period before the policy implementation, and then to compare the differences in each period with the difference in a specific time (typically chosen as one year before the implementation). If the coefficients are not statistically significant, it indicates that there was no significant change in the differences between the treatment group and the control group before the policy implementation, thus supporting that the parallel trend is likely to be satisfied. Conversely, if the coefficients are statistically significant, it suggests a significant change in the differences between the two groups, thereby the parallel trend is like to be violated.

For the analysis in this chapter, as the year of policy implementation varies from city to city, years before the policy cannot use the “natural year/absolute year” but use the difference between the observation year and the implementation year, that is, the "relative year". For example, if the policy is implemented in 2015 and individuals are interviewed in 2012, this year is marked as “3 years ago”, while individuals interviewed in 2018 is marked as "3 years later".

Referring to La Ferrara, Chong and Duryea (2012) and Li, Lu and Wang (2016), the purpose of the placebo test is to exclude the influence of other policies. Its idea is to falsify the treatment group to estimate, that is, to repeat multiple (usually 500 times) random selections of individuals as the treatment group, and check whether the coefficient of the "pseudo-policy dummy variable" is significant. If the coefficient of the pseudo-policy is significant and the original DID estimator is within the distribution of the sampling estimator, it means that the other factors has affected the outcome variable; if the sampling coefficient is not significant and the original estimator is outside the distribution of the sampling estimator, it means that evaluated policy did have an impact.

When the parallel trend is violated, multi-period staggered-DDD is used as a complementary analysis. Referring to Chetty, Looney and Kroft (2009), La Ferrara, Chong and Duryea (2012), and Wang and Wu (2019), on the basis of the DID model, the non-resident population (resident/non-resident group dummy variable) is introduced, and it interacts with the year fixed effect and the city fixed effect. The model is as follows:

$$y_{it} = \alpha + \beta treated_{it} + \sum resident_i \times year_t + \sum resident_i \times city_i + \sum city_i \times year_t + \gamma resident_i + \theta' X_{it} + \nu_i + \mu_t + \varepsilon_{it}$$

Where the new $treated_{it}$ is a dummy variable that takes the value of 1 if the integration policy has been implemented for individual i belonging to the resident group at time t, and takes the value of 0 in other cases. The coefficient of $treated_{it}$ represents the policy effect. The ν_i and μ_t are the same as before; $resident_i$ is a dummy variable to distinguish between resident and non-resident groups. The model also includes pairwise interaction terms for city, year, and resident.

8.2.5 Statistical Analysis Framework

Figure 8-2 shows the step-by-step statistical analysis process in this chapter. The first step is the descriptive statistical analysis. All outcome and control variables are described year by year in both treatment and control groups, giving an overview of the data. For continuous variables, the number of observations, mean, and standard deviation are reported; for categorical variables, the number of observations and frequencies are reported. The number of chronic diseases, the number of cancers, the number of disabilities, and the types of social activities are not suitable for reporting as categorical variables due to numerous types and large count values. In addition, due to the large number of 0 values, the mean value is too small when they are reported as continuous variables. Therefore, these four variables are reported as dummy variables, that is, whether there are chronic diseases/ cancer/ disability/ social activities.

In the second step, DID estimation is performed. Appropriate regression model specifications are selected based on the characteristics of the different outcome variables as mentioned in Chapter 5. Binary variables like the occurrence of outpatient visits and hospitalisation are analysed using logit and probit models. The selection between these two models uses the Akaike information criterion (AIC) and Bayesian Information Criterion (BIC). The count variable, the LOS, is analysed using the Poisson model and Negative Binominal Model. The likelihood ratio test is used to select between them. Cost variables, OOP costs for outpatient visits and OOP costs for hospitalisation are heavily skewed. Thus, the Generalised Linear

Model based on gamma distribution (GLM(gamma)) is considered most appropriate. Moreover, considering the correlation of individuals over different years, all models adopt the standard error of individual ID clustering correction. The predictive marginal values, that is, the marginal effect, are used to reflect the policy effect rather than using coefficients. Therefore, in subsection 8.3.2, only marginal effect results for policy effects for all outcome variables are given. The full regression results are provided in Appendix F and the results of using the "xt -" command (if applicable) are also included in Appendix F for comparison.

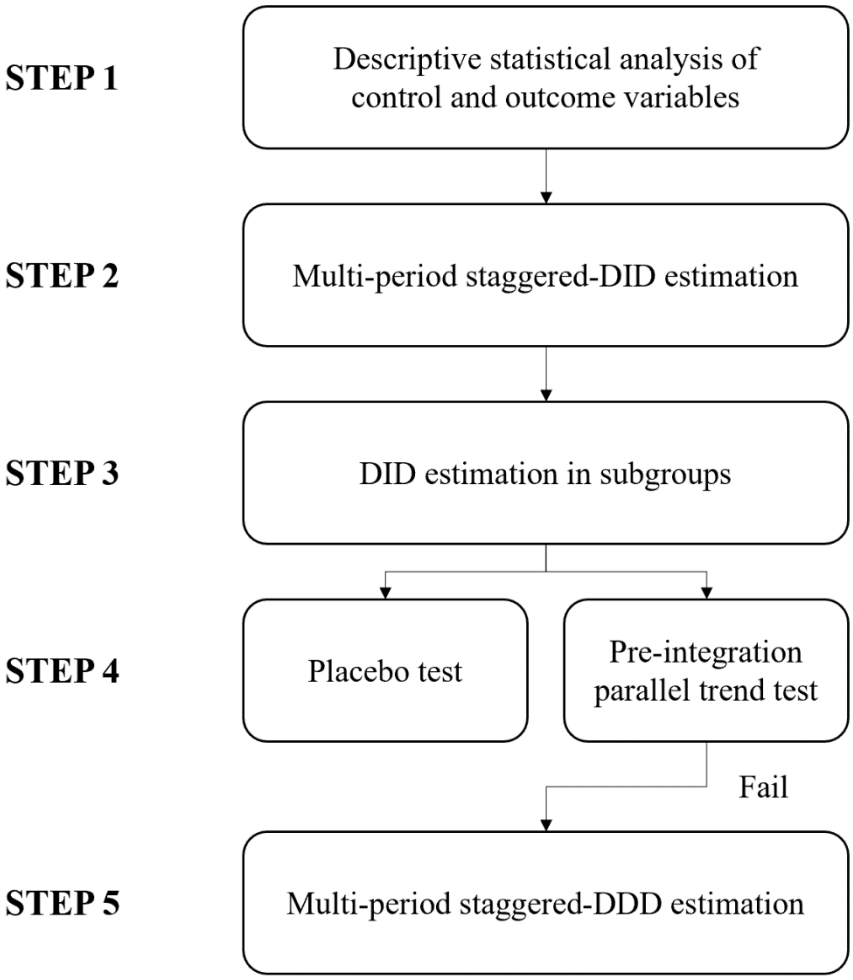

Figure 8-2 The statistical analysis framework of Chapter 8

The third step is subgroup analysis, using the same models and variables as the second step. The purpose here is for the third objective in this thesis to explore whether socioeconomic inequality leads to heterogeneity in policy effects. Therefore, income subgroups divided by

median (upper 50% and lower 50%) and regional subgroups (urban and rural) are analysed. When there are significant differences in the results, the interaction subgroups of urban/rural areas and income groups are further analysed, namely the urban high-income subgroup, urban low-income subgroup, rural high-income subgroup, and rural low-income subgroup. Similarly, marginal effect results for policy effects are given in subsection 8.3.3 and full regression results are provided in Appendix F.

The graphical results of the pre-integration parallel trend test and placebo test are reported in subsections 8.4.1. Full regression results for pre-integration parallel trend tests and descriptive statistics results for sampling in placebo tests are also provided in Appendix F. Supplementary DDD estimates are conducted only for analyses that fail the pre-integration parallel trend test. Subsection 8.4.2 reports the marginal effects estimated by DDD, and the full regression results are provided in Appendix F.

All statistical analyses are performed using STATA 16.0 software. Appendix G contains the code for outpatient visits, serving as an example, which includes DID estimation, pre-integration parallel trend test, placebo test, and DDD estimation.

8.3 Main Results

This section sequentially reports the first three parts of the five-part analysis: descriptive statistics results, DID regression results, and subgroup regression results.

8.3.1 Descriptive Statistics

Table 8-2 provides a summary of descriptive statistics for the outcome variables of the treatment and control groups of residents. Both groups exhibited similar trends in outpatient visit rate and hospitalisation rate, respectively. Although both groups showed a decline in outpatient visit rates, the control group experienced a greater decrease, resulting in the treatment group surpassing the control group after 2015. Hospitalisation rates increased in both groups, with a widening difference between them, possibly influenced by integration policies. Changes in the LOS were inconsistent between the two groups. The treatment groups displayed an overall

upward trend, while the control group increased first and then decreased, resulting in a very small overall change. In terms of OOP costs, whether outpatient or hospitalisation, both treatment and control groups showed upward trends.

Tables 8-3 to 8-5 provide summaries of descriptive statistics for control variables. Overall, the treatment and control groups exhibited similar values and trends for most variables. Both groups had an average age of around 60 years, with fluctuations in family income of around 17,000 CNY (treatment) and 16,000 CNY (control). More women than men were present in both groups, and urban residents constituted approximately 30% (treatment) and 25% (control). The education distribution was also similar between the two groups, with around 50% without formal education, around 22% with primary education, around 25% with middle education, and less than 1% with higher education. Marriage/cohabitation rates were around 86% for both groups.

Regarding health status, both groups experienced an increase in individuals reporting the best health level, from 3% to about 12%. The proportion reporting good health remained stable at around 12%, while those reporting average health increased from 30% to nearly 50%. The proportion reporting worse health decreased from over 35% to 20%, and the worst health status sharply declined from nearly 20% to approximately 5%. In terms of specific health conditions, the proportion of individuals with at least one disability decreased from about 20% to around 12%, and those with at least one chronic disease decreased from about 70% to 44%. However, the proportion of individuals with at least one cancer increased, rising from 1.04% to 1.8% in the treatment group and from 0.38% to 1.17% in the control group. These change trends suggest that residents' health status has a clear temporal trend, indicating impacts on health service use and medical costs even without policy integration. The proportion of individuals engaging in at least one social activity increased from 44% to approximately 50%. The smoking rate declined from 30% to 26%, while the drinking rate remained relatively stable, ranging from 23% to 25%. Lastly, the proportion of individuals with commercial insurance increased from 1.33% to 2.71% in the treatment group but decreased from 3.02% to 2.85% in the control group.

Table 8-2 Summary of descriptive statistics for outcome variables

Outcome	Group	Year											
		2011			2013			2015			2018		
		Obs.	Frequency		Obs.	Frequency		Obs.	Frequency		Obs.	Frequency	
Outpatient visit occurrence	Treatment	11,355	21.08%		11,986	23.74%		12,273	22.10%		12,398	16.91%	
	Control	1,310	22.06%		1,347	23.76%		1,591	19.86%		1,544	15.09%	
Hospitalisation occurrence	Treatment	11,414	9.49%		12,016	13.67%		12,270	14.04%		12,398	17.20%	
	Control	1,323	9.15%		1,356	12.98%		1,591	13.45%		1,544	15.48%	
		Obs.	Mean	S.D	Obs.	Mean	S.D	Obs.	Mean	S.D	Obs.	Mean	S.D
Length of Stay (LOS)	Treatment	1,023	11.63	10.70	1,507	12.66	14.93	1,612	12.10	16.04	2,081	12.56	20.08
	Control	108	10.81	7.50	166	12.97	9.25	191	12.42	13.03	238	10.90	7.99
Out-of-pocket costs for outpatient visit	Treatment	2,138	537	4,049	2,445	655	3,132	2,382	847	3,740	2,046	798	3,569
	Control	262	366	843	272	807	4,691	269	618	1,597	226	642	2,234
Out-of-pocket costs for hospitalisation	Treatment	967	4,719	10,608	1,323	4,750	9,238	1,288	5,194	10,532	1,973	5,784	13,358
	Control	103	2,967	4,136	143	3,455	7,249	148	4,879	8,344	224	4,809	9,776

Note : Obs.-Observations; S.D-Standard Deviation

Table 8-3 Summary of descriptive statistics for control variables-1

Control Variables		Group	Year											
			2011			2013			2015			2018		
			Obs.	Mean	S.D	Obs.	Mean	S.D	Obs.	Mean	S.D	Obs.	Mean	S.D
Age		Treatment	11,421	59.20	9.72	12,034	60.06	9.86	12,281	60.27	10.08	12,401	62.38	10.05
		Control	1,324	58.64	9.11	1,358	59.30	9.22	1,592	58.82	9.60	1,544	61.24	9.76
Family income		Treatment	11,363	17,760	45,715	11,995	15,680	59,305	12,257	13,162	95,646	12,374	22,116	101,519
		Control	1,322	19,978	26,109	1,355	15,232	30,507	1,591	8,372	42,011	1,540	21,565	57,195
			Obs.	Frequency		Obs.	Frequency		Obs.	Frequency		Obs.	Frequency	
Sex	Male	Treatment	11,421	47.17%		12,034	46.53%		12,281	46.74%		12,401	45.50%	
		Control	1,324	47.66%		1,358	46.69%		1,592	47.17%		1,544	44.75%	
Region	Urban	Treatment	11,421	30.15%		12,034	31.26%		12,281	31.19%		12,401	31.31%	
		Control	1,324	25.15%		1,358	25.11%		1,592	25.50%		1,544	26.10%	
Education	No formal education	Treatment	11,408	50.32%		12,024	49.98%		11,354	49.89%		12,401	50.01%	
		Control	1,324	52.42%		1,356	53.24%		1,431	50.73%		1,544	51.04%	
	Primary school	Treatment	11,408	23.30%		12,024	22.77%		11,354	23.46%		12,401	23.55%	
		Control	1,324	21.45%		1,356	20.87%		1,431	21.87%		1,544	21.50%	
	Middle school	Treatment	11,408	26.08%		12,024	26.83%		11,354	26.26%		12,401	26.16%	
		Control	1,324	25.91%		1,356	25.52%		1,431	27.04%		1,544	27.27%	
	Associate or Bachelor's degree	Treatment	11,408	0.30%		12,024	0.32%		11,354	0.32%		12,401	0.28%	
		Control	1,324	0.23%		1,356	0.37%		1,431	0.35%		1,544	0.19%	
	Master or Doctoral degree	Treatment	11,408	0.00%		12,024	0.11%		11,354	0.06%		12,401	0.00%	
		Control	1,324	0.00%		1,356	0.00%		1,431	0.00%		1,544	0.00%	

Note: Obs.-Observations

Table 8-4 Summary of descriptive statistics for control variables-2

Control Variables		Population	Year							
			2011		2013		2015		2018	
			Observations	Frequency	Observations	Frequency	Observations	Frequency	Observations	Frequency
Marriage or cohabitation		Treatment	11,421	87.46%	12,034	86.57%	12,281	86.74%	12,401	84.88%
		Control	1,324	87.99%	1,358	88.81%	1,592	87.50%	1,544	85.69%
Health	Best	Treatment	11,412	3.07%	11,975	4.92%	11,589	5.89%	12,364	11.18%
		Control	1,324	3.02%	1,353	4.95%	1,511	8.21%	1,541	12.72%
	Good	Treatment	11,412	11.65%	11,975	10.91%	11,589	10.50%	12,364	13.30%
		Control	1,324	11.10%	1,353	11.68%	1,511	12.11%	1,541	12.52%
	Average	Treatment	11,412	30.78%	11,975	30.99%	11,589	31.17%	12,364	48.65%
		Control	1,324	29.91%	1,353	30.23%	1,511	28.19%	1,541	47.57%
	Bad	Treatment	11,412	36.34%	11,975	36.24%	11,589	36.06%	12,364	20.75%
		Control	1,324	38.22%	1,353	35.03%	1,511	34.28%	1,541	22.00%
	Worst	Treatment	11,412	18.17%	11,975	16.94%	11,589	16.39%	12,364	6.13%
		Control	1,324	17.75%	1,353	18.11%	1,511	17.21%	1,541	5.19%
Disability		Treatment	11,418	18.92%	12,002	13.36%	12,268	14.53%	12,385	13.68%
		Control	1,324	20.62%	1,357	11.50%	1,592	15.33%	1,543	11.47%
Chronic disease		Treatment	11,418	67.57%	12,002	65.18%	12,268	63.08%	12,385	44.04%
		Control	1,324	70.92%	1,357	66.10%	1,592	59.23%	1,543	44.39%
Cancer		Treatment	11,418	1.04%	12,002	1.22%	12,268	1.09%	12,385	1.80%
		Control	1,324	0.38%	1,357	0.29%	1,592	1.07%	1,543	1.17%

Table 8-5 Summary of descriptive statistics for control variables-3

Control Variables	Population	Year							
		2011		2013		2015		2018	
		Observations	Frequency	Observations	Frequency	Observations	Frequency	Observations	Frequency
Social interactions	Treatment	11,418	44.37%	12,002	50.57%	12,268	49.93%	12,385	50.43%
	Control	1,324	44.11%	1,357	48.27%	1,592	47.86%	1,543	47.12%
Smoke	Treatment	11,418	29.15%	12,006	30.04%	12,268	28.50%	12,385	26.86%
	Control	1,324	30.82%	1,358	31.59%	1,592	30.09%	1,543	26.57%
Drink	Treatment	11,418	25.47%	12,002	25.95%	12,268	26.12%	12,385	25.28%
	Control	1,324	23.79%	1,357	25.57%	1,592	26.57%	1,543	24.11%
Commercial	Treatment	11,421	1.33%	12,034	1.75%	12,281	2.40%	12,401	2.71%
insurance	Control	1,324	3.02%	1,358	1.62%	1,592	1.44%	1,544	2.85%

8.3.2 Results from Staggered-DID

Table 8-6 summarises all staggered-DID regression results, outlining policy effects represented by marginal effects. The analysis indicates a significant positive impact of the integration policy on the average probability of outpatient visits, increasing it by 1.83% (p=0.01). Additionally, the integration led to significant reductions in OOP costs for outpatient services, with a decrease of 225 CNY (p=0.01), and also for hospitalisations, with a decrease of 1,228 CNY (p=0.02). However, no statistically significant effect was observed on the probability of hospitalisation (p=0.79) or the length of hospital stay (p=0.63). These findings provide crucial insights into the first research question of this chapter, indicating that the integration positively contributes to improved outpatient service accessibility and effectively alleviates the economic burden on patients for outpatient services and hospitalisations.

Table 8-6 Summary of staggered-DID estimation results

Outcome	Model	Policy Effect				
		M.E	S.E	P>|z|	95% CI	
Outpatient visit occurrence	probit	0.0183	0.01	**0.01**	0.004	0.032
Outpatient Out-of-Pocket (OOP) cost	glm(gamma)	-225	82	**0.01**	-386	-64
Hospitalisation occurrence	logit	0.0016	0.01	0.79	-0.010	0.013
Length of Stay (LOS)	negative-binomial	-0.3198	0.66	0.63	-1.605	0.965
Hospitalisation Out-of-Pocket (OOP) cost	glm(gamma)	-1,228	511	**0.02**	-2,230	-226

Note: M.E - Marginal Effect; S.E - Standard Error; CI - Confidence Interval

8.3.3 Results in Subgroups

Tables 8-7 to 8-11 summarise subgroup analyses for all outcomes. For the significantly affected outcome variables in Table 8-6, the policy effects did vary among different subgroups, while for the insignificantly affected outcome variables, the policy effects on them were also not significant in all subgroups.

Table 8-7 summarises subgroups results on outpatient visits. The rural treatment group exhibited greater benefits than the urban treatment group, with a significant increase of 2.05%

(p=0.022), whereas the increase observed in the urban treatment group was not statistically significant (p=0.182). Similarly, the high-income treatment group experienced better benefits compared to the low-income treatment group, with a significant increase of 2.94% (p=0.005), contrasting with the non-significant change in the low-income treatment group (p=0.469).

Table 8-7 Summary of subgroup results for outpatient visit occurrence

Subgroup	Observations	Policy Effect				
		Marginal Effect	Standard Error	P>|z|	95% Confidence Interval	
Urban	15,517	0.0167	0.0125	0.182	-0.0078	0.0412
Rural	36,161	0.0205	0.0089	**0.022**	0.0030	0.0380
High 50% income	26,304	0.0294	0.0106	**0.005**	0.0087	0.0501
Low 50% income	25,405	0.0076	0.0105	0.469	-0.0130	0.0282
Urban high income	7,823	0.0535	0.0182	**0.003**	0.0179	0.0892
Urban low income	7,630	-0.0239	0.0176	0.173	-0.0584	0.0105
Rural high income	18,401	0.0231	0.0130	0.077	-0.0024	0.0486
Rural low income	17,760	0.0190	0.0131	0.147	-0.0067	0.0446

Note: using probit model

Examining the interaction between region and income subgroups, the insignificant policy effect in urban areas can be attributed to opposing policy effects on the urban high-income and low-income subgroups. While the high-income urban treatment group benefited the most, with a 5.35% increase (p=0.003), the low-income urban treatment group experienced the only negative policy impact among all subgroups, although statistically insignificant. In rural areas, both high-income and low-income treatment groups were affected by the same direction of positive policy effects (although one was not significant and the other was only statistically significant at 10%), making the total policy effect in rural areas more significant than in urban areas.

Table 8-8 summarises policy effects on outpatient OOP costs in subgroups. The analysis indicates that, the rural treatment group derived greater benefits from the integration compared to the urban treatment group. However, when examining the policy effects on OOP costs between income subgroups, contrary effects emerge in comparison to outpatient visits. Specifically, the policy had a more significant effect in reducing OOP costs in the low-income

treatment groups compared to the high-income treatment groups. Further exploration within interaction subgroups reveals consistent cost reduction effects across all groups, but the size and significance of these effects vary. The urban low-income treatment group experienced the most significant impact, while the urban high-income treatment group was least affected.

Table 8-8 Summary of subgroup results for outpatient out-of-pocket costs

Subgroup	Observations	Policy Effect				
		Marginal Effect	Standard Error	P>|z|	95% Confidence Interval	
Urban	2,827	-152	134	0.255	-414	110
Rural	6,858	-277	91	**0.002**	-455	-100
High 50% income	4,969	-84	100	0.405	-280	113
Low 50% income	4,716	-243	101	**0.016**	-441	-45
Urban high income	1,452	-19	171	0.910	-354	316
Urban low income	1,338	-435	202	**0.032**	-833	-37
Rural high income	3,471	-219	112	0.051	-438	1
Rural low income	3,387	-137	113	0.225	-359	84

Note: using glm(gamma) model

Table 8-9 Summary of subgroup results for hospitalisation out-of-pocket costs

Subgroup	Observations	Policy Effect				
		Marginal Effect	Standard Error	P>|z|	95% Confidence Interval	
Urban	1,908	-322	748	0.667	-1,788	1,144
Rural	4,079	-1,685	612	**0.006**	-2,884	-486
High 50% income	2,920	-2,037	669	**0.002**	-3,350	-725
Low 50% income	3,067	-0	630	1.000	-1,235	1,235
Urban high income	913	-1121	957	0.242	-2,997	755
Urban low income	1,011	792	1,016	0.436	-1,201	2,785
Rural high income	2,023	-2,375	840	**0.005**	-4,020	-729
Rural low income	2.056	-550	710	0.439	-1,941	842

Note: using glm(gamma) model

Table 8-9 summarises the policy effects on hospitalisation OOP costs in subgroups. The rural treatment group and the high-income treatment group benefited more from the integration, aligning closely with the effects observed in outpatient visits. Specifically, the rural treatment group experienced a substantial decrease of 1,685 CNY ($p<0.01$), while the urban treatment

group's decrease was not statistically significant (p=0.667). The high-income treatment group showed a significant decrease of 2,037 CNY (p<0.01), whereas no decrease was observed for the low-income treatment group (p=1.000). Within the interaction subgroups, unlike the situation of the above two outcome variables, the high-income treatment subgroup in rural areas experienced the most significant policy effect on reducing hospitalisation OOP costs (2,375 CNY, p<0.01).

Tables 8-10 and 8-11 present the subgroup analysis results for two outcome variables, hospitalisation occurrence and LOS, which were not significantly influenced by the policy. The results indicate that the policy had no significant effect in any of the subgroups for the two outcome variables (all p-values were greater than 0.1). This suggests that the insignificance of the overall effect is indicative of the ineffectiveness of the policy, rather than the possibility of different subgroups offsetting each other with inconsistent or opposing effects.

Table 8-10 Summary of subgroup results for hospitalisation occurrence

Subgroup	Observations	Policy Effect				
		Marginal Effect	Standard Error	P>|z|	95% Confidence Interval	
Urban	15,591	0.0074	0.0107	0.49	-0.0136	0.0284
Rural	36,252	-0.0008	0.0072	0.91	-0.0148	0.0132
High 50% income	26,346	0.0048	0.0083	0.56	-0.0114	0.0210
Low 50% income	25,490	-0.0014	0.0090	0.87	-0.0190	0.0161

Note: using logit model

Table 8-11 Summary of subgroup results for length of stay

Subgroup	Observations	Policy Effect				
		Marginal Effect	Standard Error	P>|z|	95% Confidence Interval	
Urban	2,093	-0.5271	0.9298	0.57	-2.3495	1.2954
Rural	4,583	-0.4198	0.8280	0.61	-2.0427	1.2031
High 50% income	3,164	-0.4682	0.8767	0.59	-2.1864	1.2500
Low 50% income	3,512	-0.2052	0.8756	0.82	-1.9212	1.5109

Note: using negative-binomial model

In summary, the findings from this section provide insights into the second research question

of this chapter. The rural treatment group did experience greater benefits from the integration compared to the urban treatment group, aligning with policy expectations. However, the high-income treatment group appeared to derive greater benefits than the low-income treatment group, indicating a potential exacerbation of healthcare inequalities, particularly for those needing hospitalisation. The interactions between urban-rural areas and high-low income have complex effects on policy heterogeneity, but income appears to play a more dominant role compared to urban-rural areas.

8.4 Robustness Tests and Supplementary Results

This section sequentially reports the remaining two parts of the five-part analysis: robustness test results and DDD regression results. Robustness results can show whether the data meet the assumption requirements of the DID analysis, while DDD results are used to verify whether the results are still reliable when the assumptions are not met.

8.4.1 Results of Robustness Tests

This subsection presents the results from the pre-integration parallel trend tests, as depicted in Figures 8-3 and 8-4. Each plot point represents estimated coefficients representing the change in the difference between the treatment and control groups in a specific year relative to the control year, which is set as one year prior to the implementation of integration. The vertical lines associated with the intervals indicate the estimated 95% confidence intervals. If all the 95% CI intersect the dashed line denoting zero value, it signifies that the difference between the treatment and control groups remains consistent over time, thereby satisfying the pre-policy parallel trend assumption.

The results from Figure 8-3 indicate that all 95% CIs of estimated coefficients for the outpatient OOP cost data cross the zero-value line. This suggests that the treatment group and control group adhere to the pre-integration parallel trend. However, for the outpatient visit data, the estimated coefficient in 5 years ago (-5) is significantly lower than zero, indicating non-compliance with the pre-integration parallel trend assumption.

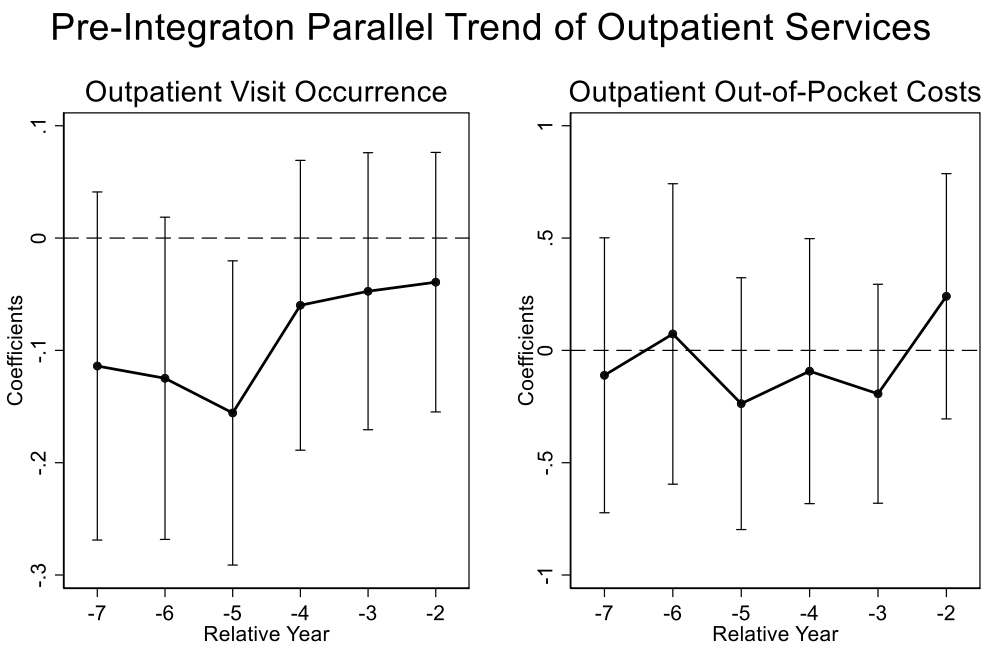

Figure 8-3 Pre-integration parallel trend of outpatient services

A similar pattern emerges from Figure 8-4. While the treatment group and control group display a parallel trend for hospitalisation OOP costs and LOS, they fail to satisfy the pre-integration parallel trend assumption for the occurrence of hospitalisation because the coefficient in 3 years ago (-3) is significantly lower than zero.

These results suggest that DID results for outpatient visits and hospitalisation may contain biases and require further DDD estimates for them.

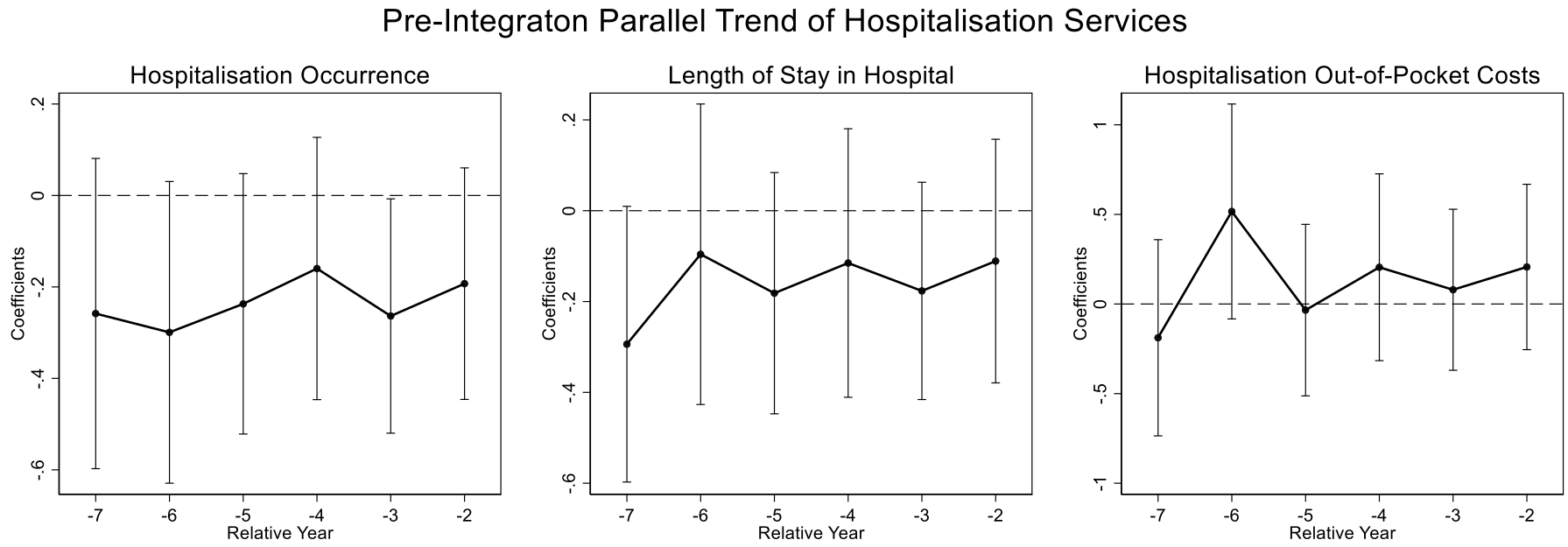

Figure 8-4 Pre-integration parallel trend of hospitalisation services

Figures 8-5 and 8-6 show the distribution of coefficients and p-values obtained from placebo (fake) policy implementation, randomly sampled over 500 iterations. The red vertical solid line marks the actual policy effect estimates from subsection 8.3.2 for comparison. The purpose of the placebo test is to demonstrate the validity of the DID estimation by verifying the absence of any policy effect when none should exist. Thus, if the coefficient distribution of the placebo policy is centred around 0 and the p-value distribution is centred around a high value (great than 0.1), it indicates that the current model settings do not produce spurious policy effects. And if an actual estimated policy effect also lies outside the sampling distribution, it provides confirmation that this estimated effect is indeed attributable to the evaluated policy, and not caused by other factors such as model setting errors or unidentified policies.

In Figure 8-5, the distributions of the effect coefficient of the placebo policy on outpatient visits and OOP costs are very close to 0, with p-value distributions around 0.98 and 0.95. These indicate that the current analysis model does not produce false policy impact. And red lines outside the distribution confirm that the previous DID estimated policy effects in subsection 8.3.2 can be attributed to integrated insurance. A similar explanation holds for the hospitalisation OOP costs results in Figure 8-6. As for the hospitalisation occurrence and LOS, since the red line is within the distribution, it cannot be ruled out that these two estimators are caused by the random error of the model, and it is not supported that the policy has a significant effect. This is consistent with the conclusion of subsection 8.3.2.

8.4.2 Results from Staggered-DDD

Tables 8-12 summarise the DDD estimation results for two outcome variables that do not exhibit parallel trends. The probability of outpatient visits increased by 2.89%, exceeding the estimated 1.82% in DID, with a consistent direction of effect but reduced significance to 10% (p=0.077). And the policy effect estimate for hospitalisation occurrence remains insignificant (p=0.74).

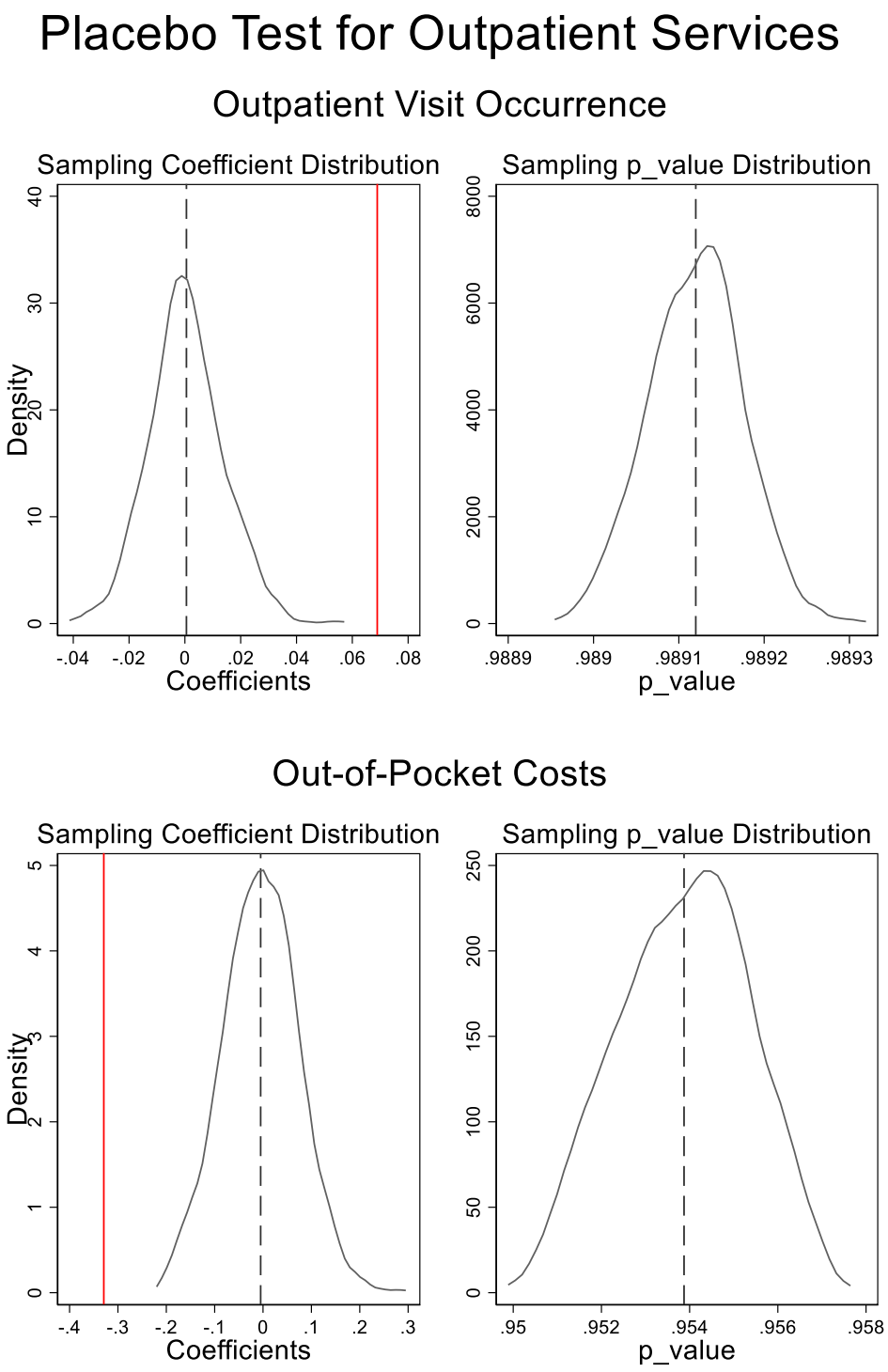

Figure 8-5 Placebo test for outpatient services
(Note: Vertical dashed line—estimated average placebo policy effect from 500 samples; solid red line—estimated true policy effect from subsection 8.3.2)

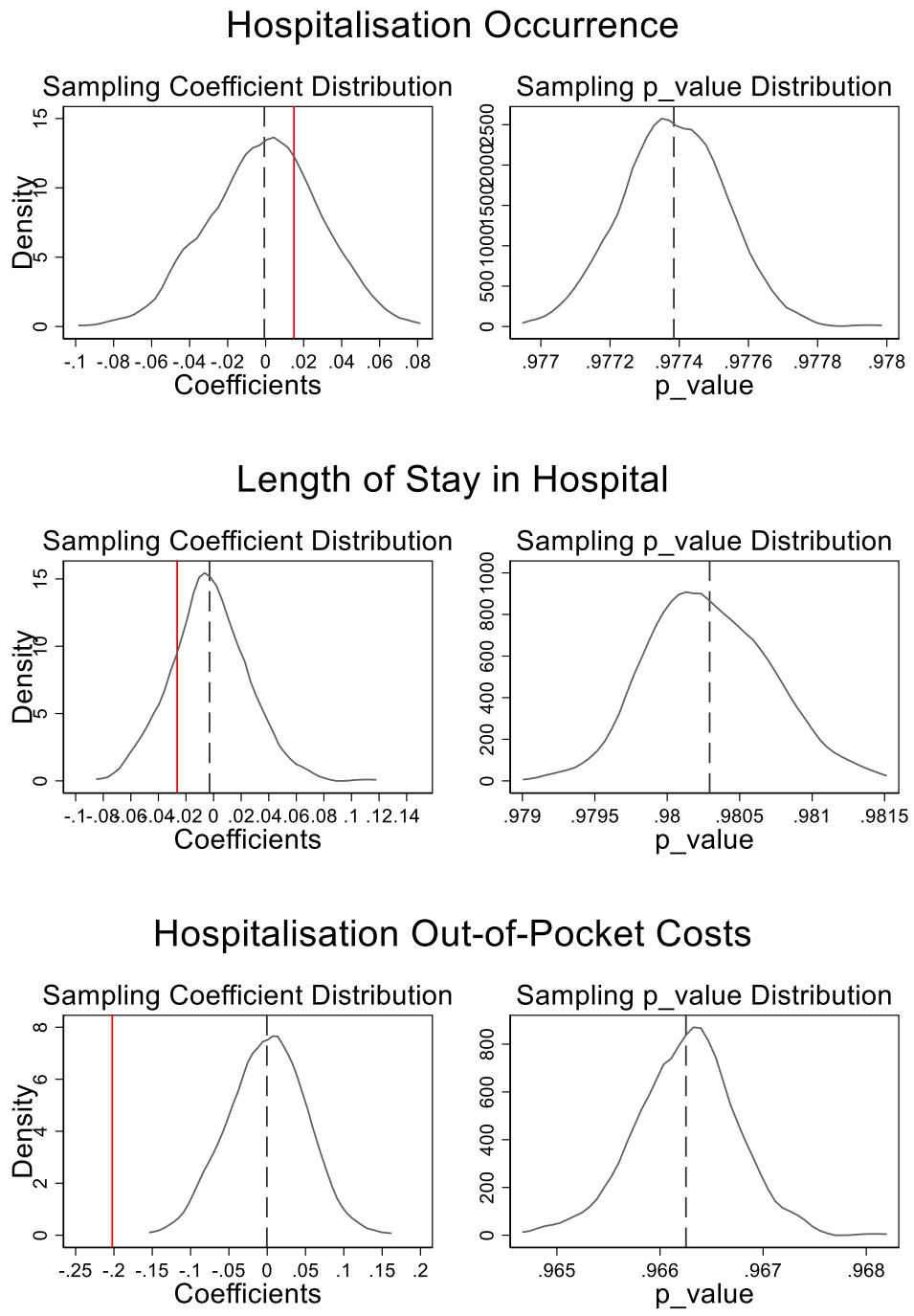

Figure 8-6 Placebo test for hospitalisation services
(Note: Vertical dashed line—estimated average placebo policy effect from 500 samples; solid red line—estimated true policy effect from subsection 8.3.2)

Table 8-12 Summary of staggered-DDD estimation results

Outcome	Model	Policy Effect				
		Marginal Effect	Standard Error	P>|z|	95% Confidence Interval	
Outpatient visit occurrence	probit	0.0289	0.02	0.077	-0.0032	0.0610
Hospitalisation occurrence	logit	-0.0044	0.01	0.74	-0.0301	0.0212

8.5 Discussion and Conclusion

This chapter investigates the impact of integrating basic medical insurance for urban and rural residents on health service use and out-of-pocket costs, using CHARLS panel data spanning 2011 to 2018. Employing the multi-period staggered Difference-in-Differences (DID) method as the primary analytical approach, the study addresses potential estimation bias caused by unobservable confounding factors. The reliability of the DID estimates is verified through pre-integration parallel trend tests and placebo tests. Moreover, the Difference-in-Difference-in-Differences (DDD) model is used to estimate policy effects for outcome variables that may not conform to parallel trends.

The findings suggest that the integration policy has increased outpatient visits and reduced out-of-pocket costs for both outpatient and hospitalisation among the treatment group, that is, residents in areas where integration was implemented. However, no significant impact on the hospitalisation service use was observed. This aligns with the conclusion drawn in Chapter 7. It was mentioned that since changes in outpatient reimbursement policies across cities were more similar, it is speculated that the impact on outpatient services at the national level might be more significant. Conversely, variations of changes in hospitalisation reimbursement policies among different cities might result in limited or insignificant policy effects on hospitalisation services at the national level.

Furthermore, this chapter also examines the differential effects of integration policies on urban-rural and income subgroups. For the significantly affected outcome variables, outpatient visits, outpatient OOP costs, and hospitalisation OOP costs, the treatment group in rural areas, that is,

the rural residents have experienced great benefits from the policy compared to urban residents. This indicates a reduction in urban-rural inequality in healthcare, aligning with the expected outcome of insurance integration. However, higher-income residents have experienced greater benefits from the policy compared to lower-income residents, in terms of outpatient visits and hospitalisation OOP costs. This means that economic-related healthcare inequalities have been exacerbated, which may not be expected from the policy reform. Since the data used in the analysis is a mixture of people from different regions of China, one possible reason is that the differences in policy effects between high and low-income groups may partly reflect policy differences caused by differences in economic development in different regions of China. Chapter 3 discussed the huge economic differences among different provinces and regions in China, and Chapter 7 also showed that the more economically developed cities, Beijing and Shanghai, have much more government subsidies for insurance than Chengdu. This point will be discussed further in the final chapter. This shows a potential trend that the local economy is better, the government invests more in insurance, and residents benefit more from insurance, which further expands the inequality in healthcare between them and residents in relatively backward areas. Therefore, further research on the effects of integrated policies across cities will help to understand the sources of these inequalities. And increasing healthcare investment and subsidies in economically backward areas may be a good measure to reduce economically related healthcare inequalities.

This study offers several strengths compared to previous research. Firstly, it used a broader dataset covering more integration regions and a longer time period, resulting in more representative findings. In addition, the model settings used in this study are more reasonable and detailed. The use of staggered DID and the control of urban fixed effects further enhance the precision of the estimates. Importantly, this study conducted pre-implementation parallel trend tests and placebo tests, which are important to support the reliability of the DID estimated results but lacking in previous studies.

However, certain limitations should be acknowledged. Firstly, the staggered DID model employed in this study has been criticised for its potential drawbacks in the latest

methodological studies. For example, it provides a weighted estimate of a single policy effect, but this value may contain negative weightings and thus it may not reflect the true policy effect (Goodman-Bacon, 2021). Another issue is about DDD estimation. The DDD estimates for outpatient visits have the same direction and a larger size policy effect than the DID estimates, which further supports that the integration promotes the use of outpatient services. However, the statistical significance of these results was diminished. This means that the current research results are still not entirely solid. Thus, better measures may need to be explored in future research to mitigate the impact of confounding factors and obtain more robust policy effect estimates.

Chapter 9 City Level Quantitative Analysis of the Impact of Integration on Health Expenses and Health Service Use: The Case of Patients with Ischemic Heart Disease

9.1 Introduction

Chapter 8 investigated the national-scale average policy effects of integration in China using the CHARLS database. The findings suggest that the integration promoted outpatient visits and reduced out-of-pocket costs for both outpatient and hospitalisation. However, it is notable that the integration operated at the city level. The results from policy comparisons in Chapter 7 have also indicated variations in policies across cities, suggesting that the effects of policies may also differ across cities. As a result, it is possible that the results in Chapter 8 are mixed with opposing city-level policy effects. Therefore, conducting research that focuses on city-level policy effects will enhance the understanding of the integration.

This chapter aims to fill research gaps regarding the lack of comparative research on the impact of various policies at the city level and the mechanisms involved after nationwide integration implementation. Specifically, this chapter explores whether integration effects on outpatient total fees, hospitalisation total fees, and length of stay are similar or various across cities during nationwide implementation. Additionally, this chapter also examines how policy effects change over time after the initial year of implementation (dynamic effects), and whether patients with varying total fees are affected differently by the integration (quantile effects). These inquiries align with the fourth and fifth objectives of this thesis and contribute to a more comprehensive understanding of the integration than in Chapter 8.

This chapter is structured as follows. Section 9.2 describes the methodology in detail. It includes details of the dataset, individual and variable selection decisions, causal identification

strategy, and statistical analysis framework. Section 9.3 shows different results for six selected cities that implemented the integration in 2017 and 2018. Section 9.4 provides more in-depth analysis results about dynamic effects and quantile effects. Lastly, Section 9.5 concludes with a summary and a brief discussion of this chapter.

9.2 Methodology

This section consists of five parts. Firstly, the source and structure of the database are described, and its strengths and weaknesses are discussed. And then, the criteria and process for patient selection are detailed, followed by the variable selection. Detailed statistical methods and model specifications are subsequently explained. Finally, the statistical analysis framework and steps of this chapter are given.

9.2.1 Data Sources

According to the data provider requirement, this database must be used anonymously for publication so that this study calls it "the third-party database". The third-party database is a reconstructed Electronic Medical Records (EMR) database[50], which means it is repeated cross-sectional data. By the end of 2018, the database covered 16 of 34 provincial-level administrative regions in China and contained more than 40 million anonymised visit records in 87 hospitals. The dataset covers various information: demographics and insurance, hospital, diagnostic, lab results, prescription, hospitalisation details, healthcare resource utilisation, and cost.

There are four key advantages of this database. First, it has a very large sample size. Therefore, models will have high statistical power. Second, the EMR data are real-world, so there is no recall bias and less prone to errors compared to survey (self-report) data (Sulieman *et al.*, 2022). Third, it involves detailed information, such as diagnostics, drugs, healthcare resources, and

[50] The raw EMR database is not allowed to be used directly because it contains patient information unless there is an ethical approval. The raw EMR database is reconstructed, such as de-identified and re-encoded, and then can be used for research without ethical approval. Such databases are prevalent in disease burden studies in China.

cost details. Fourthly, its population includes all age's patients so that it can provide a more comprehensive analysis.

This dataset also exhibits limitations. Firstly, there is a lack of consistency in the electronic medical record (EMR) systems used across different hospitals, which poses a risk when pooling data from different hospitals. This is a common flaw in all EMR data. Sulieman *et al.* (2022) highlight the lack of interoperability of electronic records across healthcare facilities in the United States. The situation in China is similar, until 2022, one national unified EMR system is still under planning discussion and has not been implemented[51]. Secondly, data access is based on the hospitals' willingness rather than being designed. Although the database covers 40 cities in 16 provinces, its representativeness in relation to the entire China is unknown. Thirdly, the database is hospital-based rather than patient-based, which leads to two issues. On one hand, when one patient seeks care at different hospitals, he/she is identified as two individuals due to the independent information systems used by each hospital. On the other hand, not all visits of one patient are included, because not all hospitals are included in the database[52]. These make the full and long-term treatments received by the patients unknown. Since the focus of this chapter is on analysing policy effects at the city level rather than the national level and examining single visits of patients rather than long-term follow-up treatment, the three identified limitations of the database are unlikely to undermine the validity of the findings in this particular context.

However, fourthly, the database lacks socioeconomic information, and therefore, only clinical information can be used for control purposes. The absence of control for economic conditions, particularly income status, may impact the identification of policy effects using the Difference-in-Differences (DID) and Propensity Score Matching (PSM) approaches, as discussed in Chapter 5. This represents a major limitation of this study.

[51] health.people.com.cn/n1/2022/0216/c14739-32353419.html

[52] However, this might not be a big problem since most people tend to go for appointments at the same hospital every time.

9.2.2 Data and Patient Selection

In this chapter, the dataset preparation processes involve two main selection steps. Firstly, the analysis dataset is filtered from the original database based on the analysis objectives. Secondly, appropriate individuals are selected based on the causal identification approaches used. Therefore, this subsection first discusses how to select target cities, hospitals, and analysis datasets from the raw data. And then, the possibility of different patient populations in the hospital as control and treatment groups are discussed. Finally, the methods used to test the rationality of the selected control group are explained.

Selection of diseases, cities and hospitals

This chapter exclusively focuses on Ischemic Heart Disease (IHD), a common chronic non-communicable disease (NCD) in China. NCDs, consisting of cardiovascular and cerebrovascular diseases (CVD), diabetes, cancer, and Chronic Obstructive Pulmonary Disease (COPD), contribute to over 60% of global deaths, estimated to rise to 75% by 2030 (Qin, 2014; Cao *et al.*, 2020). CVD is the leading cause of death globally, with IHD and stroke accounting for 84.9% of all CVD deaths (GBD, 2018).

IHD is characterised by ischemic damage to the myocardium caused by an imbalance between the blood flow in the coronary arteries and the demand on the myocardium, previously known as coronary heart disease (Zhai and Liu, 2016). It can result from the deposition of serum lipid substances on the surface of coronary arteries, leading to the narrowing or obstruction of the lumen of blood vessels, or functional changes in coronary arteries that lead to myocardial ischemia, hypoxia, or even necrosis. The IHD prevalence in China is expected to reach 189.59 per million people by 2025 (Cao *et al.*, 2020). In 2017, Years of life lost (YLL) due to premature death from IHD ranked first, marking a notable increase from the 10th position in 2007 (GBD, 2018). The annual direct medical expenses caused by IHD in China surged from 209 billion CNY in 2003 to 282 billion CNY in 2013 (Yang *et al.*, 2008; Luan, Hu and Li, 2018).

Focusing on ischemic heart disease (IHD) provides several advantages for this study. Firstly, unlike common minor ailments such as colds, IHD significantly increases medical costs and

services. This makes the policy effects more discernible if the policy does have impacts. Secondly, IHD is a very common NCD, representing a larger population compared to diseases with lower incidences. Finally, IHD is a chronic disease that patients can survive for a long time after timely intervention, ensuring in a relatively stable patient population across the years. This meets an assumption as discussed in Chapter 5 that there are no systematic differences among samples across years when using the DID approach in repeated cross-sectional data. (Luan, Hu and Li, 2018; Zhao *et al.*, 2019; Cao *et al.*, 2020; Ren *et al.*, 2020; Yu *et al.*, 2021)

Based on the analysis objectives, six cities that implemented integration in 2017 and 2018 are selected from the database (Table 9-1). They are anonymised as City 1 to City 6. Five of these cities are in the western region of China, while one is in the central region. Three cities implemented integration in 2017, while the other three implemented in 2018. Each selected city has only one hospital data, and these hospitals are labelled as Hospitals A to Hospitals F. The selected hospitals include three Tertiary-Level and three Secondary-Level. Among the Tertiary-Level hospitals, two are Grade A, and one is Grade B, while all the Secondary-Level hospitals are Grade A[53]. Notably, variations exist in the years covered by outpatient and hospitalisation data across these hospitals.

Table 9-1 Hospital and dataset information

City	Location	Year of integration	Hospital			
			Label	Level	Years of outpatient data	Years of hospitalisation data
1	West	2017	A	T-A	2016-2019	2013-2019
2	West	2018	B	S-A	2016-2019	2014-2019
3	West	2018	C	T-A	-	2016-2019
4	West	2017	D	T-B	-	2016-2017
5	West	2018	E	S-A	2016-2018	-
6	Middle	2018	F	S-A	2014-2018	-

Note: T-Tertiary; S-Secondary

[53] For Details of hospital levels see section 4.2.1, footnote 19.

Eligibility of four different patient groups

As discussed in Chapter 5, it is impossible in the EMR database to find an appropriate instrumental variable, which is correlated with the integration treatment and only affects the outcome through the treatment. Thus, this chapter primarily uses the DID approach for analysis. The DID requires distinguishing treatment and control groups among all patients. In a hospital, all patients can be divided into four groups according to their insurance status to discuss their eligibility as treatment and control groups: resident insurance, employee insurance, government free medical care, and self-paid (full out-of-pocket).

The treatment group is patients with resident insurance, which is the target group for insurance integration policies, so the control group is selected from the remaining patients. The DID requires parallel outcome trends in the control and treatment groups in the absence of treatment, but this cannot be directly observed in reality because it is a concept based on a counterfactual framework, as described in Chapter 5. However, the parallel trend assumption can be indirectly supported by individual characteristics. The theoretical reviews in Chapter 4 indicate that the outcomes, health service use and fees, are affected by various factors, which can be grouped into three categories: 1) macro factors, such as changes in medical pricing due to economic growth or inflation; 2) individual socioeconomic factors, such as income; and 3) individual physiological factors, such as age, sex, and health status. If a group of patents and the treatment group exhibit similar annual trends in these three categories of factors, then this patient group is likely to have parallel trends in health service use and fees with the treatment group and can be used as an appropriate control group. Given that the analysis in this chapter is conducted hospital-by-hospital, macro factors remain the same for all patients within a hospital. Therefore, the following discusses individual economic and physiological factors, for the remaining three patient groups, and whether the outcomes of them have parallel trends with the resident patient group.

The first to be discussed is the group of patients with government free medical care. Due to the free of charges, the use and fees of health services for patients with government free medical care are entirely independent of income and only related to their health status. In contrast, the outcomes for the treatment group, patients with resident insurance, are influenced by both wealth/income and

health status. This fundamental difference makes it nearly impossible for the two patient groups to exhibit parallel trends in outcomes. Consequently, patients with government free medical care are deemed ineligible as a control group.

The second to be considered is the group of patients with employee insurance. Observing the health status trends between non-employed residents and employees directly through statistical indicators is challenging due to the absence of health-related statistics based on employment status. However, the abundant health-related indicators in the EMR data offer a way for controlling health-related trends. This can be achieved through regression with control variables or matching methods to mitigate its impact on the parallel trend of outcomes and the policy effect estimated by DID.

For the income trend, since there are no direct statistics on the income data based on employment status, I use wage income in national statistical data to represent employee income and use operating income and transfer income to represent unemployed residents' income. Operating income refers to the income of self-employed individuals, such as farmers and small shopkeepers, who are important parts of the population covered by residential insurance. Transfer income includes government subsidies and pensions, etc., involving the poor, disabled, and retired population. They are also part of the resident insurance population. I conducted a preliminary check of China's wage data trends from 2013 to 2019[54]. Wage income increased from 10,411 CNY to 17,186 CNY, with an annual growth rate of 8.7%. Similarly, operating income and transfer income increased from 6,477 CNY to 10,927 CNY, with an annual growth rate of 9.1%. These two income groups exhibited comparable growth rates, and the proportion of operating and transfer income to wage income remained stable (ranging from 62.7% to 63.5%, correlation coefficient: 0.9999). This suggests parallel income trends between unemployed residents and employees, indicating that the income-related portion of the outcome variable will also exhibit parallel trends.

Moreover, the use of employee insurance patients as the control group of residents' insurance (or vice versa) in previous health studies using DID conducted in China further supports the parallel

[54] Data source: https://data.stats.gov.cn/english/

trends in health-related outcome variables. For example, He and Nolen (2019) examined the impact of employee insurance implementation on health status and health service use using the CHARLS data. They used employees as the employees as the treatment group and unemployed residents as the control group. Chen, Xu and Gao (2020) analysed the effects of resident medical insurance implementation on health institution visits and costs using data from the China Health and Nutrition Survey. They used employees of state-owned enterprises as a control group in DID analysis because this population was completely unaffected by the reform. Yu *et al.* (2021) investigated the influence of introducing Catastrophic Medical Insurance (CMI) into resident medical insurance in an eastern Chinese city using insurance claims data. They discussed that since there were significant structural changes in resident insurance but only marginal (annual) adjustments in employee insurance, patients with employee insurance were suitable as a control group in the DID analysis.

These three studies show that the two groups satisfy the parallel trend assumption by: 1) confirming that outcomes have similar trends before the policy is implemented; 2) controlling for individual control variables. These methods are the most common ways to ensure valid DID estimates in empirical studies and have also been discussed in the methodological review in Chapter 5. In addition, the three studies also highlight a potential advantage of comparing unemployed residents and employees in that these two populations are completely unaffected by policies aimed at one another. This can strictly meet an assumption of the DID method, that is "no spillover". This assumption requires that the policy does not affect the outcomes of individuals who should not be affected by the policy[55].

Combining the discussion of previous studies with my own preliminary checks based on national statistics, employees' insurance patients can be considered an appropriate control group for patients with resident insurance when applying the DID method.

The last to be discussed is the group of self-paid patients. This patient group includes: 1) non-

[55]https://www.publichealth.columbia.edu/research/population-health-methods/difference-difference-estimation

employed residents who choose not to enrol in resident insurance (which is a very small portion as the insurance coverage rate in China has exceeded 95% as mentioned in Chapter 2); 2) patients who have resident insurance but do not use it for this time visit. The reason is that the resident insurance enrolment is not mandatory and requires the insured to register their insurance with a hospital in person (that is related to the "designated" institution mentioned in Chapter 7)[56]. Therefore, in outpatient, especially emergency situations, there will be cases where patients have resident insurance but do not use it. In both scenarios, these self-paid patients can be regarded as having characteristics that are not significantly different from the patients with resident insurance at the population level. Hence, they satisfy the parallel trend assumption, making them suitable control groups for analysis.

Parallel trend between treatment and control group in this study

To ensure consistency and comparability of analysis results across different cities, a mixture of self-paid patients and employee insurance patients was chosen as the control group. The decision was based on the substantial variation in the proportion of self-paid patients and employee insurance patients observed in different cities throughout China. For example, in Jiayuguan City, employee insurance accounted for 38.53% of the population in 2018, while Zhangye City accounted for only 10.54% of the population in the same year[57,58]. This is also reflected in the EMR data. By using a mixture control group, fluctuations in sample size between the treatment and control groups were minimised, facilitating the attainment of optimal comparability across various cities.

---

[56] On the contrary, due to the mandatory registration of employee insurance, employees can use their insurance directly without binding with the hospital (it can be used in almost all health institutions). There are slight differences in the visit process for patients with the two types of insurance. For employee insurance patients, the hospital's patient files are established based on the employee's insurance card directly. For residents' insurance patients, the hospital firstly establishes patient files (self-paid/uninsured). Subsequently, the patient is required to annually bind his/her patient files in this hospital with resident insurance. It is important to note that patient files are not linked across different hospitals. A guide on resident insurance reimbursement can be found (in Chinese):
https://www.nanning.gov.cn/ggfw/bmfw/shbx/ylbx/t4922001.html

[57] https://www.cnstats.org/tjgb/201904/jygsjygs-2018-mbl.html

[58] https://www.zhangye.gov.cn/tjj/ztzl/tjsj/201904/t20190401_198309.html

To further verify the parallel trends between the mixed patient control group and the treatment group, I employed two approaches. Firstly, I conducted cross-validation using the CHARLS database. Since it is impossible to directly observe parallel trends in cities that have implemented integration in the EMR data, I selected cities from the CHARLS database that had not implemented integration by 2018 and shared similar macro characteristics (such as GDP per capita) with the cities included in the EHR data. Subsequently, I compared the trends in health fees between the treatment group (residents) and the control group (mixed) in these selected CHARLS cities. The results revealed that the treatment and control groups exhibited parallel trends in health fees in these selected CHARLS cities. Specifically, the annual change in the difference of outcome between the two groups was not statistically significant (For detailed information, refer to Appendix H). This cross-validation analysis strengthens the likelihood of parallel trends in medical expenses between the treatment and control groups in the EHR data. Additionally, I also conducted a pre-integration trend test on the EHR data, which will be comprehensively explained in the subsequent subsection on Causal Identification Strategies.

The full inclusion criteria are as follows:

1) Patient with a diagnosis of ischemic heart disease (Chinese in parentheses): ischemic heart disease (缺血性心脏病/心肌病/冠心病), myocardial infarction (心肌梗死/埂塞), acute coronary syndrome (急性冠脉综合征), coronary arteriosclerosis (冠状动脉硬化); or patient diagnosis ICD code contains: I21/I22/I24/I25.
2) Cities' residential insurance integration takes place after 2016;
3) Have both resident and non-resident visit data (Chinese in parentheses): resident (居民), rural (农), self-paid (自费), and employee (职).

The exclusion criteria are as follows:

1) The patient is younger than 18 years old;
2) Patient's clinical information is missing;
3) Patients with free medical care (公费医疗).

Finally, 225,294 records were included, of which 57,447 were outpatient visits records and 167,847 were hospitalisation records (Table 9-2). The sample sizes for hospitalisation records were relatively balanced between the treatment group and the control group. However, for outpatient visits records, the control group had larger sample sizes than the treatment group. These gaps may be attributed to a considerable number of patients opting to pay for their outpatient visits out-of-pocket. Among the selected hospitals, two hospitals provided both outpatient and hospitalisation records, while two hospitals exclusively had hospitalisation records, and the remaining two hospitals solely possessed outpatient records[59].

Table 9-2 Sample size

Location	City	Year of integration	Hospital		Hospitalisation record		Outpatient visits record	
			Label	Level	Control	Treatment	Control	Treatment
West	1	2017	A	T-A	10,397	11,251	16,798	1,223
West	2	2018	B	S-A	14,267	12,539	50,474	3,107
West	3	2018	C	T-A	2,551	2,855	-	-
West	4	2017	D	T-B	2,099	1,488	-	-
West	5	2018	E	S-A	-	-	39,004	1,969
Middle	6	2018	F	S-A	-	-	47,478	7,794
		Total			29,314	28,133	153,754	14,093

Note: T-Tertiary; S-Secondary

9.2.3 Variables Selection

Variables to be selected include outcome (dependent) variables and control (independent) variables, and the selection process is described separately.

Outcome (dependent) variables

The analysis in this chapter uses three outcome variables: outpatient total fees, hospitalisation total fees, and length of stay in hospital (LOS). It's important to note that the electronic medical record (EMR) system is an internal hospital system and does not involve the reimbursement process. As a result, the fee data captured in the EMR represents the actual fees incurred within

[59] This is because different hospitals have different cooperation methods with the data provider: some hospitals require to manage all data, while others only require to manage outpatient or hospitalisation data.

the hospital and does not take into account the reimbursement process. The reimbursement process is managed by the medical insurance system, which connects the hospital and the medical insurance administration. Consequently, the outpatient total fees and hospitalisation total fees variables in this study encompass the sum of all itemised fees in a single visit on the bill, including consultation fees, examination fees, treatment fees, drug fees, consumables fees, and other related fees. It does not refer to the patient's out-of-pocket payment after reimbursement. LOS is calculated as the date of discharge minus the date of admission plus one.

Control variables

The analysis used electronic medical record (EMR) data, which limited the availability of control variables to demographic information and diagnostic data. Demographic variables included sex and age at the visit, providing basic patient characteristics. Diagnostic variables were used to control the individual's health status and long-term fee trends. Three categories of relevant variables were included: IHD subtypes, specific chronic diseases, and comorbidities. The IHD subtype included four variables, namely I20 angina pectoris, I21/I22 myocardial infarction, I24 acute coronary syndrome, and I25 ischemic heart disease. To control for long-term fee trends, six specific chronic disease variables were used, including cancer, heart failure, hypertension, cerebral infarction, chronic obstructive pulmonary disease, and diabetes. These diseases were highlighted chronic diseases in China by the "Report on Nutrition and Chronic Disease Status of Chinese Residents" (National Health and Family Planning Commission, 2015). The comorbidity variables indicated the presence of other comorbidities that could impact the fees in this visit, and a total of 23 complications were classified based on the International Classification of Diseases (ICD). It should be noted that some control variables may exhibit multicollinearity, which refers to high linear correlations among them. Although multicollinearity can affect the standard error and significance of the variable coefficients, it might not be a problem in this study because control variables' coefficients are not the focus of the analysis. Additionally, the STATA software automatically eliminated perfectly collinear control variables from the regression model. This process was not performed manually due to differences in collinear variables in the different hospital data.

9.2.4 Causal Identification Strategy

This subsection provides a comprehensive explanation of all the analysis methods employed in this chapter to evaluate policy effects. It consists of two parts: the main analysis and the further analysis. The main analysis involved the DID model, the parallel trend test before integration, the use of placebo test data PSM, and secondary DID analysis. In the further analyses part, the data will be analysed for dynamic effects and quantile effects, as mentioned in Chapter 5.

As discussed in Chapter 5, DID is the primary method in this chapter. Based on the equation [4] in subsection 5.1.3, the multi-period DID regression model used in this chapter is as follows:

$$\boldsymbol{y_{it} = \alpha + \delta resident_i + \beta treated_{it} + \theta' X_{it} + \mu_t + \varepsilon_{it}}$$

Where the $\boldsymbol{resident_i}$ captures the inherent differences between the treatment (=1) and control (=0) groups; the $\boldsymbol{\mu_t}$ is the "time fixed effect", that is, a series of year dummy variables, which controls for time-dependent unobservables, such as inflation; $\boldsymbol{treated_{it}}$ is a dummy variable that takes the value of 1 if the integration policy has been implemented for individual i at time t, and takes the value of 0 in other cases; $\boldsymbol{X_{it}}$ refers to a series of control variables; $\boldsymbol{\varepsilon_{it}}$ is the error term. The coefficient of $\boldsymbol{treated_{it}}$ represents the policy effect.

After the main DID analysis, robustness tests are performed. These include the pre-integration parallel trend test and the placebo test as discussed in subsection 5.1.3. Since in a city, individuals in all treatment groups receive treatment at the same time point, the pre-integration parallel trend test is to test whether there are significant changes in the differences between the treatment group and the control group in different natural years before the implementation of the integration policy. Based on the equation [7] in subsection 5.1.3, the regression of the parallel trend test for multi-period DID used in this chapter is as follows:

$$\boldsymbol{y_{it} = \alpha + \delta resident_i + \sum \beta_t resident_i \times year_t + \theta' X_{it} + \mu_t + \varepsilon}$$

A series of interactions between grouping variable and years are introduced. Suppose a total of T years, the number of t is T-1. The coefficient for each interaction term compares the group differences between two specific years. Thus, it requires data for at least two years before implementation. To make the results more intuitive/interpretable, this study takes the year

before the policy implementation as the control year (baseline). Therefore, if all coefficients of the interaction terms before implementation are not significant, it means that the parallel trend assumption is likely satisfied. Taking 3-year pre-policy data as an example (the policy is implemented in the year 4), the difference between two groups in year 3 compared with the difference between two groups in years 1 and 2, respectively, so there are 2 interaction terms. If both coefficients of these two interaction terms are not significant, then there is no significant change in the group differences before the policy implementation, that is, the parallel trend assumption is likely satisfied. The coefficient graph of interaction terms can directly show whether the assumption is satisfied or not.

Referring to La Ferrara, Chong and Duryea (2012) and Li, Lu and Wang (2016), the purpose of the placebo test is to exclude the influence of other policies. Its idea is to falsify the treatment group to estimate, that is, to repeat multiple (usually 500 times) random selections of individuals as the treatment group, and check whether the coefficient of the "pseudo-policy dummy variable" is significant. If the coefficient of the pseudo-policy is significant and the original DID estimator is within the distribution of the sampling estimator, it means that the other factors has affected the outcome variable; if the sampling coefficient is not significant and the original estimator is outside the distribution of the sampling estimator, it means that evaluated policy did have an impact.

In cases where the pre-integration parallel trend assumption is not met, the propensity score matching (PSM) method can be utilized, as discussed in subsection 5.1.3, to select individuals with similar characteristics from the original treatment and control groups. The use of PSM is to create new treatment and control groups that exhibit no significant statistical differences in their individual characteristics. The new treatment and control groups are more likely to exhibit parallel trends compared to the original treatment and control groups. Since the analysis uses multi-year data, the PSM is performed year-by-year[60]. The large sample size allows for 1:1

[60] Similar processes can refer to the literature: Heyman F, Sjoholm F, Tingvall P G. Is There Really a Foreign Ownership Wage Premium? Evidence from Matched Employer-Employee Data[J]. Journal of

nearest-neighbour matching, and the propensity score is computed using the logit model. If initial matching results still have significant differences between treatment and control groups, alternative matching methods, such as kernel matching mentioned in Chapter 5, are also tried. If it is not possible to identify a treatment and control group without significant differences after using multiple matching methods, this set of data is excluded from the analysis. It is because DID estimates based on this set of data will contain inevitable bias due to non-parallel trends, rendering the estimates unreliable. Finally, DID analysis is performed again for matched treatment and control groups. Subsequent in-depth analyses will also be conducted using the matched groups, if applicable.

After the main DID analysis and robustness test, there are two further analyses to explore the mechanism of integration: dynamic effects and quantile effects. They will be explained one by one next.

The first further analysis is to consider the time variation in policy effects, known as "dynamic policy effects". In this study, the annual variation is considered. The equation used is as follows:

$$\boldsymbol{y_{it} = \alpha + \delta resident_i + \beta treated_{it} + \sum \gamma_t treated_{it} \times lag_t + \theta' X_{it} + \mu_t + \varepsilon_{it}}$$

where the $\boldsymbol{lag_t}$ are a series of dummy variables indicating which year after the implementation, so the interaction term $\boldsymbol{treated_{it} \times lag_t}$ indicates the effect difference between the two years. It requires data for at least two years after implementation. To make the results easier to interpret, this study set to compare the first implementation year with each subsequent year. For example, if there are data available for two years after implementation, then the coefficient $\boldsymbol{\beta}$ reflects the policy effect in the first year of implementation; the coefficient $\boldsymbol{\gamma_t}$ reflects the difference in effect policy between the second and first year, and if the coefficient is significant, there is a delay in the policy taking effect or overflow effect.

The second further analysis is the quantile effect analysis. All of the above regression models assume that the effect is the same across all patients. Given the existence of deductibles and

International Economics, 2007, 73(02): 355-376; Bockerman P, Ilmakunnas P. Unemployment and Self-Assessed Health: Evidence from Panel Data[J]. Health Economics, 2009, 18(02): 161-179.

ceilings in health insurance policies, the policy effects may be different for patients with different total fees. Thus, exploring the distribution of policy effects is more comprehensive and interpretable. As discussed in subsection 5.2.2, since the analysis is conducted in hospital-by-hospital, patients having the same disease, but different total fees can be considered to have different disease severities. Therefore, quantile analysis is used to examine the differential effects of a policy on patients with the same disease but with different levels of severity. For example, patients with serious conditions may not be able to completely receive their needed treatment, such as stopping treatment halfway, due to their own or family's economic constraints. In contrast, patients with milder conditions may not encounter such constraints. Consequently, the integration policy is expected to have a greater impact on patients with higher total fees compared to those with lower total fees. As the severity of diseases may be milder among outpatients than among hospitalisation patients, the analysis of quantile effects will be specifically performed on hospitalisation total fees.

9.2.5 Statistical Analysis Framework

The analytical framework of this chapter, as depicted in Figure 9-1, builds upon the preceding discussion. The results of main analysis and further analysis are presented in Sections 9.3 and 9.4, respectively.

In the main analysis, the first step is the descriptive statistical analysis for outcomes and control variables, results are presented in subsection 9.3.1. All outcome variables are described in terms of sample size, mean, and standard deviation for both the treatment and control groups. The yearly trends are visualised through line graphs. All control variables are also described in terms of sample size, mean, standard deviation, and median for both the treatment and control groups. This comprehensive description provides a clear understanding of the data distribution.

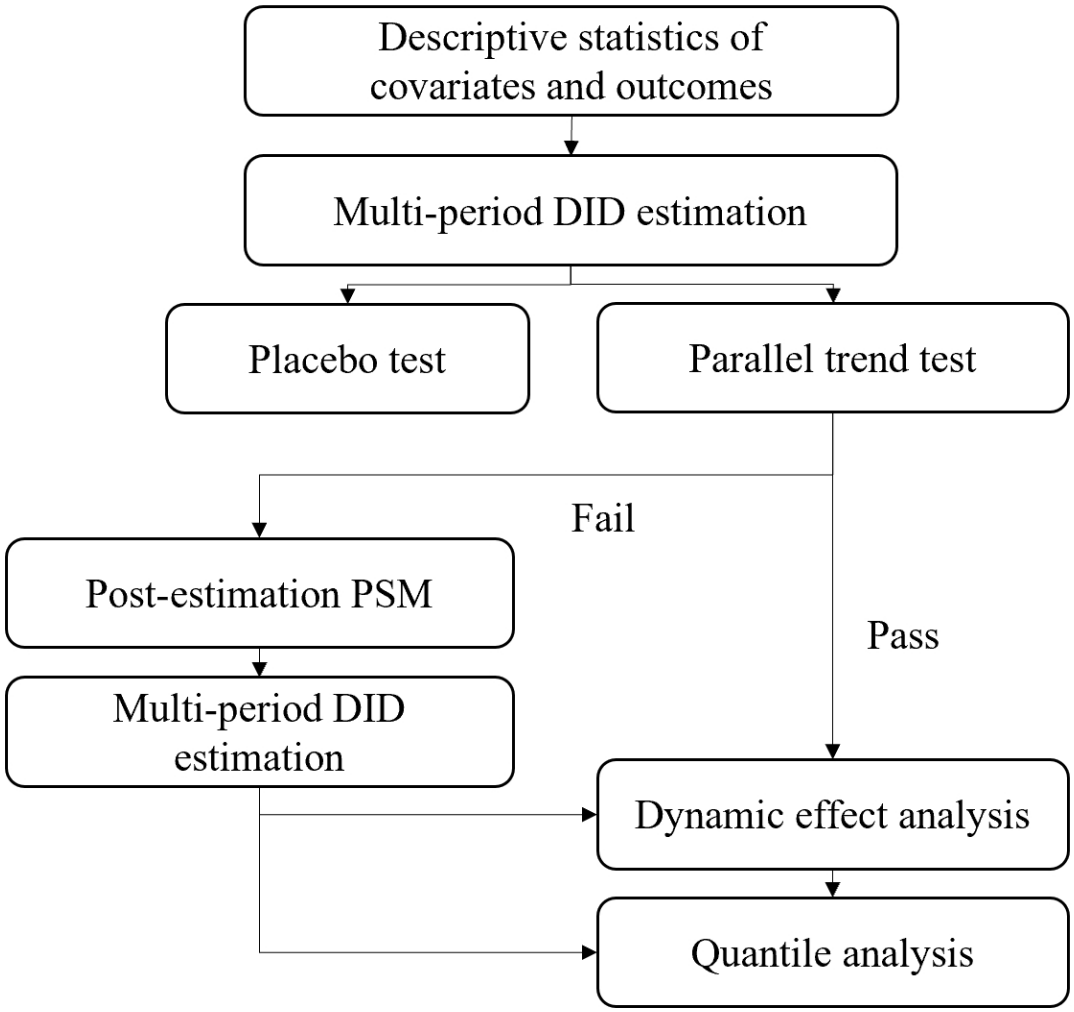

Figure 9-1 The statistical analysis framework of Chapter 9

Multi-period DID estimation is then performed, and results are presented in subsection 9.3.2. As discussed in subsection 5.2.1, appropriate regression models are selected based on the characteristics of the different outcome variables. The LOS are typical count data, so the Poisson model or Negative Binominal Model can be used. The likelihood ratio test is used to select between these two models. For fees, the Generalised Linear Model based on gamma distribution (GLM(gamma)) is used. Marginal effects are reported in the main test, and full regression results are provided in Appendix H.

The third step is the robustness test. The graphical results of the parallel trend test and the placebo test are reported in subsection 9.3.3. Regression results for parallel trend tests and descriptive statistics results for sampling in placebo tests are provided in Appendix H. If the

parallel trend assumption is likely to be violated, the PSM is employed in the fourth step. Since the purpose is to construct new control and treatment groups, the test results of the individual characteristics of the two groups before and after matching are reported in Section 9.3.4. Subsequently, the second DID analysis is conducted using the matched samples, with marginal effects reported in subsection 9.3.5 and full regression results provided in Appendix H. The results of further analysis are reported sequentially in Section 9.4, including the results of dynamic effect analysis and quantile effect analysis. Section 9.5 is the conclusion.

All statistical analyses are performed using STATA 16.0 software. The appendix provides a code sample, including DID estimation, parallel trend test, placebo test, PSM, dynamic regression and quantile regression.

9.3 Main Results

In this section, descriptive statistics results, DID analysis results, robustness test results, PSM results, and the second DID analysis results are reported in turn.

9.3.1 Descriptive Statistics

This subsection sequentially provides descriptive statistics results for outpatient control variables, hospitalisation control variables, outpatient total fees, length of stay in the hospital, and hospitalisation total fees. By comparing the control variables of the treatment and control groups, the similarity of characteristics between the two groups can be checked. Significant disparities in control variable values may indicate a potential deviation from parallel trends in the outcome variables. In this situation, the inclusion of control variables in the regression is necessary. A comparison of the outcome variables for the treatment and control groups allows for the examination of outcome trends before and after integration without the impact of control variables. If the trends in the two groups change, that would suggest a possible policy effect.

Table 9-3 summarises the descriptive statistics for control variables in outpatient visits, categorised by patient groups and hospitals. Notably, variations in the mean age of patients were observed across different hospitals, with Hospital E having the youngest patients (around

60 years old) and Hospital F having the oldest patients (around 70 years old). However, within each hospital, the mean age of the treatment and control groups were very similar. Similarly, the average number of chronic diseases and comorbidities exhibited close resemblance between the treatment and control groups. These similarities in age, chronic diseases, and comorbidities suggest a likelihood of similar trends in long-term medical fees for both groups.

However, variations were observed in the sex ratio and the diagnosis of Ischemic Heart Disease (IHD) between the treatment and control groups across different hospitals. In three hospitals, the control groups had a higher proportion of males compared to the treatment groups, with around a 10% difference. Additionally, in Hospitals A and F, the control groups had a higher proportion of angina pectoris and myocardial infarction diagnoses compared to the treatment groups, while the reverse was noted in Hospital B. A higher proportion of IHD-related diagnoses means more severe IHD symptoms, which may lead to higher medical fees. Therefore, controlling for these variables in outpatient fee analyses is necessary to obtain more accurate estimates of policy effects.

Table 9-4 summarises the descriptive statistics for the control variables in hospitalisation, categorised by patient groups and hospitals. Compared to outpatient visits, hospitalisations showed higher rates of IHD-related diagnoses, as well as a greater number of chronic diseases and comorbidities, suggesting poorer health status among these patients. However, the differences in characteristics between the treatment and control groups in hospitalisations were similar to those observed in outpatient visits. The age, number of chronic diseases, and number of comorbidities were very similar between the treatment and control groups. However, for IHD-related diagnoses, there were differences in the proportions of angina pectoris and myocardial infarction diagnoses between the treatment and control groups. Moreover, across all hospitals, the control groups consistently had a higher proportion of males compared to the treatment groups. Therefore, controlling for these variables, especially sex and IHD-related diagnoses, is crucial to obtain reliable policy effects.

Table 9-3 Descriptive statistics summary of outpatient control variables

Control Variables	Population	Hospital A			Hospital B			Hospital E			Hospital F		
		Obs.	Mean	S.D	Obs.	Mean	S.D	Obs.	Mean	S.D	Obs.	Mean	S.D
Age	Treatment	1,223	67.44	12.38	3,107	66.90	11.53	1,969	63.22	9.62	7,794	71.99	10.02
	Control	16,798	67.12	12.28	50,474	66.28	12.01	39,004	60.63	12.11	47,478	69.50	12.77
		Obs.	Frequency		Obs.	Frequency		Obs.	Frequency		Obs.	Frequency	
Sex (Male)	Treatment	1,223	40.15%		3,107	45.48%		1,969	53.43%		7,794	38.80%	
	Control	16,795	54.71%		50,471	59.29%		39,004	50.59%		47,472	48.26%	
Angina pectoris	Treatment	1,223	1.72%		3,107	0.48%		1,969	2.69%		7,794	0.53%	
	Control	16,798	3.65%		50,474	0.29%		39,004	1.32%		47,478	1.42%	
Myocardial infarction	Treatment	1,223	1.64%		3,107	1.42%		1,969	1.02%		7,794	0.22%	
	Control	16,798	3.82%		50,474	0.32%		39,004	1.18%		47,478	0.56%	
Acute coronary syndrome	Treatment	1,223	0.00%		3,107	3.54%		1,969	1.63%		7,794	0.09%	
	Control	16,798	0.12%		50,474	4.42%		39,004	0.28%		47,478	2.02%	
Ischemic heart disease	Treatment	1,223	98.86%		3,107	95.30%		1,969	95.78%		7,794	99.22%	
	Control	16,798	96.46%		50,474	95.26%		39,004	97.98%		47,478	96.21%	
Number of other chronic diseases	Treatment	1,223	0.62		3,107	1.03		1,969	0.03		7,794	0.47	
	Control	16,798	0.54		50,474	1.25		39,004	0.02		47,478	0.62	
Number of comorbidities	Treatment	1,223	0.03		3,107	0.39		1,969	0.11		7,794	0.15	
	Control	16,798	0.01		50,474	0.46		39,004	0.10		47,478	0.14	

Note: Obs.-Observations; S.D-Standard Deviation

Table 9-4 Descriptive statistics summary of hospitalisation control variables

Control Variables	Population	Hospital A			Hospital B			Hospital C			Hospital D		
		Obs.	Mean	S.D	Obs.	Mean	S.D	Obs.	Mean	S.D	Obs.	Mean	S.D
Age	Treatment	11,250	72.38	10.13	12,539	70.28	10.72	2,855	72.67	10.74	1,520	68.88	10.13
	Control	10,396	71.42	10.71	14,267	67.06	12.18	2,551	72.30	10.99	2,130	68.94	10.77
		Obs.	Frequency		Obs.	Frequency		Obs.	Frequency		Obs.	Frequency	
Sex (Male)	Treatment	11,251	27.65%		12,538	37.39%		2,855	41.12%		1,520	27.50%	
	Control	10,397	67.84%		14,267	62.64%		2,551	54.61%		2,130	69.06%	
Angina pectoris	Treatment	11,251	9.20%		12,539	14.38%		2,855	3.33%		1,520	11.05%	
	Control	10,397	12.91%		14,267	18.50%		2,551	3.41%		2,130	12.54%	
Myocardial infarction	Treatment	11,251	6.88%		12,539	7.14%		2,855	7.60%		1,520	16.97%	
	Control	10,397	9.79%		14,267	4.63%		2,551	10.86%		2,130	19.11%	
Acute coronary syndrome	Treatment	11,251	0.55%		12,539	4.47%		2,855	6.34%		1,520	1.45%	
	Control	10,397	0.45%		14,267	4.69%		2,551	4.70%		2,130	1.27%	
Ischemic heart disease	Treatment	11,251	99.58%		12,539	99.39%		2,855	93.52%		1,520	98.03%	
	Control	10,397	99.60%		14,267	99.03%		2,551	92.90%		2,130	97.70%	
Number of other chronic diseases	Treatment	11,251	2.30		12,539	3.00		2,855	2.27		1,520	2.01	
	Control	10,397	2.44		14,267	3.02		2,551	2.20		2,130	1.97	
Number of comorbidities	Treatment	11,251	2.52		12,539	1.80		2,855	3.57		1,520	1.20	
	Control	10,397	2.71		14,267	1.95		2,551	3.48		2,130	1.16	

Note: Obs.-Observations; S.D-Standard Deviation

Table 9-5 and Figure 9-2 provide descriptive results and visual representations of outpatient fees in four hospitals. Notable variations between treatment and control groups in Hospitals A, E, and F were observed after the integration. In Hospital A, the difference between the two groups narrowed as the average fees of the control group remained relatively stable, while the treatment group experienced a great decrease in fees in the year of integration. In Hospital E, the control group's average fees surpassed those of the resident group after integration. Conversely, in Hospital F, the group difference widened abruptly during the year of integration due to the treatment group's fee increase being much lower than that of the control group. These changes suggest that the policies implemented in these three cities may share similarities and may lead to a decrease in fees for residents' outpatient services.

Table 9-5 Summary of descriptive statistics for outpatient fees

Patient Group		Treatment group			Control group		
Hospital	Year	Obs.	Mean	S.D.	Obs.	Mean	S.D.
A	2016	141	371	506	2,462	261	398
	2017	293	256	398	3,060	278	409
	2018	505	263	290	7,655	242	364
	2019	284	284	302	3,621	293	431
B	2016	236	423	1,347	10,477	314	312
	2017	1,202	268	298	13,320	305	245
	2018	1,035	264	341	17,095	302	221
	2019	634	231	222	9,582	297	227
E	2016	309	223	166	16,483	225	215
	2017	975	252	445	18,863	245	358
	2018	685	277	710	3,658	289	627
F	2014	1,622	157	120	6,853	179	151
	2015	1,735	175	134	7,243	194	168
	2016	1,717	181	137	6,168	202	182
	2017	1,584	207	143	6,558	204	174
	2018	1,136	263	188	20,656	385	361

Note: Obs.-Observation; S.D.-Standard Deviation

However, it is important to note that Hospital B's treatment group had already experienced a substantial decline in average fees in the year prior to integration. This decline may be attributed to a large change in sample size, where it increased from 236 to 1202. Such a drastic change in sample size may result in the failure of subsequent pre-integration trend tests and render the

DID estimation unreliable. Furthermore, the substantial increase in the control group's sample size in Hospital F in 2018, for unknown reasons, may also impact the subsequent analysis.

These observations highlight the need to consider the potential influence of sample size changes on the reliability of the analysis, particularly in hospitals where significant fluctuations occur. It is important to carefully interpret the results of the pre-integration trend tests and subsequent estimations in light of these circumstances.

Table 9-6 and Figure 9-3 summarise and display descriptive results of Length of Stay (LOS) in four hospitals. The noteworthy changes in LOS differences between the treatment and control groups were observed in Hospitals A and B after the integration. In Hospital A, the difference between the two groups almost disappeared after integration due to the slowdown in the decline of average LOS in the treatment groups. Similarly, in Hospital B, the difference between the two groups narrowed as the decline in average LOS in the treatment group slowed down. These findings suggest that the policies in these two cities may be similar and may lead to an increase in LOS. However, Hospital C exhibited a different pattern. After integration, the LOS of the treatment group decreased while that of the control group increased. If this is influenced by integration, it indicates that the integration policy in this city was contrary to the previous two cities and had a negative effect on LOS. As for Hospital D, since the LOS of both groups increased simultaneously, it is unclear whether the policy had any impact on LOS. These observations highlight the potential varying effects of integration policies on LOS across different cities.

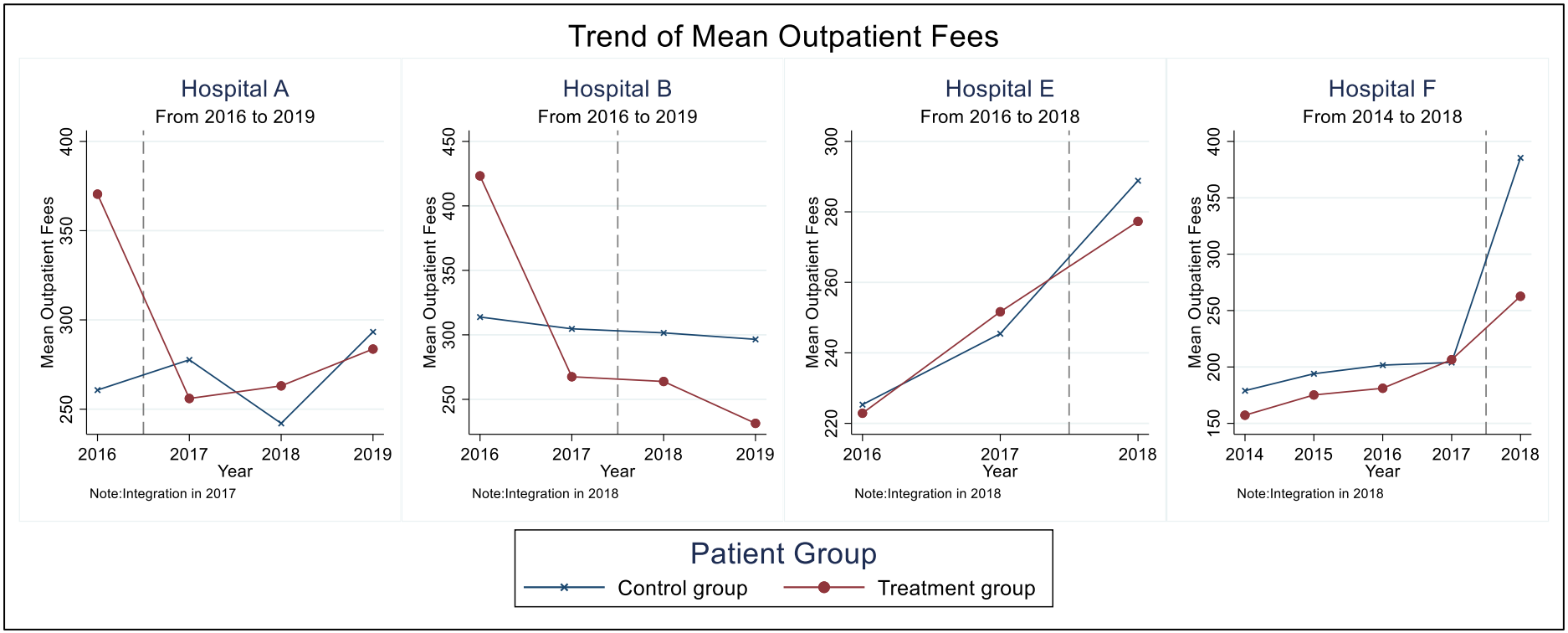

Figure 9-2 Trend of mean outpatient fees

Table 9-6 Summary of descriptive statistics for length of stay in hospital

Patient Group		Treatment group			Control group		
Hospital	Year	Obs.	Mean	S.D.	Obs.	Mean	S.D.
A	2013	1,052	11.40	6.16	1,018	12.70	7.40
	2014	1,158	11.65	5.95	1,180	12.89	7.84
	2015	1,397	11.22	7.68	1,276	11.61	7.28
	2016	1,520	10.64	4.97	1,235	11.37	7.83
	2017	1,711	10.55	5.12	1,225	10.54	5.46
	2018	2,597	9.75	4.39	2,648	9.65	4.23
	2019	1,805	9.66	4.29	1,804	9.63	4.46
B	2014	1,739	9.24	4.04	3,113	11.35	6.43
	2015	1,786	9.36	4.23	2,597	11.45	5.69
	2016	2,224	10.35	4.21	2,353	12.16	5.98
	2017	2,613	10.07	4.27	2,362	12.28	6.29
	2018	2,614	9.70	4.82	2,479	11.19	5.19
	2019	1,563	8.53	3.31	1,363	9.17	3.33
C	2016	188	11.38	6.22	705	10.88	10.57
	2017	261	11.82	9.50	955	10.66	18.63
	2018	911	10.74	8.74	562	11.31	13.24
	2019	1,483	10.26	8.60	278	10.20	10.07
D	2016	639	11.40	5.60	993	12.24	7.31
	2017	835	13.35	11.46	1,103	13.69	10.00

Note: Obs.-Observation; S.D.-Standard Deviation

Table 9-7 and Figure 9-4 present descriptive results and visual representations of hospitalisation total fees in four hospitals. Analysing the hospitalisation fees, notable changes were observed between the treatment and control groups after the implementation of integration in all hospitals. In Hospital A, the treatment group's fees surpassed those of the control group. This occurred because the fees for the treatment group decreased at a slower rate compared to the control group. Similarly, in Hospital B, the difference in fees between the two groups narrowed due to an increase in fees for the treatment group and a decrease in fees for the control group. And for Hospital D, the fees of the treatment group also surpassed those of the control group. This was a result of the treatment group experiencing a greater increase in fees compared to the control group after the integration. These findings in these three hospitals suggest that the integration policies of these three cities may be similar and may have raising effects on hospitalisation fees.

However, Hospital C exhibited a different pattern. Although both the treatment and control

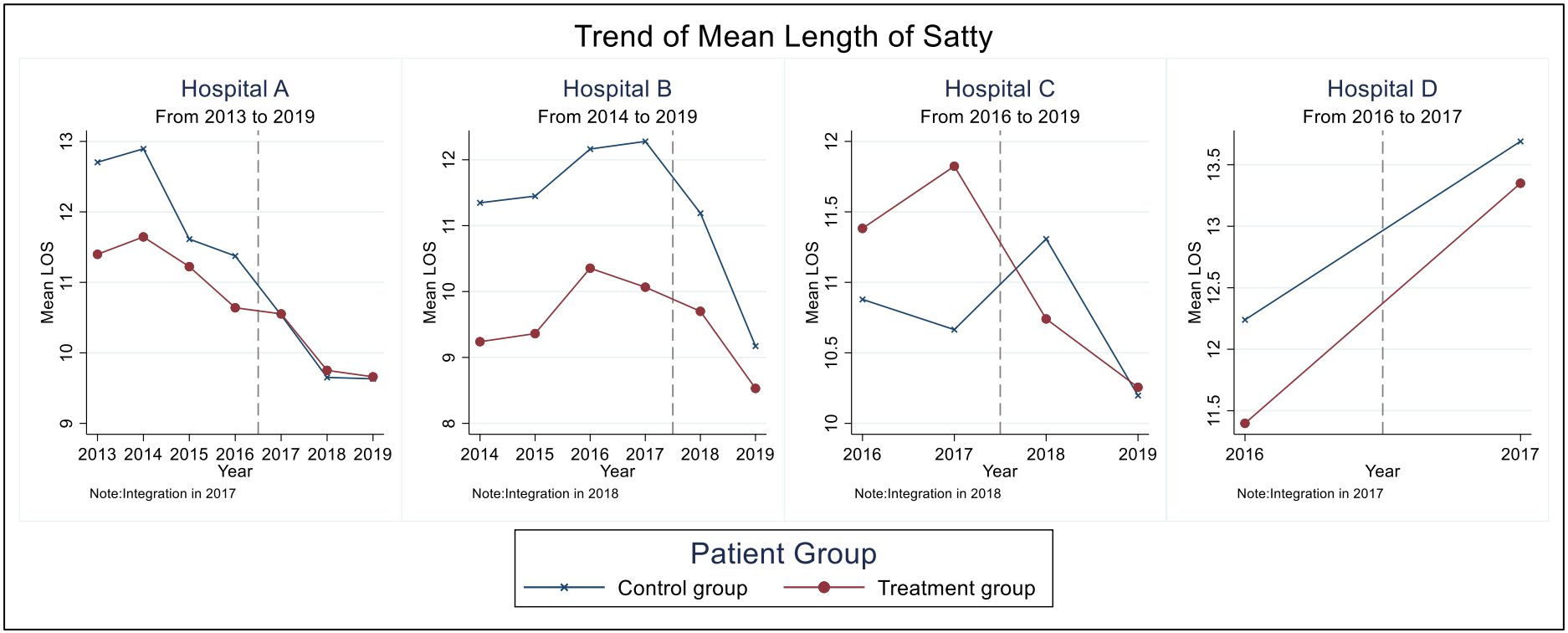

Figure 9-3 Trend of mean length of stay

groups experienced an increase in fees during the year of integration, the increase in fees was smaller for the treatment group. And in the second year of implementation, the fees continued to rise for the treatment group, while the control group's fees decreased. This implies two possible policies and effects: 1) the integration policy may have an immediate inhibitory effect on the treatment group's fees, indicating that the policy is opposite to the other three cities; 2) the integration policy may be similar to the other cities, but the raising effect of this policy may not become apparent until the second year due to various reasons such as differences in policy implementation details.

Table 9-7 Summary of descriptive statistics for hospitalisation fees

Patient Group		Treatment group			Control group		
Hospital	Year	Obs.	Mean	S.D.	Obs.	Mean	S.D.
A	2013	1,052	9,135	9,199	1,018	10,755	13,921
	2014	1,158	10,616	10,784	1,180	12,049	13,328
	2015	1,397	10,992	14,529	1,276	11,876	15,467
	2016	1,520	11,538	9,826	1,235	12,449	12,967
	2017	1,714	10,535	9,873	1,225	10,182	9,225
	2018	2,597	9,650	10,998	2,648	9,314	8,698
	2019	1,813	10,297	10,128	1,815	10,176	10,195
B	2014	1,739	4,129	6,702	3,113	6,093	10,253
	2015	1,786	4,620	5,914	2,597	6,811	7,430
	2016	2,224	5,438	6,327	2,353	7,799	8,874
	2017	2,613	6,437	7,763	2,362	9,235	11,506
	2018	2,614	6,920	9,171	2,479	8,114	9,899
	2019	1,563	6,046	6,917	1,363	7,192	8,004
C	2016	188	10,388	12,027	705	12,224	13,534
	2017	261	11,196	13,398	955	12,425	16,828
	2018	915	11,827	12,668	572	14,968	20,661
	2019	1,491	13,451	14,404	319	14,216	15,620
D	2016	639	5,841	4,910	994	7,481	9,893
	2017	849	8,103	8,714	1,105	7,878	8,806

Note: Obs.-Observation; S.D.-Standard Deviation

In summary, whether outpatient or hospitalisation, differences in control variables between treatment and control groups, especially sex and IHS-related diagnoses, must be controlled in subsequent regressions. Failing to do so may undermine the reliability of the estimated policy effects. From the changing trends of the outcome variables of the treatment and control groups,

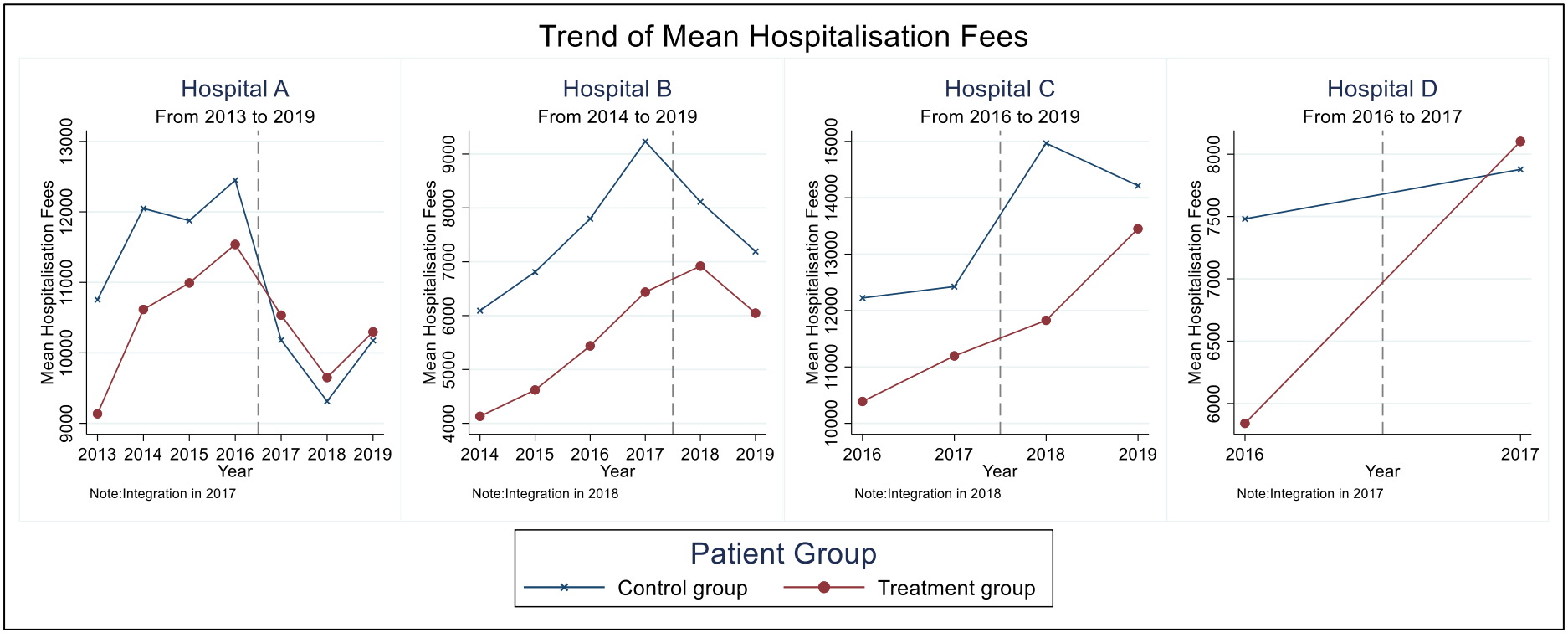

Figure 9-4 Trend of mean hospitalisation fees

it can be seen that after being promoted nationwide, the overall effects of integration policies across different cities may be similar. However, whether different policies actually produced effects, and whether the effects are really similar, can only be answered after removing the bias caused by observable and unobservable confounding factors as much as possible through subsequent DID analysis and PSM-DID analysis.

9.3.2 Results from DID

In this chapter, the focus is on examining the effects of integration policies rather than considering the effects of all variables that may influence outcomes. Thus, this subsection only provides a summary of the marginal effects of the integration implementation variable in the DID regression, as shown in Table 9-8, while the full regression results are available in Appendix H.

Table 9-8 Summary of DID estimation results

Outcome	Hospital	Policy Effect						
		Obs.	M.E.	S.E.	z	P>|z|	95% CI	
Outpatient total fees	A	18,018	-70	27	-2.58	**0.010**	-123	-17
	B	53,578	-43	9	-4.81	**0.000**	-60	-25
	E	40,973	-18	11	-1.65	0.099	-40	3
	F	55,266	-58	5	-10.49	**0.000**	-70	-47
Length of Stay	A	21,624	0.98	0.14	7.13	**0.000**	0.71	1.25
	B	26,805	0.77	0.13	6.04	**0.000**	0.52	1.02
	C	5,343	-0.76	0.45	-1.69	0.092	-1.64	0.12
	D	3,570	0.43	0.47	0.92	0.357	-0.49	1.36
Hospitalisation total fees	A	21,646	1,471	291	5.05	**0.000**	899	2,042
	B	26,805	1,162	224	5.19	**0.000**	723	1,601
	C	5,406	-977	986	-0.99	0.322	-2,909	955
	D	3,587	1,773	567	3.13	**0.002**	661	2,885

Note: fee estimation using GLM(gamma) model, LOS estimation using Negative-binomial model; Obs.-Observation; M.E.-Marginal Effect; S.E.-Standard Error; CI-Confidence Interval

For outpatient services, after controlling for sex, age, IHD-related diagnoses, chronic diseases, and comorbidities, the integration policy led to varying sizes of reductions in outpatient total fees for IHD patients across all four hospitals (ranging from -18 CNY to -70 CNY). Three of

these results were statistically significant at the 1% level, while one showed significance at the 10% level. For hospitalisation service, after adjusting for control variables, the integration policy demonstrated varying effects on total fees in three hospitals, resulting in fee increases ranging from 1,162 CNY to 1,173 CNY (all with p-values below 0.01). However, no significant effect was observed in the other hospital. Additionally, the integration policy also had significant positive effects on LOS in two hospitals, with increases of 0.98 days and 0.77 days (both with p-values less than 0.01). Conversely, one hospital experienced a decrease in LOS by 0.76 days, although this effect was only statistically significant at the 10% level (p=0.092). The other hospital showed no significant effect on LOS (p=0.357).

These estimates support the conjectures from the descriptive statistical results on the policy effect in the previous subsection. The policy effects of integration for outpatient services appear to be more similar across cities, while there is less uniformity in the effects for hospitalisation services. This finding aligns with the conclusions drawn from the qualitative analysis conducted in Chapter 7, which suggested that changes in outpatient reimbursement terms exhibit greater similarity across different cities compared to changes in hospitalisation reimbursement terms.

9.3.3 Results of Robustness Tests

This subsection enhances the reliability of the DID estimation results, building upon the appropriateness discussed in Chapter 5 and subsection 9.2.2. Two robustness tests, the pre-integration parallel trend test and the placebo test, are presented as discussed in subsection 9.2.4.

The pre-integration parallel trend test requires at least two years of pre-integration data for comparison. However, among the 12 outcomes data (3 outcome variables multiplied by 4 hospitals), Hospital A has only one-year pre-integration data of outpatient total fees, while Hospital D also has only one-year prior data of hospitalisation total fees and LOS. As a result, conducting parallel trend tests for these specific outcomes is not feasible. Matching will be employed in the subsequent subsection to examine the robustness of the results for these outcomes. For the other nine outcomes, if non-parallel pre-integration trends are observed, matching will also be conducted in the subsequent subsection.

Figures 9-5 to 9-7 display the results of the pre-integration trend tests for all outcomes (see Appendix for full regression results). The black dots in these figures represent the coefficient for the difference in the difference in outcome between the two groups (resident versus non-resident) in a specific year versus the control year (the year before integration). The vertical lines passing through the black dots are the 95% confidence intervals for the estimated coefficients. If a vertical black line intersects the horizontal dotted line which represents a coefficient of 0, the between-group difference does not change significantly between the two years. If all the vertical black lines intersect with the horizontal dotted lines within a hospital, it means that before the implementation of the integration, the outcomes of the treatment group and the control group had parallel trends.

For outpatient fees, only the treatment and control groups in Hospital E showed a pre-integration parallel trend, while Hospitals B and F did not. Regardless of hospitalisation fees and LOS, the treatment and control groups showed parallel trends before integration in all 3 tested hospitals. Satisfying the pre-integration parallel trend suggests that the parallel trend assumption is very likely to be satisfied in the absence of integration. Therefore, these results show that in the previous subsection, the DID estimations of the policy effects on the hospitalisation fees and LOS in Hospitals A, B, and C, and on outpatient fees in Hospital E are more reliable. However, the DID estimates of policy effects on outpatient fees in Hospitals B and F may be biased, and the robustness of the results needs to be checked by subsequent matching and re-estimation of DID.

Figures 9-8 to 9-10 show the results of all placebo tests. These tests involved randomly assigning treatment group samples and examining the distribution of placebo DID estimates coefficients and p-values across 500 iterations for each hospital as discussed in Chapter 5. In each figure, the black vertical dashed line represents the estimated mean of the 500-sample placebo coefficients, while the red solid line represents the true DID estimate obtained from subsection 9.3.2.

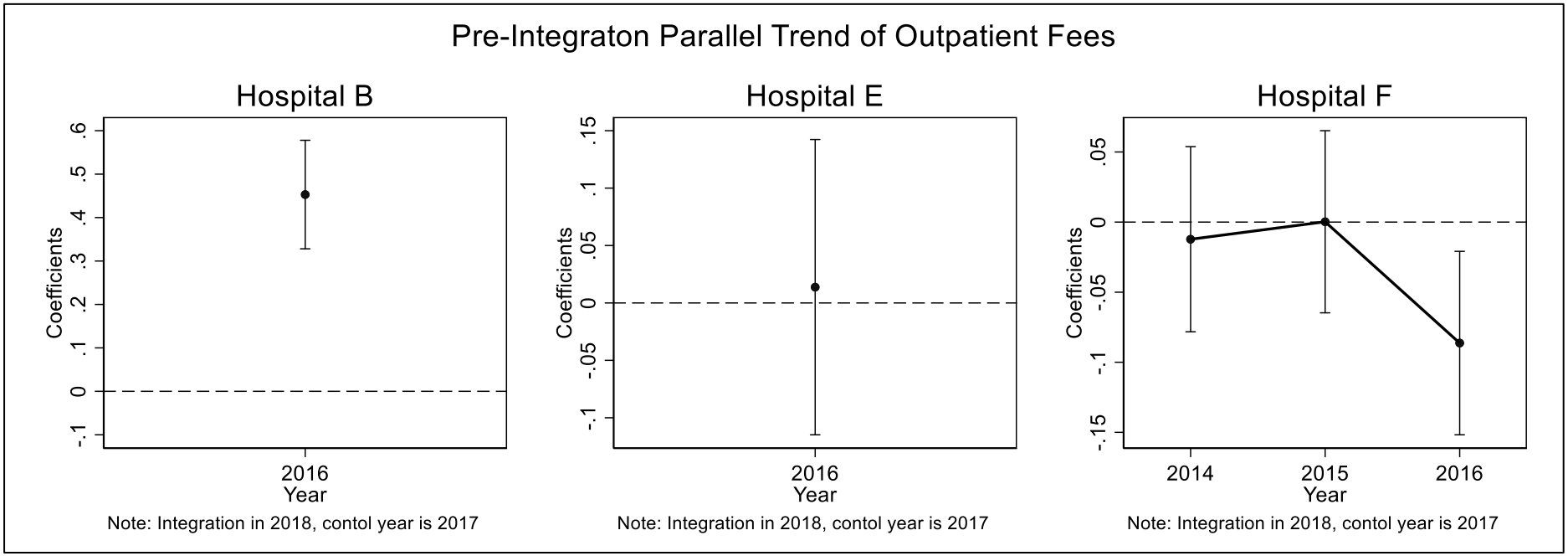

Figure 9-5 Pre-integration parallel trend of outpatient fees

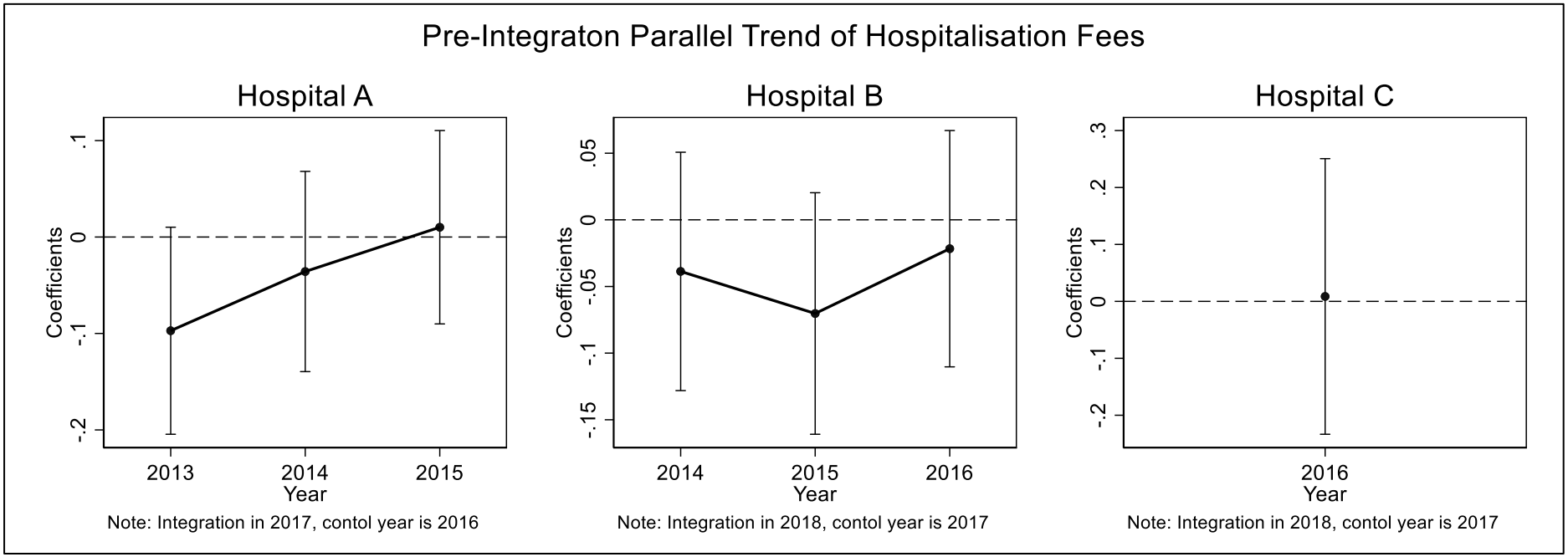

Figure 9-6 Pre-integration parallel trend of hospitalisation fees

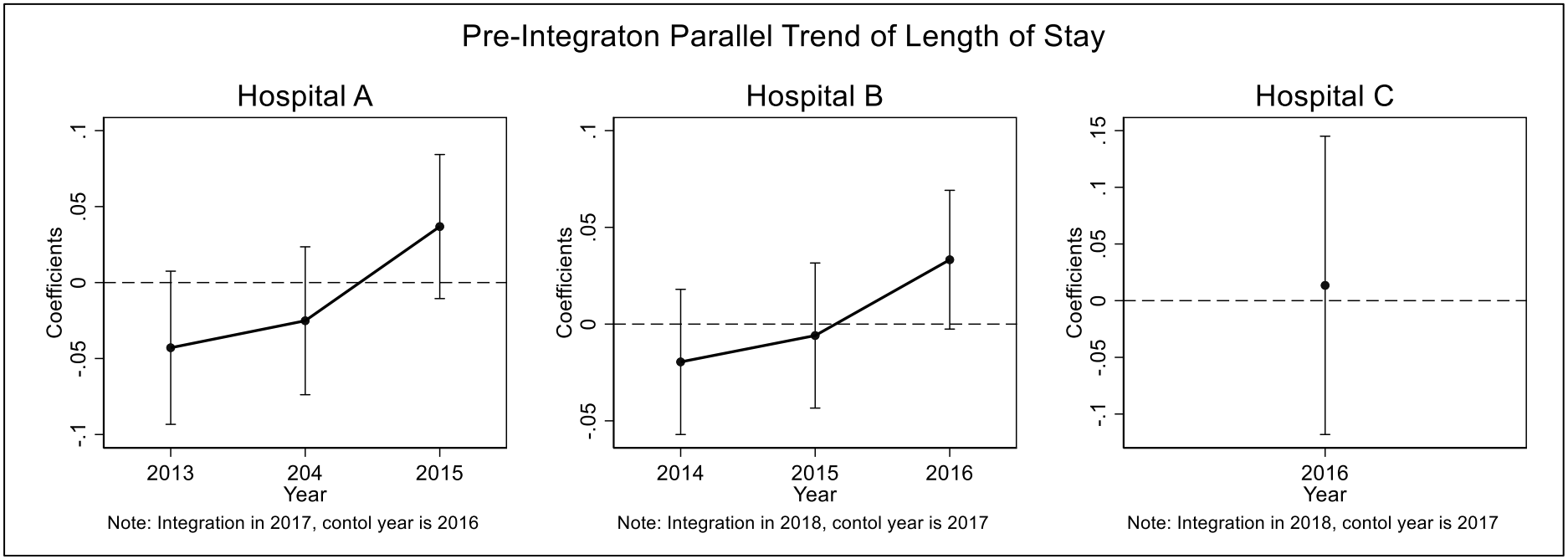

Figure 9-7 Pre-integration parallel trend of length of stay

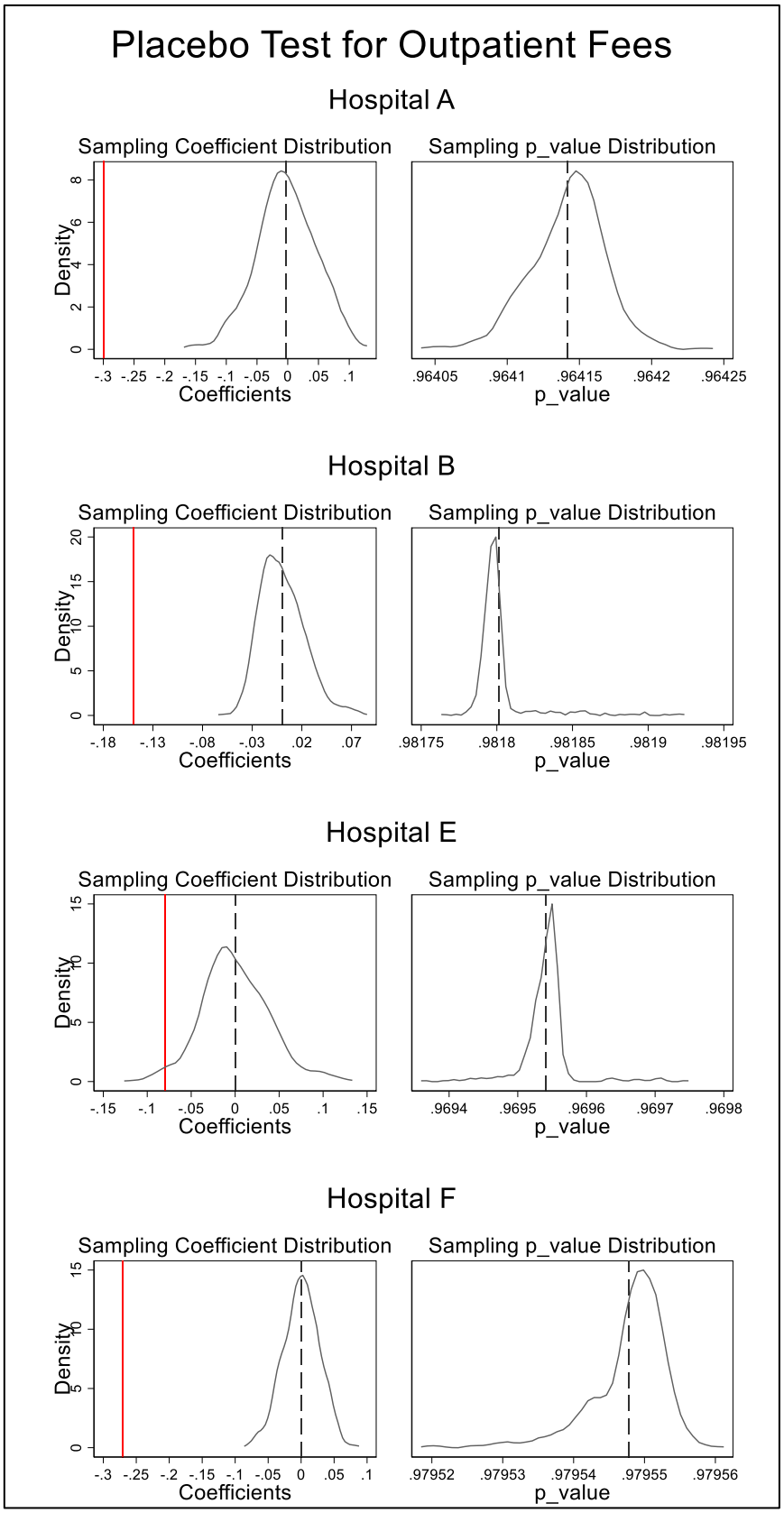

Figure 9-8 Placebo test for outpatient fees

(Note: Vertical dashed line—estimated placebo policy effect from 500 samples; solid red line—estimated true policy effect from subsection 9.3.2)

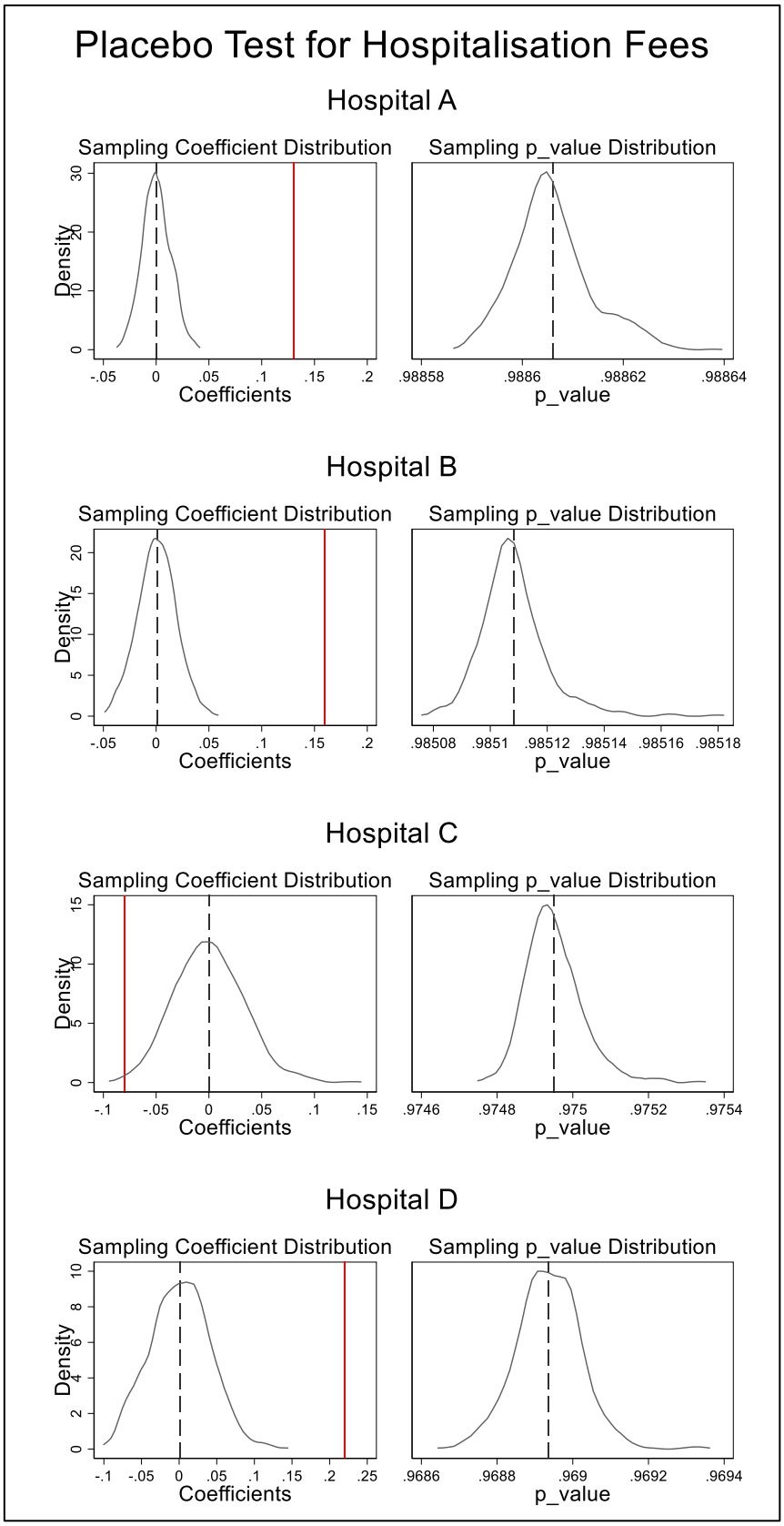

Figure 9-9 Placebo test for hospitalisation fees

(Note: Vertical dashed line—estimated placebo policy effect from 500 samples; solid red line—estimated true policy effect from subsection 9.3.2)

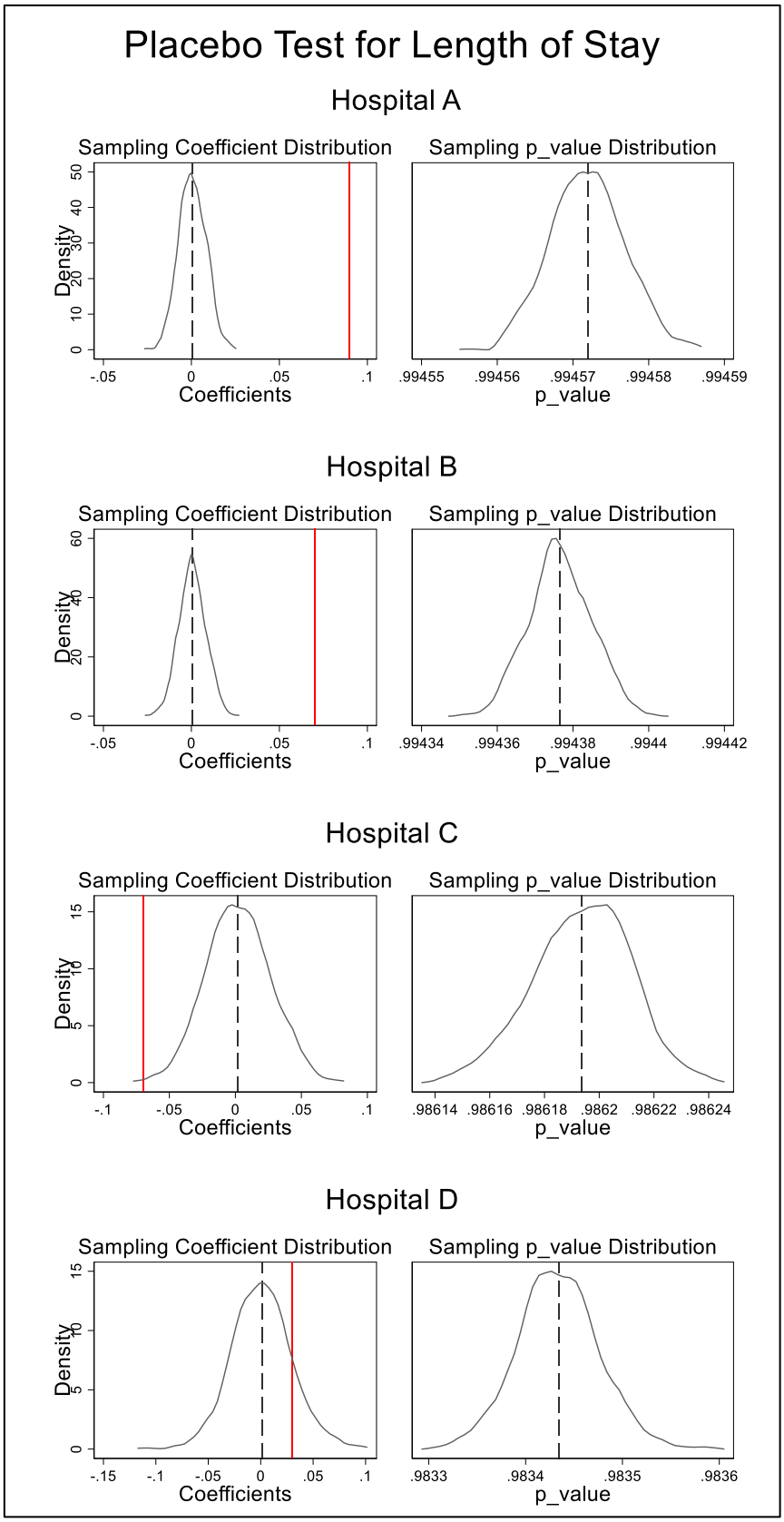

Figure 9-10 Placebo test for length of stay

(Note: Vertical dashed line—estimated placebo policy effect from 500 samples; solid red line—estimated true policy effect from subsection 9.3.2)

The distributions of the sampled placebo DID estimated coefficients consistently centred around the value of 0, indicating the absence of any policy effect for the placebo/fake policy. Additionally, all p-values associated with these placebo tests were greater than 0.95, further supporting the conclusion that no placebo policy effect was detected. In contrast, the true statistically significant DID estimates reported in subsection 9.3.2 that were deviated from the distribution of the sampling coefficients. This suggests that these policy effects are indeed attributable to the integration policy and not a result of random errors in the sample or model specification.

In summary, the tests conducted in this subsection support the reliability of 7 out of 12 DID results. For the remaining outcomes, which either exhibit non-parallel pre-integration trends (outpatient total fees in Hospitals B and F) or lack sufficient pre-integration data for testing (outpatient total fees in Hospital A, hospitalisation total fees and LOS in Hospital D), matching the treatment and control groups will be performed in the next subsection.

9.3.4 Results of Propensity Score Matching

For the 5 outcomes that may have biased estimates as indicated in the previous subsection, a nearest-neighbour PSM was employed to select treatment and control groups with similar control variable values year-by-year. In addition, alternative matching methods such as kernel matching and local linear regression matching, as discussed in Chapter 5, were also applied to data when groups with no statistically significant differences in control variables could not be obtained by nearest-neighbour PSM.

Table 9-9 presents the test results of differences in control variables between the treatment and control groups before and after nearest-neighbour PSM. The matching effects were particularly notable for the outpatient total fees data of Hospitals A and B. After matching, the annual differences in control variables between the treatment and control groups, as indicated by the mean and median values (found in the "MeanBias" and "MedianBias" columns of the table), decreased. The overall differences between the two groups were no longer statistically significant (all p-values greater than 0.1, as shown in the "p>chi2" column). However, for the

outpatient total fees data of Hospital F and the hospitalisation data of Hospital D, there were still significant overall differences in the control variables between the matched treatment and control groups in certain years. These included the outpatient total fees of Hospital F in 2016, the hospitalisation total fees of Hospital D in 2016 and 2017, and the LOS of Hospital D in 2016. Consequently, alternative matching methods were used for the data from Hospital D and Hospital F.

Table 9-10 presents the test results of differences in control variables between the treatment and control groups before and after matching using different methods. After applying kernel matching, the overall differences in control variables between the treatment and control groups in Hospital D were no longer statistically significant. This was observed for both hospitalisation total fees and LOS. However, for Hospital F, the difference between the two groups in 2016 remained statistically significant ($p < 0.05$) even after kernel matching or local linear regression matching was performed.

In summary, the matching methods employed in this subsection successfully eliminated overall differences in control variables between the treatment and control groups in 4 out of 5 outcome datasets, which had potentially biased results in previous DID estimates. Outcomes in these matched datasets are now more likely to meet the parallel trend assumption than before the matching, enhancing the reliability of DID estimations in the next subsection. However, for the outpatient data from Hospital F, no suitable match could be found between the treatment and control groups. As a result, the DID estimation using this dataset would inevitably contain bias that cannot be eliminated, rendering the results unreliable. Therefore, this particular dataset will not be used for any subsequent analyses. Furthermore, it is important to note that the matching process resulted in a reduction in sample size, as indicated in Table 9-11. This reduction may lead to a decrease in the representativeness of the population in the subsequent analysis based on these matched datasets.

Table 9-9 Post PSM sample balance tests, nearest-neighbour matching

Outcomes	Year	Sample	LR chi2#	p>chi2	Mean Bias	Median Bias
Outpatient total fees in Hospital A	2016	Unmatched	14.04	0.121	7.2	4.6
		Matched	2.71	0.911	4.2	5.7
	2017	Unmatched	49.67	0.000	10.7	8.8
		Matched	4.7	0.967	4.3	5.5
	2018	Unmatched	120.35	0.000	8.6	6.7
		Matched	6.15	0.977	2.9	2.4
	2019	Unmatched	48.02	0.000	10.4	4.8
		Matched	9.05	0.617	4.5	4.1
Outpatient total fees in Hospital B	2016	Unmatched	198.7	0.000	13.8	8.2
		Matched	8	0.966	7.1	7.3
	2017	Unmatched	611.08	0.000	9.8	3.6
		Matched	18.71	0.848	3.2	2.3
	2018	Unmatched	783.41	0.000	12.8	5.8
		Matched	19.42	0.777	4	3.4
	2019	Unmatched	466.68	0.000	13.7	8
		Matched	9.14	0.997	3.1	3.1
Outpatient total fees in Hospital F	2014	Unmatched	293.08	0.000	9.1	5.5
		Matched	5.99	0.967	1.8	1.2
	2015	Unmatched	423.65	0.000	11.6	7.9
		Matched	13.04	0.445	2.8	2.2
	2016	Unmatched	1284.56	0.000	18.8	9.7
		Matched	33.93	**0.009**	4.3	1.7
	2017	Unmatched	1547.68	0.000	21.4	14.4
		Matched	23.45	0.135	2.8	1.9
	2018	Unmatched	418.53	0.000	11.4	8.7
		Matched	5.28	0.994	2.0	1.7
Hospitalisation total fees in Hospital D	2016	Unmatched	364.21	0.000	9.3	6.1
		Matched	47.72	**0.028**	5.1	4.5
	2017	Unmatched	463.86	0.000	9	3.8
		Matched	71.78	**0.000**	5	4
LOS in Hospital D	2016	Unmatched	363.86	0.000	9.3	6.1
		Matched	58.03	**0.002**	5.7	5.3
	2017	Unmatched	467.77	0.000	8.9	4.2
		Matched	45.12	0.062	3.9	3.1

Note: #the likelihood-ratio test of the joint insignificance of all the regressors

Table 9-10 Post PSM sample balance tests, other matching methods

Outcomes	Year	Sample	LR chi2#	p>chi2	Mean Bias	Median Bias
Outpatient total fees in Hospital F	2014	Unmatched	293.08	0.000	9.1	5.5
		Matched_1	2.71	1.000	1.1	0.6
		Matched_2	5.99	0.967	1.8	1.2
	2015	Unmatched	423.65	0.000	11.6	7.9
		Matched_1	14.35	0.499	2.5	1.5
		Matched_2	13.04	0.445	2.8	2.2
	2016	Unmatched	1284.56	0.000	18.8	9.7
		Matched_1	34.19	**0.012**	3.9	2.3
		Matched_2	33.93	**0.009**	4.3	1.7
	2017	Unmatched	1547.68	0.000	21.4	14.4
		Matched_1	16.51	0.488	2.3	1.6
		Matched_2	23.45	0.135	2.8	1.9
	2018	Unmatched	418.53	0.000	11.4	8.7
		Matched_1	43.05	**0.002**	4.3	3.1
		Matched_2	5.28	0.994	2.0	1.7
Hospitalisation total fees in Hospital D	2016	Unmatched	364.21	0.000	9.3	6.1
		Matched_1	7.94	1.000	2.3	1.9
	2017	Unmatched	463.86	0.000	9.0	3.8
		Matched_1	17.48	0.983	2.7	1.8
LOS in Hospital D	2016	Unmatched	363.86	0.000	9.3	6.1
		Matched_1	7.92	1.000	2.3	1.9
	2017	Unmatched	467.77	0.000	8.9	4.2
		Matched_1	18.06	0.977	2.7	1.9

Note: Matched_1 using kernel matching; Matched_2 using local linear regression matching; #the likelihood-ratio test of the joint insignificance of all the regressors

Table 9-11 Sample size changes before and after matching

Outcomes	Sample size before matching		Sample size after matching	
	Treatment	Control	Treatment	Control
Outpatient total fees in Hospital A	1,223	16,798	1,211	9,167
Outpatient total fees in Hospital B	3,107	50,474	3,102	22,964
Outpatient total fees in Hospital F	7,794	47,478	Remove from analysis	
Hospitalisation total fees in Hospital D	1,488	2,099	1,471	2,097
LOS in Hospital D	1,488	2,099	1,457	2,094

9.3.5 Results from PSM+DID

Table 9-12 summarises the marginal effects of integration in the second DID using matched datasets and provides their first DID marginal effect results for comparison. The PSM-DID results show consistent effect directions and significances with the first DID results, with only minor variations in effect size. These findings support the robustness and reliability of the DID estimation.

Table 9-12 Summary of PSM-DID results, by hospital

Outcome	Method	Policy Effect					
		Obs.	M.E.	S.E.	P value	95% CI	
Outpatient total fees in Hospital A	DID	18,018	-70	27	0.01	-123	-17
	PSM+DID	10,378	-71	27	0.008	-123	-18
Outpatient total fees in Hospital B	DID	53,578	-43	9	0	-60	-25
	PSM+DID	26,066	-30	9	0.001	-49	-12
Length of Stay in Hospital D	DID	3,570	0.43	0.47	0.357	-0.49	1.36
	PSM+DID	3,551	0.45	0.47	0.345	-0.48	1.37
Hospitalisation total fees in Hospital D	DID	3,587	1,773	567	0.002	661	2,885
	PSM+DID	3,568	1,764	570	0.002	648	2,881

Note: fee estimation using GLM(gamma) model, LOS estimation using Negative-binomial model; Obs.-Observation; M.E.-Marginal Effect; S.E.-Standard Error; CI-Confidence Interval

In summary, this section has conducted a series of analyses and tests around the 4th objective of this thesis. Among 12 DID/PSM-DID estimation results, 11 are considered to be and reliable. Collectively, these results suggest that integration policies across cities have more similar fee-reducing effects on the outpatient total fees for IHD patients, but their effects on hospitalisation total fees and LOS are less similar. The next section will provide some possible explanations for how and why these policy effects differ.

9.4 Further Results

This section conducts two different analyses (dynamic effects and quantile effects) around the fifth objective of this thesis to explore whether and how there are differences in the mechanism of action of policies in different cities.

9.4.1 Results from Dynamic Effect Analyses

Since dynamic effects analysis requires at least two years of data after policy implementation as discussed in subsection 9.2.4, only outpatient data from Hospitals A and B, and hospitalisation data from Hospitals A, B, and C, can be analysed. Tables 9-13 summarise the results of the dynamic effects.

For outpatient total fees in both Hospitals A and B, the analysis reveals no statistically significant temporal changes in the policy effect, indicating a stable impact of integration over time. In Hospital A, where integration spanned three years, outpatient fees were affected by -87 CNY, -53 (=-87+34) CNY, and -69 (=-87+18) CNY, respectively. Although the fee reduction effect faded over time, the differences were not significant (p=0.231 and 0.564). In Hospital B, where integration spanned two years, outpatient fees were affected by -27 CNY, and -37 (=-27-10) CNY, respectively. However, the difference was not significant (p=0.473).

For hospitalisation services, the dynamic effects in the three hospitals exhibited notable disparities. Hospital A sustained a consistent and stable impact on both fees and length of stay (LOS) over time. In Hospital B, the impact of policy on fees was stable over time, but the impact on LOS expanded over the two years, with an increase from 0.53 days in the first year to 1.19 days in the second (p=0.002). This expansion may be attributed to more conservative behaviours by doctors or other decision makers at this hospital than those at other hospitals. Thus, a likely scenario is that in the first year of implementation, doctors did not make major adjustments to the treatment plan for all patients, but only for some patients. In the second year, doctors first assessed the effects of the first-year adjustments and then changed the treatment plan for all patients. This cautious process likely caused a delay in the full display of policy effects.

In Hospital C, the impact on fees reversed dramatically in the second year of implementation, while the impact on LOS, though not reversed, was also substantially reduced. The impact on hospitalisation fees was -2,136 CNY in the first year of integration but was 668 (=-2,136+2,804) CNY in the second year. The two-year change was significant (p=0.02). This reversion explains

Table 9-13 Summary of dynamic effect analyses results

Outcome	Hospital	Observation	Policy Effect					
			Marginal Effect	Standard Error	z	P>|z|	95% CI	
Outpatient Fees	A	10,378	Integration year	-87	29	-2.97	**0.003**	-144	-30
			Difference of 2nd year	34	28	1.20	0.231	-21	89
			Difference of 3rd year	18	32	0.58	0.564	-44	81
	B	26,066	Integration year	-27	11	-2.46	**0.014**	-48	-5
			Difference of 2nd year	-10	15	-0.72	0.473	-39	18
Length of Stay	A	23,298	Integration year	0.71	0.21	3.36	**0.001**	0.30	1.12
			Difference of 2nd year	0.35	0.23	1.54	0.124	-0.10	0.80
			Difference of 2 years after	0.40	0.25	1.60	0.109	-0.09	0.88
	B	26,805	Integration year	0.53	0.15	3.63	**0.000**	0.244	0.82
			Difference of 2nd year	0.66	0.21	3.11	**0.002**	0.243	1.08
	C	5,343	Integration year	-1.16	0.50	-2.31	**0.021**	-2.142	-0.17
			Difference of 2nd year	1.02	0.57	1.79	0.074	-0.10	2.14
Hospitalisation Fees	A	21,646	Integration year	1,174	446	2.63	**0.008**	301	2,047
			Difference of 2nd year	367	462	0.79	0.427	-538	1,272
			Difference of 3rd year	374	497	0.75	0.452	-600	1,348
	B	26,805	Integration year	1,217	266	4.58	**0.000**	696	1,739
			Difference of 2nd year	-134	343	-0.39	0.697	-806	539
	C	5,406	Integration year	-2,136	1,106	-1.93	0.053	-4,303	31
			Difference of 2nd year	2,804	1,202	2.33	**0.020**	448	5,161

Note: fee estimation using GLM(gamma) model, LOS estimation using Negative-binomial model; CI-Confidence Interval

why the city-level average treatment effect was not significant in the subsection 9.3.2. The effect on LOS was -1.16 days (p=0.021), in the first year of implementation, but changed to -0.14 (=-1.16+1.02) days by the second year, although the difference was only significant at the 10% level (p =0.074). The attenuation of the effect over time explains the low significance of the city-level mean treatment effect in the previous analyses. However, without being able to link to the policy clause details (due to anonymity), it is difficult to speculate on the reason for this reversal.

In short, the dynamic effect analysis confirmed again that the integration effects were more similar in outpatient but more different in hospitalisation.

9.4.2 Results from Quantile Effect Analyses

The quantile effect analysis aimed to investigate diverse policy effects of integration on patients with varying total fees/disease severity. As mentioned in subsection 9.2.5, this analysis focused on hospitalisation, as hospitalisation patients' conditions are typically more severe compared to outpatients, and their treatments are more likely to be influenced by financial constraints. Therefore, quantile regressions were applied specifically to hospitalisation total fees. Table 9-14 displays the marginal effects of the integration on hospitalisation total fees for patients across different quantiles, ranging from the 5th to the 95th percentile, in each hospital. For example, the value "500" in the 50th quantile of Hospital A indicates that patients with total hospitalisation fees at the median (50%) of all patients experienced a 500 CNY increase in their total fees due to the integration. The presence of three stars denotes statistical significance at the 1% level for this result.

The integration policy effects on hospitalisation total fees depend on the pre-integration fees of patients. As pre-integration fees increased, the policy had greater effects. In Hospital A, all patients experienced fee increases, which provided evidence to support that the integration policy released the previously unmet medical needs of patients. Similarly, in Hospitals B and D, the integration also had significant effects, but only for patients with medium and high total fees. These results also support that the unmet medical needs of these subgroup patients were

released. However, some estimated coefficients in Hospital D were not significant. This could be attributed to the smaller sample size in Hospital D, resulting in larger standard errors and insignificant p-values[61]. Another possible reason, as discussed in Chapter 7, is that patients with very high and very low fees may not be significantly affected by the policy due to the adjustment of deductibles and ceilings, while patients with medium fees may be significantly affected by the policy. Additionally, low-fee patients in Hospital B experienced fee reductions. One possible reason is that the integration policy increases the deductible. In this case, if the treatment remains the same, the total fees also remain the same, but the out-of-pocket fees increase. Some patients with less severe conditions may therefore choose to reduce their treatment duration or dose to avoid higher out-of-pocket fees. Moreover, most of the results of Hospital C were not significant. As mentioned in subsection 9.4.1, it could be due to the policy effect reversal caused by unknown reasons.

Table 9-14 Summary of quantile effect analyses results on hospitalisation total fees

Hospitals	Statistics	Quantiles						
		5%	10%	25%	50%	75%	90%	95%
A	Coef.	209 **	229 ***	262 ***	500 ***	1,025 ***	2,407 ***	4,096 ***
	S.E.	-93	-65	-84	-138	-265	-498	-888
B	Coef.	-249 ***	-231 ***	-2	317 ***	902 ***	1,661 ***	2,678 ***
	S.E.	-62	-48	-46	-65	-160	-382	-559
C	Coef.	160	-64	-12	-498	-1,223*	-4,121	-10,312 **
	S.E.	-262	-223	-315	-420	-679	-2,956	-5,095
D	Coef.	31	209	190*	297**	588*	1,289	2,017
	S.E.	-164	-174	-105	-140	-334	-1,162	-2,753

Note: Coef. -Coefficients; S.E.-Standard Error; *** p<0.01, ** p<0.05, * p<0.1

[61] Hospital A had a total sample size of 21,646, while Hospital D had a smaller sample size of 3,568. Looking at the sample size for different quantiles, Hospital A had 1,149 patients with fees exceeding 90% and 575 patients with fees exceeding 95%. In contrast, Hospital D had 163 patients with fees exceeding 90% and 82 patients with fees exceeding 95%.

9.5 Discussion and Conclusion

This chapter aims to examine the effect of integrating basic medical insurance for urban and rural residents on three key outcomes: outpatient total fees, hospitalisation total fees, and length of hospital stay (LOS) in six cities using Electronic Medical Records (EMR) data. To address potential biases stemming from unobservable factors, the multi-period difference-in-differences (DID) method is employed as the primary approach for identifying policy effects. Drawing on the previous empirical literature, the chapter incorporates several checks and tests, including pre-analysis data checks, pre-integration parallel trend tests, placebo tests, and combined propensity score matching (PSM) and DID analyses, all aimed at reinforcing the robustness of the DID estimates. Additionally, the chapter explores the underlying mechanisms of different policies by examining the dynamic effect and the quantile effect of the policies.

The primary findings from the DID and PSM-DID analyses in this chapter indicate that the implementation of integrated policies across different studied cities resulted in similar reductions in outpatient total fees. However, the effects on LOS and hospitalisation total fees varied among the cities. These conclusions align with the qualitative analysis conducted in Chapter 7, which revealed that changes in outpatient reimbursement policies were more consistent across cities, while changes in hospitalisation reimbursement policies exhibited less uniformity. Two further in-depth analyses revealed the sources and mechanisms underlying the variations in policy effects. The dynamic effect analysis reaffirmed the cross-city consistency of the policy effects on outpatient fees and found three distinct dynamic effects on hospitalisation: the delayed display of policy effects that may be due to doctors' conservative behaviour; policy effect reversal which is temporarily unexplained in this study; and policy effect decay. Moreover, the findings from the quantile effect analysis on hospitalisation total fees demonstrated that the integration policy effects intensify as patients' pre-integration total fees (indicating the severity of illness) increase. This supports the notion that integration policies help address the unmet medical needs of patients who face economic constraints. Additionally, it was noted that policies in some cities may have raised deductibles, which may lead to reduced treatment for patients with less severe conditions, which aligns with the conclusion in Chapter 7.

The main advantage of this study compared with previous pilot city studies is that it combines the data comparison and check before analysis and the parallel trend test and placebo test after analysis to support the reliability of DID analysis results, which are lacking in previous studies. Furthermore, a notable strength lies in the inclusion of two in-depth analyses aimed at unravelling the mechanisms underlying policy effects. Mechanism analysis represents an important advancement in causal inference, with applications observed in various fields such as psychology and education. However, in the realm of insurance research, this approach has yet to be extensively explored.

However, certain limitations should be acknowledged. However, it is crucial to acknowledge the limitations of this study. The primary concern revolves around the quality of the data used for analysis. Due to fluctuations in sample size within certain hospitals across different years, both control and outcome variables exhibit complex trends, rendering the identification of policy effects more challenging. At the same time, although the rationality of the DID analysis is supported by macro indicators and cross-checks in the methodology section, due to the lack of individual-level socioeconomic variables to control confounding and matching, the satisfaction of the parallel trend assumption may still be challenging. Additionally, as the integrated policies in these cities are relatively new, the study is unable to observe the long-term effects of the policies.

Another limitation arises from the anonymisation requirement imposed on the data. While this chapter and Chapter 7 arrive at similar conclusions, it should be noted that the analyses in these two chapters are based on the policies implemented in different cities. Consequently, the findings may not be able to support each other, and also be directly applicable or generalisable to national policies and populations. Additionally, also due to the anonymisation, the policy effects investigated in this study could not be directly linked to specific changes in reimbursement terms and lists. As a result, although there are certain speculations regarding the mechanism of the policies based on the analysis results, they cannot be definitively proved within the scope of this study.

Chapter 10 Discussion

Over the past few decades, China has overhauled its social health insurance system. The latest insurance reform policy is the integration of basic medical insurance for urban and rural residents. Its main expectation is to reduce the inequality of healthcare between urban and rural residents and reduce the burden on patients. Previous studies have explored whether this reform works in different pilot regions. However, due to differences in data, analysis methods, and objectives, their results vary. In this thesis, I combined qualitative policy document analysis and quantitative econometric analysis at the national and city levels to comprehensively assess integration policies' effects, mechanisms, and their cross-city differences. More specifically, the objectives of the study are:

1. To summarise the different regional integration policies and to discuss the potential benefits and problems of integration;
2. To estimate the effects of integration on health service use and costs, at the national level;
3. To explore the differences in policy effects among different populations in socio-economic subgroups, at the national level;
4. To assess the impact of different integration policies on health service use and costs at the city-level, using ischaemic heart disease as an example;
5. To investigate the mechanisms of action of different integration policies at the city-level and how they differ, using ischaemic heart disease as an example.

Section 10.1 discusses and interprets the findings of the three studies included in this thesis in detail, by placing the results within the context of existing knowledge. This discussion includes the benefits of integration, potential problems, and ways in which these problems might be mitigated in the future. Section 10.2 then reflects on the strengths and limitations of this study. Combined with the research finds and limitations, suggestions in future research on how to further analyse the insurance reform and how to improve the limitations of this thesis are put forward. Section 10.3 is the conclusion of the full thesis, emphasizing its contributions.

10.1 Benefits, Issues, and Possible Improvement Measures

This section provides comprehensive discussions and interpretations of all results of the three studies in this thesis in order. In the discussion of each subsection, the benefits of integration revealed by this part of the analysis, the potential problems, and possible measures to deal with these problems in the future are included.

10.1.1 The Implementation of Integration

The qualitative analysis presented in Chapter 7 provided an overview of the timeline for implementing integration. The process began with the introduction of 2 integrated insurance schemes in 2007, followed by 9 years of various pilot integrations across 75 cities. By 2016, the integration was being nationally promoted. Subsequently, over half of the cities completed their integration in 2017 and 2018, leading to the full completion of integration in 337 prefecture-level administrative regions at the end of 2020. The reasons behind the 14-year duration of the entire integration reform can be attributed to the significant regional disparities in economic and medical resources, as evident from the comparison of macroeconomic and medical information in Chapter 3. This explanation is further supported by the timing of policy implementation in provincial capital cities in Chapter 7. The early pilot areas are primarily economically developed regions. For example, Chongqing and Chengdu are the economic centre cities of western China, Tianjin is the foreign trade and economic centre of northern China, and Hangzhou is the economic centre of eastern China. And many of the cities that implemented integration later, such as Hefei, Haikou, and Lhasa, are located in provinces with underdeveloped economies and relatively limited medical resources (Anhui, Hainan, Xinjiang provinces) as highlighted in Chapter 3. The reason is that the integration requires increased investment in personnel, information systems, and other management aspects while increasing subsidies to residents is sometimes also unavoidable. Regions with better economic conditions generally enjoy stronger local government finances, resulting in less fiscal pressure to implement reforms.

Combining the 4 case studies in Chapter 7 with the wider literature, two main benefits of integration for residents can be drawn: 1. Improvement of reimbursement terms; 2. Expansion of reimbursement scope.

The integration policies of outpatients were similar across cities, but the integration policies of hospitalisation were quite different. The outpatient reimbursement terms in all four case cities improved greatly. This improvement can lead to more outpatient visits by reducing self-paid expenses for patients. In Tianjin, both urban and rural residents experienced notable changes, with urban residents transitioning from no benefits to a reimbursement ceiling of 3,000 CNY per year, and rural residents seeing an increase from an annual subsidy of 20 CNY to the 3,000 CNY reimbursement ceiling. In Beijing, the reimbursement ceiling for urban and rural residents rose to 4,000 CNY from their previous values of 2,000 CNY and 3,000 CNY, respectively. The reimbursement rate rose slightly and the deductible for urban residents decreased. The situation in Shanghai is different in that the benefits for urban residents have changed even more, with a significant increase in the outpatient reimbursement rate, while the deductible has dropped from 1,000 CNY to 500 CNY.

Hospitalisation reimbursement terms also improved for some people, but the overall effect was likely to be small. One reason is that the potential impacts of integration on rural and urban residents may offset each other. For example, due to the implementation of 3 levels of optional insurance in Tianjin, no matter which level the rural residents choose, their treatment will be improved compared to before the integration, but if the urban residents choose a lower level of insurance, the reimbursement terms will become worse than before the integration. In Shanghai, except for the ceiling, the reimbursement rates of rural residents increased but of urban residents did not. In addition, such as in Beijing, although the reimbursement ceiling was raised to 250,000 CNY for both urban and rural residents, it was difficult to say whether rural or urban residents benefitted more from the integration due to the complexity of reimbursement calculations before the integration of rural residents. This may also lead to unclear overall policy effects.

The fiscal savings of the government may be one main reason why outpatient reimbursement terms improved significantly but hospitalisation reimbursement terms did not. The ceiling of outpatient reimbursement is only a few thousand CNY, while the ceiling of hospitalisation reimbursement is as high as hundreds of thousands CNY. Therefore, even if outpatient reimbursement doubles, it is only equivalent to an increase of a few per cent in hospitalisation reimbursement. For example, after the integration of resident medical insurance in Nanning (the provincial capital of Guangxi), the medical insurance reimbursement in 2019 totalled 3,796 million CNY, of which the hospitalisation reimbursement was 3,536 million CNY (accounting for 93.13% of the total), while the outpatient reimbursement was 260 million CNY (accounting for only 7.37% of the hospitalisation reimbursement amount)[62]. Therefore, prioritising funds towards outpatients can, on the one hand, keep the medical insurance fund stable; on the other hand, it can also make patients feel that they are benefiting more and improve their satisfaction with insurance integration.

The second main benefit of integration can be seen in the expanded scope of reimbursement. The list of designated institutions and the reimbursement catalogue are both greatly expanded after integration. As can be seen from Chapter 7, the providers for rural residents who were eligible for reimbursement before the integration were mostly located in counties, towns, and villages, which are mostly township health centres and village clinics, while the providers for urban residents were hospitals. Rural residents, like those in Tianjin, are entitled to 40-60% reimbursement at township medical institutions, but only 10-20% if they go to more advanced out-of-county institutions before the integration. In China, there is a huge gap between the standard of care in clinics and high-level hospitals: poorer facilities, lower skill levels and a large number of village doctors who are not licensed regular doctors. After the integration, in most cities, the former URBMI- and NRCMS-designated medical institutions, are all included in the URRBMI-designated institutions of the agreement management scope. For example, there were 833 designated medical institutions in Beijing before the integration, which increased to more than 3,000 after the integration; the number of designated medical institutions

[62] http://ybj.nanning.gov.cn/xxgk_164/sjfb/t4380260.html

for rural residents in Tianjin also expanded from about 30 to more than 1,400[63,64]. Moreover, although the changes in the reimbursement catalogue were not covered in this thesis's case study, it was also found during the review of the literature that the integration has also led to a huge expansion of the catalogue's content. For example, in Shandong, Guangdong and Ningxia, the types of drugs reimbursable to rural residents have expanded from 1,100, 1,083 and 918 to 2,400, 2,450 and 2,100 respectively after the merger of urban and rural medical insurance, all more than doubling[65].

The above-mentioned benefits are all expected by the reform, but at the same time, there are several potential problems in the integration policy: 1. high premiums; 2. adverse selection; 3. patient expectations.

Firstly, in terms of premiums, this poses a problem: rapidly increasing premiums may weaken rural residents' willingness to participate in insurance and their satisfaction with it. In this study, the out-of-pocket premiums for rural residents in Tianjin and Beijing increased from 30 CNY before integration to minimums of 60 CNY (doubled) and 180 CNY (sixfold) respectively after integration, and from a minimum of 270 CNY to 680 CNY (2.5 times) after integration in Shanghai. This is in line with the findings of other regional studies: the out-of-pocket premiums for rural residents in Zhongmou County were 120 CNY in 2016, 180 CNY in 2017 (after integration), 220 CNY in 2018, and 250 CNY in 2019, a cumulative increase of 108% in four years; the out-of-pocket premiums for rural residents in Xiangyin county had an increase of 50 CNY in 2012 to 180 CNY in 2018, an increase of over 300% in six years; Shijiazhuang rural residents' out-of-pocket premiums increased tenfold in 10 years (Hao, 2019; Peng, 2019; Wu, 2020). Since resident insurance is registered on a household basis, based on a rural family of 4-5 people, the annual premium for a family may exceed 1,000 CNY in some rural areas. However, the annual cost of healthcare for rural residents in 2018 was 1,240 CNY (see Appendix of Chapter 3). It is clear that, on average, the difference in health expenditure between those who

63 http://health.people.com.cn/n1/2017/1214/c14739-29706396.html

64 https://www.chinacourt.org/article/detail/2016/08/id/2059119.shtml

65 https://www.chinacourt.org/article/detail/2016/08/id/2059119.shtml

are insured and those who are not, is not significant, so the willingness of residents to participate is not high. This was confirmed by resident interviews in other regional studies: “The premium is 180 CNY per person this year, so with a family of six, I would have to pay more than 1,000 CNY, which is my family's income for half a month. In the past year or two, there have not been many illnesses in my family, so this amount is enough for medical treatment when I get sick” (Hao, 2019). Although government subsidies also increased in tandem (but less than the increase in premiums paid by residents), residents see more of an increase in their burden as the subsidies go directly into the insurance fund pool and are not directly perceived by the insured.

Secondly, the multi-level design of premiums and benefits may lead to adverse selection. Not only Chengdu and Tianjin, but also Chongqing, Hangzhou, Jinan, Harbin, Lasa, and many other cities have chosen to adopt multi-level insurance as shown in Chapter 7. However, from previous research, multi-level insurance appears that may introduce adverse selection, whereby healthy people choose the cheaper level of premium and unhealthy people choose the more expensive level (Sun *et al.*, 2014; Chen *et al.*, 2015; Shang *et al.*, 2016). In the end, the proportion of high-health-risk groups among the insured population is increasing, the risk-sharing function of insurance is getting worse, and the sustainability of insurance funds is getting worse.

Thirdly, due to the constraints of the reimbursement catalogue, deductible, and ceiling, the actual reimbursement may not meet the patient's expectations. As mentioned in the literature review, many insured residents found that the reimbursement rate they presupposed was different from the actual reimbursement rate when making claims. A study shows 43% of the residents surveyed were dissatisfied with the reimbursement terms (Peng, 2019); interviews in another study also showed the dissatisfaction of the insured: "The total cost of the hospitalisation was 34,764 CNY and only 11,892 CNY was reimbursed, after being discharged from the hospital, I bought a lot of medicines from the pharmacy in the city, which were not reimbursed and could not be bought in my county" "The reimbursement rate for this hospitalisation was said to be 60%, but according to the actual situation, only 34% was

reimbursed" (Hao, 2019). Similarly, it was also mentioned in a study that during the insurance integration process in Changsha, due to the adjustment of the reimbursement catalogue, some commonly used drugs in rural areas that were reimbursed by the NRCMS in the past were no longer reimbursed in the integrated residents' insurance, which caused great dissatisfaction among rural residents (Zhu, 2019).

In future reforms, in order to alleviate or solve the three potential problems mentioned above, possible measures are: 1. Increase government investment in the health sector; 2. Continue to deepen insurance integration to a single-level design; 3. Strengthen the localization of dynamic adjustments to drug reimbursement catalogues.

As mentioned in Chapter 3, China's health expenditure continues to grow but remains low as a proportion of GDP compared to developed countries. Therefore, there is a need to continue to increase investment in the health field. Increasing the government's subsidy for residents to participate in insurance, that is, premium subsidies, on the one hand, can reduce the burden on residents, especially rural residents, and increase residents' willingness to participate in insurance. On the other hand, insurance funds can be expanded so that insurance offers better reimbursement terms or at least avoid reducing existing reimbursement items to reduce resident dissatisfaction. Moreover, investments can also be used to strengthen the promotion of insurance, as previous research (Sun, Zhao and Yin, 2017) has shown that knowledge of insurance policies is positively correlated with willingness to enrol.

For those cities that implement multi-level insurance, it is a good reform direction to continue to deepen the integration, and finally achieve a single-level of insurance. Su *et al.* (2019) also mentioned that the 'one system, two standards/levels' model should only be used as a transitional measure in the pilot areas and should achieve "one system, one standards/levels". In fact, this has already happened or is happening in some cities. For example, Chengdu implemented 3 levels of insurance when it first integrated in 2009, and then implemented 2 levels of insurance in 2012; Yinchuan implemented 3 levels of insurance when it first integrated

in 2012 and transitioned to a single insurance design by 2020[66,67]. In addition to solving the adverse selection problem, as stated in the announcement of the Yinchuan Municipal Government, changing from multi-level to single-level can also solve "the imbalance of urban and rural benefits, the incompatibility between the level of insurance payment and the reform of the household registration system, the insufficient fund revenue, fragmentation of policies" and many other issues. Single-level insurance can reduce income-induced healthcare inequality. Yi and Zhan (2018) found that compared to multi-level design, single-level design significantly improves the insured's ability to pay for health services, making their hospitalisation less constrained by personal income.

Finally, the adjustment of the drug list should take into account the local disease spectrum. In 2020, China's National Healthcare Security Administration released the "Interim Measures for the Administration of Medicines in Basic Medical Insurance" and the "Interim Measures for the Administration of Medical Consumables in Basic Medical Insurance", establishing a dynamic adjustment mechanism for the reimbursement catalogue. Based on these documents, evidence from health technology assessment (HTA) is increasingly important in determining which and how much drugs and consumables can be reimbursed in China. However, the current drug adjustment is based on the national disease situation and does not take into account the epidemiological differences of diseases across different regions. As a result, the local common medicines mentioned above are not allowed to be reimbursed after the insurance integration. Therefore, further investment should be made in the field of HTA, and the drug access process of medical insurance should be optimised in combination with epidemiology in different regions. More practical drugs and consumables can be reimbursed, and medical efficiency and patient satisfaction can be improved.

---

66 http://www.cdwh.gov.cn/wuhou/c106229/2011-12/01/content_a9b4df3ad4dc4d92b3e7329b413729ff.shtml

67 https://www.nx.gov.cn/zwxx_11337/sxdt/202001/t20200103_1914434.html

10.1.2 Nationwide Policy Impact and Subgroup Heterogeneity

Overall, the statistical analysis results in Chapter 8 show that nationwide, the integration boosts outpatient visits and reduces outpatient and hospitalisation self-paid costs. These are the expected outcomes of the reform, namely improving access to health services and reducing the financial burden on patients. The research results of this thesis may be not comparable with the 9 policy evaluation studies in the literature review, whether the results are consistent or inconsistent. This is because the data used in this study is larger than that used by previous studies, and includes more newly integrated regions. Therefore, the results of previous studies can actually be equivalent to the results of a certain regional subgroup of this study, while the results of this study should be considered more "representative" at the national level.

The subgroup analysis results in Chapter 8 show that the policy is pro-rural. This means that rural residents benefit more from integration reforms than urban residents. This is in line with policy expectations to reduce urban-rural healthcare inequality. However, this result is contradictory to the case study in Chapter 7, where urban residents may benefit more than rural residents. For example, in Beijing, deductibles fell for urban residents but remained the same for rural residents, while urban residents' reimbursement ceilings were also raised more than that of rural residents. One possible reason is that the case studies are not well representative of the national situation since they are all provincial capital cities. And possibly in other non-provincial capital cities, integration may be more beneficial to rural residents.

In addition, these subgroup analysis results also show that the policy is pro-high-income in outpatient visits and hospitalisation OOP costs. This means that high-income residents benefit more from integration reforms than low-income residents and this may not be what the reform anticipated. The interaction subgroup results of urban and rural areas and income show that the impact of income on policy effects is greater than of that the impact of urban and rural areas. As a result, for outpatient visits, it is the urban high-income residents who are most affected by the policy (increased by 5.35%), followed by the rural high-income residents (increased by 2.32%), and then the rural low-income (increased by 1.9%). The urban poor are affected by the opposite policy effect (decreased by 2.39%) to the other three subgroups. As a result, Because

the urban poor and the rich are affected oppositely, the overall impact of the policy on urban residents is not significant; while the poor and the rich in rural areas are affected in the same direction, which eventually makes the overall impact of the policy on rural residents more significant than that of urban residents.

In summary, these results demonstrate that the pro-rural nature of the integration policy reduces healthcare inequalities between urban and rural areas, but its pro-high-income nature further widens healthcare inequalities between income groups. This conclusion is comparable with other studies. Most previous studies have demonstrated inequalities in health care use/cost between urban and rural residents and between residents of different incomes. For example. Ye and Liu (2020) divided rural residents into 5 income quintiles and found that without adjusting for confounding factors, although the 20% lowest income population had the highest rate of medical reimbursement at 25% (22.3% for the lower 20% income, 21.4% for the middle 20% income, 20.1% for the higher 20% income, and 18.2% for the highest 20% income), the financial burden of disease was highest (27.4%) for the 20% lowest group after taking into account income (9.1% for 20% lower income, 3.4% for middle 20% income, 3% for higher 20% income, 1% for highest 20% income). After adjusting for confounding factors, the financial burden of disease would be 132.3% higher in the lowest income group than in the highest income group, 25.1% higher in the lower income group than in the highest income group, and 14.1% higher in the middle-income group than in the highest income group.

Two potential reform directions may help to address these problems. One direction is to increase the geographic scope of insurance integration. There are great differences in insurance benefits in different cities. For example, in 2018, the ceiling of reimbursement for hospitalisation in Beijing City was 250,000 CNY, while the ceiling of reimbursement in Zhuzhou City (in Hunan province) was only 150,000 CNY[68]. Therefore, in addition to converting multi-level to single-level insurance, raising the level of integration may be a good choice: from urban-rural integration within a city to multi-city integration, and then develop

[68] http://zhuzhou.bendibao.com/live/2021728/2779.shtm

into province-wide insurance integration, and ultimately achieve unified national health insurance. This may reduce healthcare inequalities due to cross-regional economic and income inequalities. On the other hand, while the integration of insurance has narrowed the gap between urban and rural areas, it may also bring pressure on fund revenue and expenditure balance as discussed in Chapter 7. Against the social backdrop of rapidly rising medical costs and an accelerating ageing population, the increase in premiums needs to be accompanied by a sustained increase in financial subsidies. And integrating the residents' and employees' insurance is also a meaningful exploration. Sharing the fund pools of both can effectively increase the stability and sustainability of the insurance fund.

10.1.3 Mechanism of Different Policies

The analysis results of IHD patients in six hospitals in different cities in Chapter 9 show that after the nationwide promotion, the integration policies across cities have shown a similar fee-decreasing effect on outpatient services, but various effects on the length of stay and hospitalisation fees. This finding re-emphasises the conclusion of the qualitative policy document analysis in Chapter 7 that changes in outpatient reimbursement terms are more similar across cities, while changes in hospitalisation reimbursement terms have less commonality. In some hospitals, the integration increased the length of stay and total hospitalisation costs for IHD patients. In combination with the possibly better hospitalisation reimbursement terms after integration in the case analysis in Chapter 7 and the reduction of hospitalisation self-paid costs in Chapter 8, a reasonable explanation is that integration releases the medical needs of patients previously constrained by economic conditions by reducing patients' self-paid costs.

In the analysis of dynamic effects in Chapter 9, it is found that the policy takes effect immediately and stably in outpatient fees, but in LOS and hospitalisation fees, there are 3 different situations: delay increase, decay, and reversal. These dynamic effects are reasonable. In China's policy practice, there is usually a time lag ranging from 0 to 1 year between the release of policies and their effectiveness (Cao, 2020). As Meyer-Sahling (2007) explained, this time difference reflects the strategic choices made by decision-makers based on political,

management, and other considerations. The long-term feedback path created by policies through their formulation and implementation promotes the re-understanding and re-adjustment of decision-makers, implementers, and recipients, thus forming the continuity and nesting of policies over time (Pierson, 2000; Mettler and Sorelle, 2018). Therefore, in this thesis, the diverse dynamic effects are the results of the interaction between decision-makers (government staff, medical insurance administrators, and hospital administrators), implementers (hospital administrators and doctors), and recipients (patients). Doctors may change the treatment plan for some patients in the first year of policy implementation, and make adjustments in the second year after receiving feedback from patients. However, exactly which doctors' decisions were changed by the integration policy, and what changes were made to patients' treatment plans, is beyond the scope of this study since they are treatment pattern studies, but it is worth exploring in the future.

Another important part of the analyses in Chapter 9 is quantile regression on hospitalisation fees. As discussed in Chapter 5 Methodology, within a hospital, patients with the same disease are subject to consistent treatment guidelines and pricing for health services and drugs. Therefore, varying quantiles of hospitalisation fees can be used to indicate different levels of disease severity, enabling the reflection of treatment effects within distinct patient subgroups based on their health status. Patients with low hospitalisation total fees are considered to be relatively lee serious condition and the with high total fees reflects more serious condition. Patients with severe conditions may not be able to get the treatment they need due to economic constraints (medical need is suppressed), so when the reimbursement policy becomes better, their self-paid amount is reduced. Then, they are likely to adjust their treatment and receive more necessary treatment. Therefore, their total fees increased significantly (but the self-paid amount after the policy may be lower or equal to the amount before the policy). Many studies have revealed that economic status is highly positively correlated with the use of health services, and poor people with the same health status receive fewer health services than rich people, especially in China, income has been shown to be a key factor influencing healthcare use (Xie, 2009; Wang and Wang, 2010). In the 3 hospitals affected significantly by the integration policies, the size of policy effects was closely correlated with the increase in patient fee

quantiles. In 2 hospitals, the integration consistently led to increased fees for all patients, with larger effects observed as the patients' fee quantile increased. In another hospital, the integration initially exhibited a gradually weakened effect of fee-decreasing as the patients' fee quantile increased (the size of policy effects ranged from -249 CNY to -2 CNY from the 5% quantile to the 25% quantile), followed by a gradually enhanced effect of fee-increasing as the patients' fee quantile increased (the size of policy effects ranged from 317 CNY to 2,678 CNY from the 50% quantile to the 95% quantile). If all patients are adequately treated, patients in the same hospital should receive similar (or proportional[69]) policy effects that do not scale with disease severity. The above-estimated policy effects that increase with severity fully demonstrate that the more severe patients are, the more restrained their medical needs is due to economic constraints, and the integration has alleviated such need constraint.

The above results explain and highlight the effect of integration and the source of the difference, but due to anonymity, they cannot be linked to the specific reimbursable list of reimbursement terms and thus cannot provide new policy suggestions. However, it can be seen that the reimbursement catalogue may also played an important role in policy work, which again emphasizes the importance of increasing investment in the medical field, especially HTA, to improve access to medicines.

In summary, this thesis contributes to the understanding of China's health insurance reform and healthcare inequality through five key findings. Firstly, by combining qualitative and quantitative analysis, this thesis establishes that the integration policy terms and effects on outpatient services were generally consistent across different cities, whereas there was less uniformity in terms of hospitalisation. Secondly, through nationwide quantitative analysis, the thesis provides evidence that the integrated policy effectively enhanced the accessibility of outpatient services and alleviated the financial burden for patients in both outpatient and hospitalized settings. Thirdly, by examining policy effects on urban-rural and income

[69] The proportion of policy effects in pre-policy fees should be similar for different quantile patients. However, taking hospital A as an example, from low-quantile fees patients to high-quantile fees patients, this proportion increases from 5% to more than 13%.

subgroups at the national level, the thesis verifies that integration helped alleviate medical inequality between urban and rural areas, but simultaneously exacerbated inequality among different income groups. Fourthly, through quantitative analysis at the city level, the thesis elucidates that disparities in policy effects can be attributed to the complex interplay among policymakers, implementers, and recipients, as well as the variety in the adjustment of reimbursement terms and lists across cities. Finally, through the quantitative analysis at the city level, this thesis demonstrates that integration relieves the unmet medical needs of patients due to economic constraints.

10.2 Strengths, Limitations, and Future Research

This section provides reflections on the strengths and limitations of this study. Subsection 10.2.1 highlights the strengths of this paper over past studies, such as the inclusion of multiple analyses to explore policy mechanisms. However, limitations of this study are also acknowledged in subsection 10.2.2, such as the inability to directly link policy impacts to specific reimbursement terms and lists due to data anonymization requirements. Combining the considerations about the limitations and the previous discussion of the results in Section 10.1, Subsection 10.2.3 proposes how to further analyse insurance or other health reforms in future research and how to improve the limitations of this paper.

10.2.1 Strengths

The advantages of this thesis are mainly in three aspects: the use of mixed methods in the research design, the advantages of the latest CHARLS and EMR data, and more advanced statistical models and analysis frameworks.

The first strength of this thesis lies in the use of mixed methods. Unlike previous studies in literature reviews that solely relied on qualitative methods, such as field surveys, or quantitative methods, such as statistical analysis, this thesis combines qualitative document analysis and quantitative statistical analysis. This comprehensive approach encompasses various aspects, such as the analysis of the integration implementation timeline, comparison of policy terms pre- and post-integration, estimation of average policy effect at the national level, and comparison

of policy effects within a single hospital across cities. The mixed method is widely used in disciplines like sociology, political science, economics, and health sciences due to its inherent advantages (Tashakori and Creswell, 2007; Migiro and Magangi, 2011; Regnault, Willgoss and Barbic, 2017). The mixed method enhances the research rigour by combining high representativeness through large sample sizes for quantitative methods and high confidence for qualitative methods. By employing this approach, a study effectively achieves a more comprehensive understanding of research goals and questions than either method alone (Cojocaru, 2009; Tariq and Woodman, 2013; Yardley and Bishop, 2015). This thesis proved the first of the 5 key findings/conclusions mentioned in the previous section by combining the analyses of Chapter 7 and Chapter 9.

The second strength lies in the use of the latest CHARLS data and EMR data. CHARLS data have been used in many previous studies as shown in Chapter 6. However, these studies have limited coverage of integrated regions due to the gradual and slow implementation of the integration policy, which was only in its pilot phase before 2016. Specifically, CHARLS data encompass a total of 126 cities across 28 provinces in China. Ma, Zhao and Gu (2016) using 2012 data includes only 13 integrated cities; Liu (2017), Li (2017), and Liu and Lin (2018) using 2013 data cover 15 integrated cities; Ma and Li (2018) and Su *et al.* (2019) using 2015 data include 29 integrated cities. This thesis used the latest available data from 2018 CHARLS, covering 109 integrated cities. Therefore, as mentioned in section 10.1.2, the results of previous studies can be considered equivalent to findings from specific regional subgroups within this study, highlighting the importance of considering the results of this study as more "representative". Furthermore, as mentioned in Chapter 9, using EMR data can provide models with higher statistical power than survey data due to larger samples, avoidance of recall bias and fewer errors, and greater population representation (Sulieman *et al.*, 2022).

The third advantage lies in the statistical methodology and empirical analysis framework, including using the extended DID method and providing robustness testing, and introducing dynamic effect analysis and quantile effect analysis. Since this thesis uses multi-period data (panel data and repeated cross-sectional data) for policy evaluation, it adopts staggered DID

and multi-stage DID as the main causal identification methods. Compared with the two-stage DID used in past studies, staggered and multi-stage DID can provide more accurate estimates by covering more information and controlling for all time-invariant biases. At the same time, due to the use of two-stage DID, previous studies cannot statistically support the rationale for using the DID method. However, this thesis combines cross-database verification, national statistical data comparison verification, parallel trend test before policy, and placebo test to more strongly support the rationality and reliability of the DID model, as well as uses the DDD and PSM+DID to support the robustness of the results. More importantly, this thesis introduces the mechanism analysis, such as the quantile effect analysis, in examining hospitalisation fees and LOS within the context of integration. Mechanism analysis is a significant advancement in the field of causal inference, and its applications have been observed in diverse disciplines like psychology and education, as discussed in Chapter 5. However, its applications in insurance or health research remain relatively underexplored.

10.2.2 Limitations

The main limitations of this thesis include: 1. Health service supply and the interaction of different policies; 2. Secondary, anonymity, and extrapolation of data; 3. Defects in DID method.

Firstly, the present study only addresses the demand side of healthcare. In economics, both supply and demand sides determine the actual demand for commodities (health services). The health insurance reform is only demand-side, which may lead to policy effects that are not significant in a short time period until the supply of health services changes in the future. In addition, China's healthcare reform involves many aspects, which means that many different reforms are being carried out at the same time, not only the health insurance integration but also payment method reforms (payment method transfer from Fee-for-Service to Diagnosis Related Group), public hospitals reform (abolish medicine makeup policy, improve the price of services that reflect the value of staff techniques, such as consultant, surgery, nursing, etc.), drug access reform (introduce health technology assessment in the decision making of essential drug list) and these separate effects are difficult to disentangle in this type of analysis.

The second limitation pertains to the secondary nature, anonymity, and extrapolation of data. Using secondary data, researchers may encounter disadvantages such as potential misunderstandings due to the lack of involvement in data collection, code, and modification processes, ultimately impacting study quality (Flintermann, 2014; Johnston, 2014; Olabode, Olateju and Bakare, 2019). To avoid misinterpretation of CHARLS data, I carefully reviewed the survey process description and coding instructions provided by the CHARLS survey team. However, the initial encoding by the hospital and subsequent re-encoding by the data provider of electronic medical records (EMR) data remain confidential. This leads to the non-aggregatable characteristic of EHR data mentioned in Chapter 9. Another disadvantage of secondary data is its lack of specificity to the later researcher's research question, potentially resulting in incomplete or missing information on the data and variables of interest (Doolan and Froelicher, 2009; Smith *et al.*, 2011; Johnston, 2014). Insufficient data and variables posed the greatest challenge in the statistical analysis in Chapter 9. Socioeconomic factors, crucial in determining whether residents and non-residents exhibit parallel trends in healthcare use and medical fees, were not included in the EMR data. Therefore, in Chapter 9, I relied on both national statistics data and CHARLS cross-validation to support the parallel trend assumption. Anonymity also presents other issues. In EMR data, the anonymisation of patients protects their privacy, which prevents follow-up analysis. However, because my study was performed on a single visit, the analysis was not affected. But on the other hand, the anonymity of the data source (hospital name and city name) makes it impossible to link the analysis results with the specific integrated reimbursement terms and reimbursement catalogue, so it is impossible to give more precise and deeper interpretations of the findings. In addition, the data used in this thesis are only representative of two specific populations: middle-aged and elderly people and people with ischaemic heart disease. It raises concerns about the extrapolation of the results. Attendance behaviour and treatment spending may differ substantially between the younger cohort and the middle-aged and older cohorts. And there may be systematic differences in the impact of policies on the ischaemic heart disease population and other disease populations.

Thirdly, recent advancements in statistical techniques have highlighted inherent flaws in the two-way fixed effect (staggered-DID) model (Callaway and Sant'Anna, 2021; Goodman-

Bacon, 2021). When the policy effects change over time, the estimated outcome using the staggered DID method may not accurately represent the true Average Treatment Effect (ATE). The reason is that the staggered-DID model provides a weighted average of the ATE estimates obtained for multiple standards DID estimates, but the weight assigned to each estimate can be negative. This means that even if the policy itself has a positive effect on the treatment groups at all time points, the estimation coefficient of the staggered DID may still be negative. In other words, under the situation of heterogeneity treatment effects, the single coefficient estimation result of the staggered DID model loses credibility.

10.2.3 Future Research

Based on the findings and limitations of this thesis, future research directions mainly include expanding data sources, improving statistical methods, and analysing the interactive effects of different insurance reforms

Firstly, an important direction is to obtain greater data permission through government, academia, and business cooperation. On the one hand, if survey data, hospital data, and insurance data can be combined, then the shortcomings of using either data alone can be overcome. As Sulieman *et al.* (2022) say, integrating self-reported data from survey and EMR data is an emerging field of research that will address issues of aggregation of multiple data types. Self-report data can identify and complement missing information in EMR data, and negative answers provided by self-report can support the absence of a disease, which is a shortcoming of EMR data: absence of diagnosis does not mean the absence of disease (Pendergrass and Crawford, 2018). These efforts will create accurate cohorts for health and medical research. On the other hand, links to data and policy documents can be obtained without the constraints of anonymity. As mentioned in Section 10.1.3, by analysing the details of reimbursement terms, reimbursement catalogues, and changes in hospital implementation rules in corresponding cities, and also comparing patients' medication (prescription) differences after receiving data permission, researchers can better answer how the integration or other insurance reforms change the behaviours of hospital managers, doctors, and patients.

The second is the improvement of the causal identification method. For the DID approach alone, developments in two directions have emerged: heterogeneous treatment effects and nonparallel trends. After Goodman-Bacon (2021) discussed the heterogeneity issue of two-way fixed effects, an important development direction of DID is the heterogeneity of policy effects. For example, Callaway and Sant'Anna (2021) proposed different aggregation schemes that can be used to highlight the heterogeneity of treatment effect across different dimensions and to summarize the overall effect of participating treatments ("csdid" in Stata); Borusyak, Jaravel and Spiess (2021) proposed an efficient estimator when the heterogeneity of treatment effects is not restricted ("did_imputation" in Stata); Chaisemartin, D'Haultfoeuille and Guyonvarch (2023) proposed a DID design with multiple groups and periods, which can calculate dynamic policy effects ("did_multiplegt" in Stata). These methods can all be used in future research to obtain more precise policy effect estimates.

In addition, although the pre-policy parallel trend test is the most intuitive and commonly used method to test the rationality of DID, some research suggests that this test may have low power (Bilinski and Hatfield, 2020; Roth, 2022). Therefore, the latest study, by Rambachan and Roth, (2023), shows that instead of requiring DID to satisfy parallel trends, it regulates the impact of non-parallel trends on estimation. The authors propose two approaches to arrive at efficient causal inferences. The first method is based on the empirical Bayes method, which calibrates the distribution of the treatment effect by correcting for the difference in trend before treatment. The second method is based on a pre-test of the size of parallel trend violations before treatment, which can be used to construct confidence intervals that are valid in the case of treatment violations of parallel trends. Using them in future research can more strongly support the reliability of DID estimation results.

Finally, China's healthcare system and insurance system are still under constant reforms. These provide researchers with ample opportunity to explore what policy works and how to improve access to care, improve health, reduce patient burden, and alleviate inequalities. A notable example is the implementation of Diagnosis-Related Groups (DRGs) reform in China. The DRGs are "a patient classification scheme which provides a means of relating the type of

patients a hospital treats (i.e., its case mix) to the costs incurred by the hospital" (CMS, 2019). China started to implement DRG on a pilot basis in 30 cities at the end of 2019[70], and it extended to 71 cities in 2020. Official documents issued at the end of 2021 indicate that nationwide promotion begins in 2022 and will be completed by the end of 2025[71]. One of the goals of DRG reform is to finally realise that 70% of the health insurance funds are used through DRG instead of "fee-for-service", that is, hospitals are reimbursed a fixed DRG-specific amount for each patient treated. This change in payment method is worth studying in the future because it can greatly change the treatment path and treatment/reimbursement fees of patients, such as "After multiple consultants, doctors can easily determine what kind of patients are more likely to save insurance reimbursement funds when they treat patients. This may cause doctors to have to choose between the patient's treatment effect and the hospital's benefit", "Some critical care doctors claim that treating every patient is a loss of money (the actual expenses of treatment exceed the amount that can be reimbursed by DRG)"[72]. Moreover, what kind of interaction will there be between the dynamic adjustment of the drug reimbursement catalogue and the deeper integration of insurance in the future? What will be the outcome of the integration of resident and employee insurance? These are also possible research directions.

10.3 Conclusion

This thesis provided a comprehensive assessment of China's recently completed reform, the integration of urban and rural basic medical insurance, by combining quantitative and qualitative analyses. The five main findings and contributions of this thesis are:

1. The integration policies and effects of outpatients were similar across cities, but the integration policies and effects of hospitalisation were quite different;

2. From the national level, the integration policies promoted outpatient visits and reduced outpatient and hospitalisation costs;

3. Urban-rural healthcare inequality was reduced through integration, but inequality among different income groups was expanded;

[70] https://www.gov.cn/zhengce/zhengceku/2019-11/18/content_5562261.htm

[71] http://www.gov.cn/zhengce/zhengceku/2021-11/28/content_5653858.htm

[72] https://www.thepaper.cn/newsDetail_forward_23583451

4. Differences in the impact of integration in different cities on hospitalisation may be due to differences in doctor behaviours, reimbursement terms, and reimbursement catalogues;

5. For hospitalisation, integration did release patients' pent-up medical needs due to economic constraints.

This research has advantages over previous studies due to the use of mixed methods, statistical analysis at the national and city levels, and more advanced and comprehensive causal identification methods and analysis framework. But at the same time, there are many limitations such as not considering the medical supply side and the interaction of multiple policies, the disadvantages of data types, and the inherent defects of DID. Future research can focus on expanding data sources, improving causal identification methods, and analysing different policies interaction.

Appendix A: Definition of Health System, Health Security System, and Social Health Insurance

In the definition of the WHO, "A health system includes the resources, actors, and institutions related to the financing, regulation, and provision of health actions. A health action is defined to be any set of activities whose primary intent is to improve or maintain health" (Murray and Frenk, 2000).

According to China's health care reform documents, the Chinese health system can be divided into four parts (Figure A-1): the Public Health System, the Health Service System, the Health Security System, and the Drug Supply System (The State Council of China, 2009). The public health system includes disease prevention and control, health education, maternal and child healthcare, mental health, emergency treatment, blood collection and supply, health supervision, family planning (birth control), and so on. The health services system includes county hospitals/township hospitals/village clinics in rural areas, and city hospitals/community health centres in urban areas. The drug supply system includes the national essential drug system (national drug list and production and supply of essential drugs) and medicine production and circulation system (market access and drug registration approval).

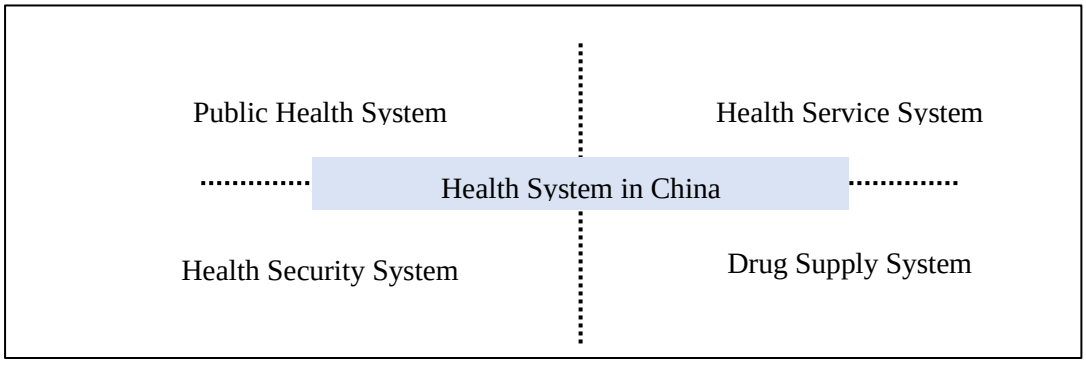

Figure A-1 The structure of the Chinese healthcare system

The health security system encompasses various government-implemented systems that provide compensation to individuals for income loss and medical expenses caused by diseases or other incidents like childbirth or disability (Changxue *et al.*, 2016). These systems can be categorised into different models worldwide: National Health Service (NHS) System (e.g., United Kingdom, Canada, Australia), Social Health Insurance System (e.g., Germany, Japan,

China), Compulsory Saving Health Security System (e.g., Singapore), Private Health Insurance and Public Medical Assistance (e.g., United States). The dominant models are the NHS model, the social insurance model, and the private model.

The NHS model, also known as the "Beveridgean System", ensures that all citizens have free access to healthcare services through general taxation, guaranteeing universal healthcare (Or *et al.*, 2010). This model is based on the principle that all citizens are entitled to socio-economic rights, irrespective of their social class, market status., or political rights. It considers it a state's responsibility to provide services to all its citizens, ensuring necessities such as health, education, food and shelter by offsetting some basic social risks such as disability, old age, unemployment, illness and maternity. In this respect, the model aims to promote the benefits of equality.

The social health insurance model, also referred to as the "Bismarckian System", is defined by the World Health Organisation (WHO) as a form of healthcare financing and management based on risk pooling (WHO, 2003; Or *et al.*, 2010). While its implementation may vary across countries, it shares common features, including individuals making regular contributions (premiums) to a health insurance fund based on their income rather than their disease risks, contributions from both employers and employees, government subsidies for low-income individuals, the existence of multiple schemes within a country, compulsory enrolment for some schemes, and so on (Jacobs and Goddard, 2000). Health insurance has been shown to distribute individual health risks, reduce the cost of healthcare services, improve accessibility, and ultimately contribute positively to individual health (Trujillo, Portillo and Vernon, 2005).

The private insurance system operates on a market-driven approach with significant involvement from private players. It believes that reducing the state's role in healthcare leads to a more efficient system, where funding and provision of services are managed by market leaders. This system involves various private health insurance contracts paid for by employers or individuals. Individuals voluntarily enter into agreements with insurance providers, and premium levels are based on individual health risk rather than income. Patients are responsible

for co-payments or deductibles to be reimbursed for all or part of their medical expenses. In the private healthcare system, health services are provided by hospitals and doctors as entrepreneurs. The system allows patients the freedom to choose their providers.

All three models face a complex policy challenge of balancing quality, equity and cost control in healthcare. Many developed countries experience escalating healthcare spending, posing a threat to the sustainability of their systems. Governments must strive to maintain quality and equity while controlling costs as healthcare remains a significant political concern. The NHS system prioritises equality and cost control but may face challenges in maintaining quality. The private health insurance system emphasises quality above the other values. The social insurance system prioritises quality and, to a lesser extent, equity, while cost control becomes a concern. The table below provides a rough comparison of these models.

Table A-1 Comparison of three health security system

Models	Quality	Equity	Cost
National Health Service	Lowest	Highest	Lowest
Social Health Insurance	Medium	Medium	Medium
Private Health Insurance	Highest	Lowest	Highest

In the past few years, China has established its multi-level health security system with social health insurance system as the main component, while also including commercial insurance, medical assistance, and supplementary medical insurance (Table A-2).

Table A-2 The structure of the Chinese health security system

Hierarchy	Module	Scheme	Sponsor
1st: Basic Medical Security	Main Body: Social Health Insurance	(1) Urban Employee Basic Medical Insurance (2) Urban and Rural Resident Basic Medical Insurance	Government
	Extension: Serious Disease Insurance	(3) Urban and Rural Resident Serious illness insurance (4) Subsidy for Employee's High Medical Expenses (5) Medical Allowance of Officeholder	
	Bottom	(6) Medical Assistance	
2nd: Enterprise Supplementary Medical Insurance	Supplementary	(7) Enterprise Supplementary Medical Insurance and Group Health Insurance	Employer
3rd: Commercial Health Insurance		(8) General Commercial Health Insurance (9) Tax-advantaged Commercial Health Insurance for Individuals	Individual/Family
4th: Charity & Mutual Aid		(10) Charitable Donations (11) Medical Mutual Aid	Society/Enterprise

Appendix B: the Opinions of the State Council on the Integration of the Basic Medical Insurance System for Urban and Rural Residents (in Chinese)

(Sources: http://www.gov.cn/zhengce/content/2016-01/12/content_10582.htm)

国务院关于整合城乡居民基本医疗保险制度的意见

国发〔2016〕3号

各省、自治区、直辖市人民政府，国务院各部委、各直属机构：

整合城镇居民基本医疗保险（以下简称城镇居民医保）和新型农村合作医疗（以下简称新农合）两项制度，建立统一的城乡居民基本医疗保险（以下简称城乡居民医保）制度，是推进医药卫生体制改革、实现城乡居民公平享有基本医疗保险权益、促进社会公平正义、增进人民福祉的重大举措，对促进城乡经济社会协调发展、全面建成小康社会具有重要意义。在总结城镇居民医保和新农合运行情况以及地方探索实践经验的基础上，现就整合建立城乡居民医保制度提出如下意见。

一、总体要求与基本原则

（一）总体要求。

以邓小平理论、“三个代表”重要思想、科学发展观为指导，认真贯彻党的十八大、十八届二中、三中、四中、五中全会和习近平总书记系列重要讲话精神，落实党中央、国务院关于深化医药卫生体制改革的要求，按照全覆盖、保基本、多层次、可持续的方针，加强统筹协调与顶层设计，遵循先易后难、循序渐进的原则，从完善政策入手，推进城镇居民医保和新农合制度整合，逐步在全国范围内建立起统一的城乡居民医保制度，推动保障更加公平、管理服务更加规范、医疗资源利用更加有效，促进全民医保体系持续健康发展。

（二）基本原则。

1.统筹规划、协调发展。要把城乡居民医保制度整合纳入全民医保体系发展和深化医改全局，统筹安排，合理规划，突出医保、医疗、医药三医联动，加强基本医保、大病保险、医疗救助、疾病应急救助、商业健康保险等衔接，强化制度的系统性、整体性、协同性。

2.立足基本、保障公平。要准确定位，科学设计，立足经济社会发展水平、城乡居民负担和基金承受能力，充分考虑并逐步缩小城乡差距、地区差异，保障城乡居民公平享有基本医保待遇，实现城乡居民医保制度可持续发展。

3.因地制宜、有序推进。要结合实际，全面分析研判，周密制订实施方案，加强整合前后的衔接，确保工作顺畅接续、有序过渡，确保群众基本医保待遇不受影响，确保医保基金安全和制度运行平稳。

4.创新机制、提升效能。要坚持管办分开，落实政府责任，完善管理运行机制，深入推进支付方式改革，提升医保资金使用效率和经办管理服务效能。充分发挥市场机制作用，调动社会力量参与基本医保经办服务。

二、整合基本制度政策

（一）统一覆盖范围。

城乡居民医保制度覆盖范围包括现有城镇居民医保和新农合所有应参保（合）人员，即覆盖除职工基本医疗保险应参保人员以外的其他所有城乡居民。农民工和灵活就业人员依法参加职工基本医疗保险，有困难的可按照当地规定参加城乡居民医保。各地要完善参保方式，促进应保尽保，避免重复参保。

（二）统一筹资政策。

坚持多渠道筹资，继续实行个人缴费与政府补助相结合为主的筹资方式，鼓励集体、单位或其他社会经济组织给予扶持或资助。各地要统筹考虑城乡居民医保与大病保险保障需求，按照基金收支平衡的原则，合理确定城乡统一的筹资标准。现有城镇居民医保和新农合个人缴费标准差距较大的地区，可采取差别缴费的办法，利用2—3年时间逐步过渡。整合后的实际人均筹资和个人缴费不得低于现有水平。

完善筹资动态调整机制。在精算平衡的基础上，逐步建立与经济社会发展水平、各方承受能力相适应的稳定筹资机制。逐步建立个人缴费标准与城乡居民人均可支配收入相衔接的机制。合理划分政府与个人的筹资责任，在提高政府补助标准的同时，适当提高个人缴费比重。

（三）统一保障待遇。

遵循保障适度、收支平衡的原则，均衡城乡保障待遇，逐步统一保障范围和支付标准，为参保人员提供公平的基本医疗保障。妥善处理整合前的特殊保障政策，做好过渡与衔接。

城乡居民医保基金主要用于支付参保人员发生的住院和门诊医药费用。稳定住院保障水平，政策范围内住院费用支付比例保持在75%左右。进一步完善门诊统筹，逐步提高门诊保障水平。逐步缩小政策范围内支付比例与实际支付比例间的差距。

（四）统一医保目录。

统一城乡居民医保药品目录和医疗服务项目目录，明确药品和医疗服务支付范围。各省（区、市）要按照国家基本医保用药管理和基本药物制度有关规定，遵循临床必需、安全有效、价格合理、技术适宜、基金可承受的原则，在现有城镇居民医保和新农合目录的基础上，适当考虑参保人员需求变化进行调整，有增有减、有控有扩，做到种类基本齐全、结构总体合理。完善医保目录管理办法，实行分级管理、动态调整。

（五）统一定点管理。

统一城乡居民医保定点机构管理办法，强化定点服务协议管理，建立健全考核评价机制和动态的准入退出机制。对非公立医疗机构与公立医疗机构实行同等的定点管理政策。原则上由统筹地区管理机构负责定点机构的准入、退出和监管，省级管理机构负责制订定点机构的准入原则和管理办法，并重点加强对统筹区域外的省、市级定点医疗机构的指导与监督。

（六）统一基金管理。

城乡居民医保执行国家统一的基金财务制度、会计制度和基金预决算管理制度。城乡居民医保基金纳入财政专户，实行“收支两条线”管理。基金独立核算、专户管理，任何单位和个人不得挤占挪用。

结合基金预算管理全面推进付费总额控制。基金使用遵循以收定支、收支平衡、略有结余的原则，确保应支付费用及时足额拨付，合理控制基金当年结余率和累计结余率。建立健全基金运行风险预警机制，防范基金风险，提高使用效率。

强化基金内部审计和外部监督，坚持基金收支运行情况信息公开和参保人员就医结算信息公示制度，加强社会监督、民主监督和舆论监督。

三、理顺管理体制

（一）整合经办机构。

鼓励有条件的地区理顺医保管理体制，统一基本医保行政管理职能。充分利用现有城镇居民医保、新农合经办资源，整合城乡居民医保经办机构、人员和信息系统，规范经办流程，提供一体化的经办服务。完善经办机构内外部监督制约机制，加强培训和绩效考核。

（二）创新经办管理。

完善管理运行机制，改进服务手段和管理办法，优化经办流程，提高管理效率和服务水平。鼓励有条件的地区创新经办服务模式，推进管办分开，引入竞争机制，在确保基金安全和有效监管的前提下，以政府购买服务的方式委托具有资质的商业保险机构等社会力量参与基本医保的经办服务，激发经办活力。

四、提升服务效能

（一）提高统筹层次。

城乡居民医保制度原则上实行市（地）级统筹，各地要围绕统一待遇政策、基金管理、信息系统和就医结算等重点，稳步推进市（地）级统筹。做好医保关系转移接续和异地就医结算服务。根据统筹地区内各县（市、区）的经济发展和医疗服务水平，加强基金的分级管理，充分调动县级政府、经办管理机构基金管理的积极性和主动性。鼓励有条件的地区实行省级统筹。

（二）完善信息系统。

整合现有信息系统，支撑城乡居民医保制度运行和功能拓展。推动城乡居民医保信息系统与定点机构信息系统、医疗救助信息系统的业务协同和信息共享，做好城乡居民医保信息系统与参与经办服务的商业保险机构信息系统必要的信息交换和数据共享。强化信息安全和患者信息隐私保护。

（三）完善支付方式。

系统推进按人头付费、按病种付费、按床日付费、总额预付等多种付费方式相结合的复合支付方式改革，建立健全医保经办机构与医疗机构及药品供应商的谈判协商机制和风险分担机制，推动形成合理的医保支付标准，引导定点医疗机构规范服务行为，控制医疗费用不合理增长。

通过支持参保居民与基层医疗机构及全科医师开展签约服务、制定差别化的支付政策等措施，推进分级诊疗制度建设，逐步形成基层首诊、双向转诊、急慢分治、上下联动的就医新秩序。

（四）加强医疗服务监管。

完善城乡居民医保服务监管办法，充分运用协议管理，强化对医疗服务的监控作用。各级医保经办机构要利用信息化手段，推进医保智能审核和实时监控，促进合理诊疗、合理用药。卫生计生行政部门要加强医疗服务监管，规范医疗服务行为。

五、精心组织实施，确保整合工作平稳推进

（一）加强组织领导。

整合城乡居民医保制度是深化医改的一项重点任务，关系城乡居民切身利益，涉及面广、政策性强。各地各有关部门要按照全面深化改革的战略布局要求，充分认识这项工作的重要意义，加强领导，精心组织，确保整合工作平稳有序推进。各省级医改领导小组要加强统筹协调，及时研究解决整合过程中的问题。

（二）明确工作进度和责任分工。

各省（区、市）要于2016年6月底前对整合城乡居民医保工作作出规划和部署，明确时间表、路线图，健全工作推进和考核评价机制，严格落实责任制，确保各项政策措施落实到位。各统筹地区要于2016年12月底前出台具体实施方案。综合医改试点省要将整合城乡居民医保作为重点改革内容，加强与医改其他工作的统筹协调，加快推进。

各地人力资源社会保障、卫生计生部门要完善相关政策措施，加强城乡居民医保制度整合前后的衔接；财政部门要完善基金财务会计制度，会同相关部门做好基金监管工作；保险监管部门要加强对参与经办服务的商业保险机构的从业资格审查、服务质量和市场行为监管；发展改革部门要将城乡居民医保制度整合纳入国民经济和社会发展规划；编制管理部门要在经办资源和管理体制整合工作中发挥职能作用；医改办要协调相关部门做好跟踪评价、经验总结和推广工作。

（三）做好宣传工作。

要加强正面宣传和舆论引导，及时准确解读政策，宣传各地经验亮点，妥善回应公众关切，合理引导社会预期，努力营造城乡居民医保制度整合的良好氛围。

国务院

2016年1月3日

（此件公开发布）

Appendix C: Supplemental Tables for Chapter 3

Table A-3 Trend of nationwide population and economy in China, 2006-2018

Population and Economy Indicators	Year												
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Total Population (million)	1,314	1,321	1,328	1,335	1,341	1,347	1,354	1,361	1,368	1,375	1,383	1,390	1,395
age 65+(million)	104	106	110	113	119	123	127	132	138	144	150	158	167
Aging Rate (%)	7.9	8.1	8.3	8.5	8.9	9.1	9.4	9.7	10.1	10.5	10.8	11.4	11.9
Birth Rate (‰)	12.09	12.10	12.14	11.95	11.90	11.93	12.10	12.08	12.37	12.07	12.95	12.43	10.94
GDP (¥, billion)	21,944	27,009	31,924	34,852	41,212	48,794	53,858	59,296	64,356	68,886	74,640	83,204	91,928
Per Capita GDP (¥)	16,738	20,494	24,100	26,180	30,808	36,302	39,874	43,684	47,173	50,237	54,139	60,014	66,006
PCDI (¥)	-	-	-	-	-	-	-	18,311	20,167	21,966	23,821	25,974	28,228
PCCE (¥)	-	-	-	-	-	-	-	13,220	14,491	15,712	17,111	18,322	19,853
Food, Tobacco and Liquor (¥)	-	-	-	-	-	-	-	4,127	4,494	4,814	5,151	5,374	5,631
Clothing (¥)	-	-	-	-	-	-	-	1,027	1,099	1,164	1,203	1,238	1,289
Residence (¥)	-	-	-	-	-	-	-	2,999	3,201	3,419	3,746	4,107	4,647
Household Facilities, Articles and Services (¥)	-	-	-	-	-	-	-	806	890	951	1,044	1,121	1,223
Transport and Communications (¥)	-	-	-	-	-	-	-	1,627	1,869	2,087	2,338	2,499	2,675
Education, Cultural and Recreation (¥)	-	-	-	-	-	-	-	1,398	1,536	1,723	1,915	2,086	2,226
Health Care and Medical Services (¥)	-	-	-	-	-	-	-	912	1,045	1,165	1,307	1,451	1,685
Miscellaneous Goods and Services (¥)	-	-	-	-	-	-	-	325	358	389	406	447	477

Abbreviations: GDP-Gross Domestic Product; PCDI-Per Capita Disposable Income; PCCE-Per Capita Consumption Expenditure

Source: China Statistical Yearbook

Note: Total population is sampled data; Due to changes in statistical methods, residents' income and consumption data before 2013 are not available

Table A-4 Comparison of GDP per capita among China and OECD countries, 2006 and 2018 (USD, thousand)

Country	2006	Rank in 2006	2018	Rank in 2018	Country	2006	Rank in 2006	2018	Rank in 2018
Luxembourg	89.74	1	116.64	1	Italy	33.50	19	34.48	21
Switzerland	57.58	3	82.80	2	Korea, Rep.	20.89	24	31.36	22
Norway	74.15	2	81.70	3	Spain	28.37	20	30.37	23
Ireland	54.31	5	78.81	4	Slovenia	19.67	26	26.12	24
Iceland	56.67	4	73.19	5	Portugal	19.82	25	23.41	25
United States	46.30	7	62.79	6	Estonia	12.63	29	23.27	26
Denmark	52.03	6	61.35	7	Czech Republic	15.18	27	23.08	27
Australia	36.04	17	57.37	8	Greece	24.80	22	20.32	28
Sweden	46.28	8	54.61	9	Slovak Republic	13.16	28	19.44	29
Netherlands	44.86	9	53.02	10	Lithuania	9.24	33	19.15	30
Austria	40.64	12	51.46	11	Latvia	9.66	31	17.86	31
Finland	41.19	11	50.15	12	Hungary	11.48	30	16.16	32
Germany	36.32	16	47.60	13	Chile	9.46	32	15.92	33
Belgium	38.67	14	47.52	14	Poland	9.04	35	15.42	34
Canada	40.39	13	46.23	15	China	2.10	37	9.77	35
United Kingdom	44.60	10	42.94	16	Mexico	9.07	34	9.67	36
New Zealand	26.67	21	41.95	17	Turkey	8.04	36	9.37	37
Israel	21.84	23	41.72	18					
France	36.44	15	41.46	19					
Japan	35.43	18	39.29	20					

Source: World Bank Open Data

Table A-5 Trend of urban-rural population and economy in China, 2006-2018

Area	Population and Economy Indicators	Year												
		2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
Urban Areas	Population (million)	583	606	624	645	670	691	712	731	749	771	793	813	831
	Population (% of total)	44.34	45.89	46.99	48.34	49.95	51.27	52.57	53.73	54.77	56.1	57.35	58.52	59.58
	PCDI (¥)	11,620	13,603	15,549	16,901	18,779	21,427	24,127	26,467	28,844	31,195	33,616	36,396	39,251
	PCCE (¥)	8,697	9,998	11,243	12,265	13,472	15,161	16,674	18,488	19,968	21,392	23,079	24,445	26,112
	Health Care and Medical Services (¥)	620.5	699	786	856	872	969	1,064	1,136	1,306	1,443	1,631	1,777	2,046
	Health Care and Medical Services (% of PCDI)	5.34	5.14	5.06	5.07	4.64	4.52	4.41	4.29	4.53	4.63	4.85	4.88	5.21
Rural Areas	Population (million)	732	715	704	689	671	657	642	630	619	603	590	577	564
	Population (% of total)	55.66	54.11	53.01	51.66	50.05	48.73	47.43	46.27	45.23	43.90	42.65	41.48	40.42
	PCDI (¥)	-	-	-	-	-	-	-	9,430	10,489	11,422	12,363	13,432	14,617
	PCCE (¥)	2,829	3,224	3,661	3,994	4,382	5,221	5,908	7,485	8,383	9,223	10,130	10,955	12,124
	Health Care and Medical Services (¥)	191.5	210	246	288	326	437	514	668	754	846	929	1,059	1,240
	Health Care and Medical Services (% of PCDI)	-	-	-	-	-	-	-	7.09	7.19	7.41	7.52	7.88	8.48

Abbreviations: PCDI-Per Capita Disposable Income; PCCE-Per Capita Expenditure

Source: China Statistical Yearbook

Note: Total population is sampled data; Due to changes in statistical methods, rural residents' PCDI data before 2013 are not available

Table A-6 Trend of nationwide and urban-rural health expenditure in China, 2006-2018

Area	Health Expenditure Indicators	Year												
		2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
Nationwide	Total HE (¥, billion)	984	1,157	1,454	1,754	1,998	2,435	2,812	3,167	3,531	4,097	4,634	5,260	5,912
	Government HE (¥, billion)	178	258	359	482	573	746	843	955	1,058	1,248	1,391	1,521	1,640
	Social HE (¥, billion)	321	389	507	615	720	842	1,003	1,139	1,344	1,651	1,910	2,226	2,581
	OOP HE (¥, billion)	485	510	588	657	705	847	966	1,073	1,130	1,199	1,334	1,513	1,691
	Total HE (% of GDP)	4.52	4.32	4.59	5.08	4.89	5.03	5.26	5.39	5.55	5.98	6.23	6.36	6.57
	Government HE (% of total)	18.07	22.31	24.73	27.46	28.69	30.66	29.99	30.14	29.96	30.45	30.01	28.91	27.74
	Social HE (% of total)	32.62	33.64	34.85	35.08	36.02	34.57	35.67	35.98	38.05	40.29	41.21	42.32	43.66
	OOP HE (% of total)	49.31	44.05	40.42	37.46	35.29	34.77	34.34	33.88	31.99	29.27	28.78	28.77	28.61
	Per Capita HE, ¥	749	876	1,095	1,314	1,490	1,807	2,077	2,327	2,582	2,981	3,352	3,784	4,237
Urban Areas	Total HE (¥, billion)	717	897	1,125	1,354	1,551	1,857	2,128	2,364	2,658	3,130	3,546	-	-
	Per Capita HE (¥)	1,248	1,516	1,862	2,177	2,316	2,698	2,999	3,234	3,558	4,059	4,472	-	-
Rural Areas	Total HE (¥, billion)	267	261	328	401	447	577	684	802	874	968	1,089	-	-
	Per Capita HE (¥)	362	358	455	562	666	879	1,065	1,274	1,412	1,604	1,846	-	-

Abbreviations: HE-Health Expenditure; OOP-Out-Of-Pocket; GDP-Gross Domestic Product

Source: China Health Statistical Yearbook

Note: Urban-Rural Health Expenditure data for 2017 and 2018 have not yet been released.

Table A-7 Trend of nationwide health care resources in China, 2006-2018

Health Care Resources Indicators	Year												
	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
Health Care Institutions (thousand)	918	912	891	917	937	954	950	974	981	984	983	987	997
Hospitals (thousand)	19	20	20	20	21	22	23	25	26	28	29	31	33
Tertiary Hospitals	1,045	1,182	1,192	1,233	1,284	1,399	1,624	1,787	1,954	2,123	2,232	2,340	2,548
Grass-root Institutions (thousand)	885	879	858	882	902	918	913	915	917	921	927	933	944
Medical Personnel (thousand)	6,681	6,964	7,252	7,781	8,208	8,616	9,116	9,790	10,234	10,694	11,173	11,749	12,300
Medical Technical Personnel (thousand)	4,728	4,913	5,174	5,535	5,876	6,203	6,676	7,211	7,590	8,008	8,454	8,988	9,529
Licensed (Assistant) Doctors (thousand)	2,099	2,123	2,202	2,329	2,413	2,466	2,616	2,795	2,893	3,039	3,191	3,390	3,607
Registered Nurse (thousand)	1,426	1,559	1,678	1,855	2,048	2,244	2,497	2,783	3,004	3,241	3,507	3,804	4,099
Village Doctors and Assistants (thousand)	957	932	938	1,051	1,092	1,126	1,094	1,081	1,058	1,032	1,000	969	907
Sickbeds (thousand)	3,512	3,701	4,039	4,417	4,787	5,160	5,725	6,182	6,601	7,015	7,411	7,941	8,404
in Hospitals (thousand)	2,560	2,675	2,883	3,121	3,387	3,705	4,162	4,579	4,961	5,331	5,689	6,121	6,520
in Basic Medical Institutions (thousand)	762	850	971	1,100	1,192	1,234	1,324	1,350	1,381	1,414	1,442	1,529	1,584
Number of Residents per Hospitals and Basic Medical Institutions	1,454	1,470	1,513	1,479	1,453	1,433	1,447	1,447	1,450	1,449	1,447	1,442	1,429
Number of Residents per Tertiary Hospitals (thousand)	1,258	1,118	1,114	1,082	1,044	963	834	761	700	647	619	594	548
Medical Technical Personnel per 1000 Residents	3.60	3.72	3.90	4.15	4.39	4.61	4.94	5.27	5.56	5.83	6.12	6.47	6.83
Licensed (Assistant) Doctors per 1000 Residents	1.60	1.61	1.66	1.75	1.80	1.83	1.94	2.04	2.12	2.22	2.31	2.44	2.59
Registered Nurse per 1000 Residents	1.09	1.18	1.27	1.39	1.53	1.67	1.85	2.04	2.20	2.37	2.54	2.74	2.94
Sickbeds per 1000 Residents	2.70	2.83	3.05	3.32	3.58	3.84	4.24	4.55	4.85	5.11	5.37	5.72	6.03

Source: China Health Statistical Yearbook

Table A-8 Comparison of health care resources among China and OECD countries, 2017

Country	Doctor per 1000 persons	Nurse per 1000 persons	Hospital Bed per 1000 persons	Country	Doctor per 1000 persons	Nurse per 1000 persons	Hospital Bed per 1000 persons
China	2.01	2.7	4.34	Korea	2.34	6.91	12.27
Australia	3.68	11.68	-	Latvia	3.21	4.57	5.57
Austria	5.18	6.85	7.37	Lithuania	4.56	7.71	6.56
Belgium	3.08	-	5.66	Luxembourg	2.98	11.72	4.66
Canada	2.72	9.96	2.52	Mexico	2.43	2.9	1.38
Chile	-	-	2.11	Netherlands	-	-	3.32
Czech Republic	-	8.06	6.63	New Zealand	3.27	10.17	2.71
Denmark	-	-	2.61	Norway	4.66	17.67	3.6
Estonia	3.47	6.19	4.69	Poland	2.38	5.1	6.62
Finland	-	-	3.28	Portugal	-	6.7	3.39
France	3.37	10.48	5.98	Slovak Republic	3.4	5.65	5.82
Germany	4.25	12.93	8	Slovenia	3.1	9.92	4.5
Greece	-	3.31	4.21	Spain	3.88	5.74	2.97
Hungary	3.32	6.51	7.02	Sweden	-	-	2.22
Iceland	3.87	14.5	3.06	Switzerland	4.3	17.23	4.53
Ireland	3.07	12.16	2.96	Turkey	1.87	2.07	2.81
Israel	3.14	5.08	3.02	United Kingdom	2.81	7.83	2.54
Italy	3.99	6.71	3.18	United States	2.61	11.74	-
Japan	-	-	13.05				

Source: OECD data; Abbreviations: OECD- The Organisation for Economic Co-operation and Development

Note: The data of nurse in OECD and China Yearbook is similar (2.70 vs. 2.74), but the data of doctor and bed of China in OECD data are much lower than in China Health Statistical Yearbook. After check, the possible reasons are 1) Only licensed doctors (excluding assistant doctors) are counted in OECD (2.01 in OECD data vs. 2.04 licensed doctors in Yearbook); 2) Only hospital beds.(excluding basic medical institutions' beds) are counted (4.34 in OECD data vs. 4.4 hospital bed calculated based on Yearbook).

Table A-9 Trend of urban-rural health care resources in China, 2006-2018

Area	Health Care Resources Indicators	Year												
		2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
Urban Areas	Medical Technical Personnel per 1000 Residents	6.09	6.44	6.68	7.15	7.62	6.68	8.54	9.18	9.70	10.21	10.42	10.87	10.91
	Licensed (Assistant) Doctors per 1000 Residents	2.56	2.61	2.68	2.83	2.97	2.62	3.19	3.39	3.54	3.72	3.79	3.97	4.01
	Registered Nurse per 1000 Residents	2.22	2.42	2.54	2.82	3.09	2.62	3.65	4.00	4.30	4.58	4.75	5.01	5.08
	Sickbeds per 1000 Residents	-	4.90	5.17	5.54	5.94	6.24	6.88	7.36	7.84	8.27	8.41	8.75	8.70
Rural Areas	Medical Technical Personnel per 1000 Residents	2.70	2.69	2.80	2.94	3.04	2.66	3.41	3.64	3.77	3.90	4.08	4.28	4.63
	Licensed (Assistant) Doctors per 1000 Residents	1.26	1.23	1.26	1.31	1.32	1.10	1.40	1.48	1.51	1.55	1.61	1.68	1.82
	Registered Nurse per 1000 Residents	0.66	0.70	0.76	0.81	0.89	0.79	1.09	1.22	1.31	1.39	1.50	1.62	1.80
	Sickbeds per 1000 Residents	-	2.00	2.20	2.41	2.60	2.80	3.11	3.35	3.54	3.71	3.91	4.19	4.56

Source: China Health Statistical Yearbook

Table A-10 Trend of nationwide health service use and cost in China, 2006-2018

Health Service Use and Cost Indicators	Year												
	2006	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**
Health Institutions Visits (times, billion)	4.46	4.72	4.90	5.49	5.84	6.27	6.89	7.31	7.60	7.69	7.93	8.18	8.31
Hospital Visits (times, billion)	1.47	1.64	1.78	1.92	2.04	2.26	2.54	2.74	2.97	3.08	3.27	3.44	3.58
Basic Medical Institutions Visits (times, billion)	2.87	2.94	2.96	3.39	3.61	3.81	4.11	4.32	4.36	4.34	4.37	4.43	4.41
Average Number of Visits per year (times)	1.9	2.2	2.7	4.2	4.4	4.7	5.1	5.4	5.6	5.6	5.7	5.9	6.0
Health Institutions Hospitalisations (persons, million)	79	98	115	133	142	153	179	192	204	211	227	244	255
Hospitalisations in Hospitals (persons, million)	56	65	74	85	95	108	127	140	154	161	175	189	200
Hospitalisations in Basic Medical Institutions (persons, million)	19	28	35	41	40	38	43	43	41	40	42	45	44
Annual Hospitalisation Rate (%)	6.14	7.54	8.7	9.95	10.59	11.37	13.21	14.12	14.97	15.32	16.46	17.6	18.27
Average Outpatient Costs in Hospital per time (¥)	129	125	138	152	167	180	193	206	220	234	246	257	274
Drug Costs (¥)	65	63	71	78	86	91	97	102	106	111	112	110	112
Drug Costs (%)	50	50.7	51.3	51.5	51.3	50.6	50.3	49.3	48.3	47.2	45.5	42.7	40.9
Average Hospitalisation Costs in Hospital per person (¥)	4,669	4,734	5,234	5,684	6,194	6,632	6,980	7,442	7,832	8,268	8,605	8,891	9,292
Drug Costs (¥)	1992.04	2,015	2,276	2,481	2,670	2,771	2,867	2,939	2,999	3,042	2,978	2,765	2,622
Drug Costs (%)	42.67	42.6	43.5	43.6	43.1	41.8	41.1	39.5	38.3	36.8	34.6	31.1	28.2

Source: China Health Statistical Yearbook

Note: Health Institution Visits include Outpatient, Emergency, and Other Visits, such as Health Management/Consultant.

Table A-11 Provincial population and economy in China, 2018

Area (Provincial)	GDP (Billion, ¥)	Population (Million)	Per Capita GDP (¥)	PCDI (¥)	PCCE (¥)	PCCE, Healthcare (¥)
Beijing	3,311	21.54	153,095	62,361	39,843	3,275
Tianjin	1,336	15.60	85,757	39,506	29,903	2,677
Hebei	3,249	75.56	43,108	23,446	16,722	1,541
Shanxi	1,596	37.18	43,010	21,990	14,810	1,635
Neimenggu	1,614	25.34	63,772	28,376	19,665	1,848
Liaoning	2,351	43.59	53,872	29,701	21,398	2,257
Jilin	1,125	27.04	41,516	22,798	17,200	2,012
Heilongjiang	1,285	37.73	33,977	22,726	16,994	2,235
Shanghai	3,601	24.24	148,744	64,183	43,351	3,070
Jiangsu	9,321	80.51	115,930	38,096	25,007	2,016
Zhejiang	5,800	57.37	101,813	45,840	29,471	2,059
Anhui	3,401	63.24	54,078	23,984	17,045	1,224
Fujian	3,869	39.41	98,542	32,644	22,996	1,235
Jiangxi	2,272	46.48	49,013	24,080	15,792	1,000
Shandong	6,665	100.47	66,472	29,205	18,780	1,628
Henan	4,994	96.05	52,114	21,964	15,169	1,541
Hubei	4,202	59.17	71,109	25,815	19,538	1,908
Hunan	3,633	68.99	52,809	25,241	18,808	1,706
Guangdong	9,995	113.46	88,781	35,810	26,054	1,521
Guangxi	1,963	49.26	40,012	21,485	14,935	1,365
Hainan	491	9.34	52,801	24,579	17,528	1,236
Chongqing	2,159	31.02	69,901	26,386	19,248	1,660
Sichuan	4,290	83.41	51,556	22,461	17,664	1,569
Guizhou	1,535	36.00	42,767	18,430	13,798	1,083
Yunnan	2,088	48.30	43,366	20,084	14,250	1,268
Xizang	155	3.44	45,476	17,286	11,520	460
Shaanxi	2,394	38.64	62,195	22,528	16,160	1,749
Gansu	810	26.37	30,797	17,488	14,624	1,574
Qinghai	275	6.03	45,739	20,757	16,557	1,842
Ningxia	351	6.88	51,248	22,400	16,715	1,727
Xinjiang	1,281	24.87	51,950	21,500	16,189	1,593

Abbreviations: GDP-Gross Domestic Product; PCDI-Per Capita Disposable Income; PCCE-Per Capita Consumption Expenditure
Source: China Statistical Yearbook

Table A-12 Provincial healthcare institutions in China, 2018

Area (Provincial)	The Number of Inhabitants per Health Care Institution	Rank	The Number of Inhabitants per Tertiary Hospitals	Rank
East	1,554	-	555,005	-
Middle	1,403	-	791,071	-
West	1,214	-	570,767	-
Beijing	2,142	26	211,176	1
Tianjin	2,744	30	362,791	6
Hebei	888	3	1,064,225	31
Shanxi	884	2	652,281	24
Neimenggu	1,030	7	316,750	4
Liaoning	1,210	10	325,299	5
Jilin	1,192	9	530,196	17
Heilongjiang	1,854	24	393,021	7
Shanghai	4,580	31	515,745	14
Jiangsu	2,421	28	500,062	12
Zhejiang	1,752	22	428,134	10
Anhui	2,537	29	930,000	29
Fujian	1,428	17	511,818	13
Jiangxi	1,272	13	673,623	25
Shandong	1,233	12	555,083	19
Henan	1,346	15	1,021,809	30
Hubei	1,622	21	455,154	11
Hunan	1,227	11	873,291	28
Guangdong	2,205	27	550,777	18
Guangxi	1,460	18	639,740	22
Hainan	1,754	23	424,545	9
Chongqing	1,511	19	646,250	23
Sichuan	1,023	6	417,050	8
Guizhou	1,283	14	620,690	21
Yunnan	1,936	25	700,000	26
Xizang	503	1	286,667	2
Shaanxi	1,095	8	603,750	20
Gansu	945	5	712,703	27
Qinghai	943	4	301,500	3
Ningxia	1,546	20	529,231	16
Xinjiang	1,348	16	518,125	15

Calculation: The Number of Inhabitants per Health Care Institution=Total Population/Total Healthcare Institution; The Number of Inhabitants per Tertiary Hospitals= Total population/Total Tertiary Hospitals
Note: Rank from the most abundance (1) to the scarcest (31).

Table A-13 Provincial health care personnel and sickbeds in China, 2018

Area (Provincial)	Medical Technical Personnel per 1000 inhabitants	Rank	Licensed (Assistant) Doctors per 1000 inhabitants	Rank	Registered Nurse per 1000 inhabitants	Rank	Sickbeds per 1000 inhabitants	Rank
Beijing	11.9	1	4.6	1	5	1	5.74	20
Tianjin	6.7	16	2.8	8	2.5	25	4.37	31
Hebei	6.1	26	2.8	9	2.3	30	5.58	23
Shanxi	6.6	20	2.7	12	2.8	19	5.6	22
Neimenggu	7.4	6	2.9	4	3	12	6.27	12
Liaoning	7	11	2.8	10	3.1	10	7.21	1
Jilin	6.8	13	2.9	5	2.8	20	6.18	13
Heilongjiang	6.1	27	2.4	21	2.5	26	6.63	8
Shanghai	8.1	4	3	3	3.6	2	5.74	21
Jiangsu	7.3	9	2.9	6	3.2	6	6.11	15
Zhejiang	8.5	2	3.3	2	3.5	4	5.79	19
Anhui	5.3	30	2	30	2.4	27	5.19	26
Fujian	6.3	23	2.3	25	2.8	21	4.88	27
Jiangxi	5.3	31	1.9	31	2.4	28	5.37	24
Shandong	7.4	7	2.9	7	3.2	7	6.06	16
Henan	6.5	21	2.5	17	2.7	23	6.34	11
Hubei	6.9	12	2.6	14	3.2	8	6.65	7
Hunan	6.3	24	2.6	15	2.7	24	6.99	5
Guangdong	6.7	17	2.4	22	2.9	15	4.56	30
Guangxi	6.5	22	2.2	28	2.9	16	5.2	25
Hainan	6.8	14	2.4	23	3.2	9	4.8	29
Chongqing	6.7	18	2.5	18	3.1	11	7.1	4
Sichuan	6.7	19	2.5	19	3	13	7.18	3
Guizhou	6.8	15	2.3	26	3	14	6.82	6
Yunnan	6.2	25	2.1	29	2.8	22	6.03	17
Xizang	5.5	29	2.4	24	1.6	31	4.88	28
Shaanxi	8.5	3	2.6	16	3.6	3	6.57	9
Gansu	6	28	2.3	27	2.4	29	6.17	14
Qinghai	7.4	8	2.7	13	2.9	17	6.49	10
Ningxia	7.7	5	2.8	11	3.4	5	5.96	18
Xinjiang	7.1	10	2.5	20	2.9	18	7.19	2

Source: China Health Statistical Yearbook
Note: Rank from the most abundance (1) to the scarcest (31).

Table A-14 Provincial health service use in China, 2018

Area (Provincial)	Average Number of Health Institution Visits Per Year (times)	Rank	Annual Hospitalisation Rate (%)	Rank
East	7.34	-	16.5	-
Middle	4.74	-	18.7	-
West	5.21	-	20.3	-
Beijing	10.92	29	16.4	13
Tianjin	7.69	28	10.4	2
Hebei	5.71	19	16.1	11
Shanxi	3.49	2	13.3	4
Neimenggu	4.16	5	15.2	8
Liaoning	4.56	9	17.0	14
Jilin	4.09	4	15.0	6
Heilongjiang	2.96	1	15.5	9
Shanghai	11.15	31	17.3	15
Jiangsu	7.38	26	18.0	18
Zhejiang	10.94	30	17.8	17
Anhui	4.70	11	16.0	10
Fujian	5.93	20	14.6	5
Jiangxi	4.57	10	18.6	21
Shandong	6.53	25	18.3	19
Henan	6.10	23	20.0	24
Hubei	5.94	21	22.3	28
Hunan	3.90	3	22.3	29
Guangdong	7.45	27	15.1	7
Guangxi	5.19	16	18.9	22
Hainan	5.44	18	12.8	3
Chongqing	5.15	15	22.7	31
Sichuan	6.19	24	22.0	27
Guizhou	4.54	8	22.6	30
Yunnan	5.35	17	19.9	23
Xizang	4.77	12	9.0	1
Shaanxi	5.08	14	20.7	25
Gansu	5.02	13	18.5	20
Qinghai	4.20	6	16.3	12
Ningxia	6.02	22	17.6	16
Xinjiang	4.31	7	21.8	26

Source: China Health Statistical Yearbook
Note: Rank from the lowest (1) to the highest (31); Health Institution Visits include Outpatient, Emergency, and Other Visits, such as Health Management/Consultant.

Table A-15 Provincial health service cost in China, 2018

Area (Provincial)	Average Outpatient Costs in Hospital per time (¥)	Rank	Average Hospitalisation Costs in Hospital per person (¥)	Rank
Beijing	545	31	22,619	31
Tianjin	339	29	17,148	29
Hebei	240	11	8,915	19
Shanxi	256	16	9,005	20
Neimenggu	251	14	8,529	14
Liaoning	308	27	9,446	23
Jilin	276	23	9,802	24
Heilongjiang	265	21	8,771	18
Shanghai	379	30	18,390	30
Jiangsu	281	24	11,216	26
Zhejiang	265	22	11,517	27
Anhui	239	10	7,451	7
Fujian	262	19	8,204	13
Jiangxi	257	18	8,166	12
Shandong	254	15	9,275	22
Henan	194	3	8,157	11
Hubei	244	13	9,089	21
Hunan	291	26	8,021	8
Guangdong	282	25	11,549	28
Guangxi	206	4	8,652	16
Hainan	263	20	9,908	25
Chongqing	320	28	8,145	10
Sichuan	256	17	8,073	9
Guizhou	236	9	5,731	1
Yunnan	214	5	6,423	3
Xizang	183	1	8,732	17
Shaanxi	243	12	7,424	6
Gansu	188	2	5,804	2
Qinghai	217	6	8,621	15
Ningxia	225	7	7,284	4
Xinjiang	232	8	7,329	5

Source: China Health Statistical Yearbook

Note: Rank from the cheapest (1) to the most expensive (31).

Appendix D: Literature Review of the Effects of Social Health Insurance in China

Searching Strategy

For the search on China's social health insurance study, keywords included "(China OR Chinese) AND (health insurance OR medical insurance OR urban resident basic medical insurance OR URBMI OR new cooperative medical scheme OR NCMS OR NRCMS)". The published date of the research was limited from January 2007 to June 2019. According to the title and abstract, quantitative studies related to health services use and health expenses were included, while quantitative studies unrelated to these two topics and all qualitative studies were excluded.

Results

According to the targeted insurance, these studies can be divided into two categories: one is to consider social health insurance as a whole without considering specific insurance plans, such as the cost difference study for patients with or without social health insurance; the other one is to study the impact of a single type of insurance or to compare the differences in the impact of different insurance schemes.

For health service use, in the earliest research, Wagstaff and Lindelow (2008) used 3 different data sources: one had all population in 9 provinces 1991, 1993, 1997 and 2000 data, and one had children in a village in one province in 2000 and 2003 data, and one had 7 provinces villages and towns 1998 data. They used the FE model and IV to solve the endogeneity problem. They found that insured patients are more likely to use health services (promote effect) and visit more high-level healthcare institutions. Considering the specific insurance schemes, the performance of UEBMI is the best in all studies. It significantly increases both outpatient and hospitalisation visits. While the performances of URBMI and NRCMS are ambiguous, some studies have shown that the increase in visits and some studies have shown no effect (Lei and Lin, 2009; Wagstaff *et al.*, 2009; Yu *et al.*, 2010; Li and Zhang, 2013; Chen, Liu and Xu, 2014; Zhou, Zhou, *et al.*, 2014; Cheng *et al.*, 2015). Even in subgroup analysis, this inconsistency still exists. For example, Wagstaff *et al.* (2009) found that the wealthiest people received the greatest impact from insurance, while the poorest are not affected significantly. This means that

insurance is pro-rich. Yu *et al.* (2010) found that high-income people benefit more from NRCMS but Cheng *et al.* (2015) found the low-income people benefit more. This inconsistency in subgroup analysis results has led to a discussion of whether the implementation of insurance reduced or enlarged healthcare inequality. In addition, the study also found that insurance increased the use of preventive care, particularly general physical examinations (Lei and Lin, 2009).

For health expenses, Wagstaff and Lindelow (2008) showed that insured people have higher cash outlays regardless of insurance type. For NRCMS, only one study thought it decreased out-of-pocket (OOP) expenses (Wagstaff *et al.*, 2009), all other studies showed that NRCMS has no effect on OOP expenses and the occurrence of catastrophic medical expenses (Lei and Lin, 2009; Q. Sun *et al.*, 2009; Li and Zhang, 2013; Long *et al.*, 2013; Cheng *et al.*, 2015). The impact of URBMI is also not significant (Liu and Zhao, 2014). The total costs are increased (Wagstaff *et al.*, 2009; Liu and Zhao, 2014). From the cost perspective, it is hard to say whether it is a good thing or a bad thing: it is good if patients get better service they need due to the dropped marginal costs; it is bad if the doctor gives the patient more unnecessary examination and treatment, this is a supply-side moral hazard. From the perspective of life expectancy and mortality, Huang and Gao (2010) argued that increased total medical costs and healthcare use due to insurance lead to increased life expectancy and decreased mortality, and are therefore effective demand rather than excessive demand.

The inconsistency of the research results mainly comes from three aspects: research time, research methods, and research population. How long the policy has been implemented may be an influential factor of policy effect. For example, NRCMS started in 2003, Wagstaff *et al.*(2009) used data from the implementation of the policy two years later: 2003 and 2005; Q. Sun *et al.*(2009)and Yu *et al.* (2010) data three years later: from mid-2005 to mid-2006; Jiang *et al.*(2012) used data five years later: 2008 data. Different statistical methods can also lead to different estimation results. Previous studies have involved many methods: OLS, fix-effect panel analysis, PSM-DID, panel model with IV, and two-part model. Another important reason is population. For example, Wagstaff *et al.*(2009) and Jiang *et al.*(2012) conducted a national-

level analysis, which represented the population of the whole China. Q. Sun *et al.*(2009) and Yu *et al.* (2010) only focused on two provinces: Shandong and Ningxia. And Li and Zhang (2013) only focused on two provinces: Zhejiang and Gansu. Regional differences may be a possible reason to explain to some extent why insurance is pro-rich in some areas and pro-poor in some areas. The poor population in one region (developed areas) may belong to the rich population in another region (poor areas). In addition, some studies only focused on old people, such as Li and Zhang (2013) and Cheng *et al.*, (2015); and some only focused on chronic disease patients, such as Q. Sun *et al.*(2009) and Jiang *et al.*(2012).

In summary, considering the huge differences in the implementation of China's health insurance in different regions, it is worthwhile to pay attention to the difference between the national level and the regional level, and the difference between one region and another when analysing the impact of insurance. It is also necessary to consider the selection of appropriate statistical methods.

Appendix E: Supplemental Tables for Chapter 7

Table A-16 Summary of policy document websites

City	Policy Document	Web Link
Chengdu	Rules for the Implementation of Basic Medical Insurance for Urban Residents in Chengdu	http://www.cdzgh.com/htm/detail_02050603_6473.html
	Trial Measures for Basic Medical Insurance for Urban Residents in Chengdu	http://www.chinalawedu.com/falvfagui/fg23051/252236.shtml
	Interim Measures for Basic Medical Insurance for Urban and Rural Residents in Chengdu	http://gk.chengdu.gov.cn/govInfoPub/detail.action?id=12529&tn=6
	Notice on the Funding of 2011 Basic Medical Insurance for Urban and Rural Residents and Supplementary Medical Insurance for Major Diseases	http://gk.chengdu.gov.cn/govInfoPub/detail.action?id=224853&tn=2
	Notice on the Funding of 2010 Basic Medical Insurance for Urban and Rural Residents	https://wenku.baidu.com/view/b1024381d4d8d15abe234e04.html
	Interim Measures for Urban and Rural Basic Medical Insurance Outpatient Coordination in Chengdu	https://wenku.baidu.com/view/75419feab8f67c1cfad6b8a8.html
	Notice on the adjustment of the starting payment standard and the reimbursement ratio of medical expenses for urban and rural residents' basic medical insurance	http://gk.chengdu.gov.cn/govInfoPub/detail.action?id=372889&tn=2
	Notice on Further Improving the Outpatient Coordination of Basic Medical Insurance for Urban and Rural Residents	http://gk.chengdu.gov.cn/govInfoPub/detail.action?id=113243&tn=6
	Notice of Shuangliu County People's Government on the Issuance of Supplementary Provisions to the Implementation Plan of the New Rural Cooperative Medical System of Shuangliu County in 2008	http://gk.chengdu.gov.cn/govInfoPub/detail.action?id=21886&tn=6

City	Policy Document	Web Link
Tianjin	List of Medical Insurance Policies for Urban and Rural Residents in Chengdu (2013)	https://www.doc88.com/p-9949386686232.html
	List of Medical Insurance Policies for Urban and Rural Residents in Chengdu (2018)	http://cd.bendibao.com/news/20171213/94542.shtm
	Notice on the Issuance of Provisions on Basic Medical Insurance for Urban and Rural Residents in Tianjin	https://www.tj.gov.cn/zw/zfgb/qk/2009/10_3186/201811/t20181116_3637821.html
	Notice on the Issuance of the Rules for the Implementation of the Provisional Provisions on Basic Medical Insurance for Urban Residents in Tianjin	https://code.fabao365.com/law_205185.html
	Notice on the Issuance of the Administrative Measures for the Administration of Basic Medical Insurance for Urban Residents in Tianjin	https://si12333.cn/policy/pbyi.html
	Notice of the General Office of Tianjin Municipal People's Government on the work related to the New Rural Cooperative Medical Scheme in the city in 2009	https://wenku.baidu.com/view/9100cf184b2fb4daa58da0116c175f0e7cd119c4.html?_wkts_=1670378335284&bdQuery=%E5%A4%A9%E6%B4%A5%E6%96%B0%E5%9E%8B%E5%86%9C%E6%9D%91%E5%90%88%E4%BD%9C%E5%8C%BB%E7%96%97
Shanghai	Shanghai Urban Residents' Basic Medical Insurance Trial Measures	https://yljk.sufe.edu.cn/7a/32/c1400a31282/page.htm
	Notice on the Issuance of the Shanghai Basic Medical Insurance Measures for Urban and Rural Residents	https://www.shanghai.gov.cn/nw12344/20210105/0099fc847579431e85dc33e8fd3626e7.html
	Notice of the General Office of Shanghai Municipal People's Government on Matters Relating to the Basic Medical Insurance for Urban and Rural Residents in the City in 2017	http://www.shpt.gov.cn/rensheju/gkrs-baoxian/20171106/272593.html
	Shanghai Medical Insurance Benefits	https://www.kankanews.com/a/2016-01-25/0037348807.shtml

City	Policy Document	Web Link
Beijing	Shanghai Basic Medical Insurance for Urban and Rural Residents (2015)	https://www.yjbys.com/shebao/yiliao/620957.html
	Notice on Matters Relating to the City's Basic Medical Insurance for Urban and Rural Residents in 2016	www.hanghai.gov.cn/nw2/nw2314/nw2319/nw12344/u26aw45326.html?date=2015-10-28
	Measures for Basic Medical Insurance for Urban and Rural Residents in Beijing	http://www.mohrss.gov.cn/yiliaobxs/YILIAOBXSgongzuodongtai/201712/t20171222_284834.html
	Rules for the Implementation of the Basic Medical Insurance Scheme for Urban and Rural Residents in Beijing	http://www.beijing.gov.cn/zhengce/zhengcefagui/201905/t20190522_60735.html
	Urban and rural residents' health insurance contribution rates for 2018	http://www.beijing.gov.cn/zhengce/zcjd/201905/t20190523_78174.html
	Announcement on urban and rural residents' basic medical insurance enrolment and payment	http://www.beijing.gov.cn/fuwu/bmfw/bmzt/cxyb/
	Notice on issues related to the payment of insurance premiums for the 2017 Beijing urban residents' basic medical insurance	http://www.beijing.gov.cn/zhengce/zhengcefagui/201905/t20190522_59442.html
	Notice on the Adjustment of Funding Standards for Basic Medical Insurance for Urban Residents in Beijing	http://www.beijing.gov.cn/zhengce/zhengcefagui/201905/t20190522_58062.html

Appendix F: Supplemental Tables for Chapter 8

Table A-17 Summary of original questions of outcome variables

Variable	Number	Question	Option	Recode
Outpatient visit occurrence	ED001	In the last month have you visited a public hospital, private hospital, public health center, clinic, or health worker's or doctor's practice, or been visited by a health worker or doctor for outpatient care?	(1) Yes (2) No	1="Yes" 0="No"
Outpatient out-of-pocket costs	ED024	How much did you pay out-of-pocket, after reimbursement from insurance?	(Fill in the blanks)	-
Hospitalisation occurrence	EE003	Have you received inpatient care in the past year?	(1) Yes (2) No	1="Yes" 0="No"
Length of stay in hospital	EE016	How many nights were you hospitalised there?	(Fill in the blanks)	
Hospitalisation out-of-pocket costs	EE027	How much did you or will you eventually pay out-of-pocket for the total costs of hospitalisation?	(Fill in the blanks)	-

Table A-18 Summary of original questions of control variables-1

Variable	Number	Question and options	Recode
Age	BA002	When were you born? (Fill in the blanks)	age=interview year - birth year
Sex	CV005	(Record Gender by Interviewer) (1) Male (2) Female	1="Male" 0="Female"
Urban/rural	-	(Record Region by Interviewer) (1) Urban (2) Rural	1="Urban" 0="Rural"
Education	BD001	What is the highest level of education you have attained? (1) No formal education (illiterate) (2) Did not finish primary school but capable of reading and/or writing (3) Sishu/home school (4) Elementary school (5) Middle school (6) High school (7) Vocational school (8) Two-/ Three-Year College/ Associate degree (9) Four-Year College/Bachelor's degree (10) Master's degree (11) Doctoral degree/Ph.D.	1="illiterate, not finish primary school, no formal education" 2="home school, primary school" 3="middle school,high school,vocational school" 4="Associate degree,Bachelor's degree" 5="Master,Doctoral"
Marriage or cohabitation	BE001	What is your marital status? (1) Married with spouse present (2) Married but not living with spouse temporarily for reasons such as work (3) Separated (4) Divorced (5) Widowed (6) Never married	1="Married with spouse present, Married but not living with spouse temporarily for reasons such as work, Living with a partner"
	BE002	Are you unmarried but Living with a partner? (1) Yes (2) No	0=other cases
Self-report health	DA002	Would you say your health is very good, good, fair, poor or very poor? (1) Very good (2) Good (3) Fair (4) Poor (5) Very poor	-

Table A-19 Summary of original questions of control variables-2

Variable	Number	Question and options	Recode
Number of disabilities	DA005	Do you have one of the following disabilities? (1) Physical disabilities (2) Brain damage/mental retardation (3) Vision problem (4) Hearing problem (5) Speech impediment	-
Number of chronic diseases	DA007	Have you been diagnosed with by a doctor? (1) Hypertension (2) Dyslipidemia (elevation of low density lipoprotein, triglycerides (TGs),and total cholesterol, or a low high density lipoprotein level) (3) Diabetes or high blood sugar (4) Cancer or malignant tumor (excluding minor skin cancers) (5) Chronic lung diseases, such as chronic bronchitis , emphysema (excluding tumors, or cancer) (6) Liver disease (except fatty liver, tumors, and cancer) (7) Heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems (8) Stroke (9) Kidney disease (except for tumor or cancer) (10) Stomach or other digestive disease (except for tumor or cancer) (11) Emotional, nervous, or psychiatric problems (12) Memory-related disease (13) Arthritis or rheumatism (14) Asthma	-

Table A-20 Summary of original questions of control variables-3

Variable	Number	Question and options	Recode
Number of cancers	DA017	In which organ or part of your body do you have cancer? Including the origins and metastasis of tumor.（circle all that apply） (1) Brain (2) Oral cavity (3) Larynx (4) Other pharynx (5) Thyroid (6) Lung (7) Breast (8) Oesophagus (9) Stomach (10) Liver (11) Pancreas (12) Kidney (13) Prostate (14) Testicle (15) Ovary (16) Cervix (17) Endometrium (18) Colon or rectum (19) Bladder (20) Skin (21) Non-Hodgkin lymphoma (22) Leukemia (23) Other organ	-
Kind of social activities	DA056	Have you done any of these activities in the last month? (Code all that apply) (1) Interacted with friends (2) Played Ma-jong, played chess, played cards, or went to community club (3) Provided help to family, friends, or neighbors who do not live with you and who did not pay you for the help (4) Went to a sport, social, or other kind of club (5) Took part in a community-related organization (6) Done voluntary or charity work (7) Cared for a sick or disabled adult who does not live with you and who did not pay you for the help (8) Attended an educational or training course (9）Stock investment (10）Used the Internet (11）other (12) None of these	-
Smoke	DA061	Do you still have the habit or have you totally quit? (1) Still have (2) Quit	-
Drink	DA067	Did you drink any alcoholic beverages, such as beer, wine, or liquor in the past year? How often? (1) Drink more than once a month (2) Drink but less than once a month	-

Table A-21 Summary of original questions of control variables-4

Variable	Number	Question and options	Recode
Commercial insurance	EA001	Are you the policy holder/primary beneficiary of any of the types of health insurance listed below? (circle all that apply) (1) Urban employee medical insurance (yi-bao) (2) Urban resident medical insurance (3) New cooperative medical insurance (he-zuo-yi-liao) (4) Urban and rural resident medical insurance (5) Government medical insurance (gong-fei) (6) Medical aid (7) Private medical Insurance: Purchased by R's union (8) Private medical Insurance: Purchased by Individual (9) Other medical insurance (specify) (10)No insurance	1= "Private medical Insurance" 0=other cases
Family income	GA005/ GA006	Did [preload household member name] receive any wage and bonus income in the past year? How much did he/she receive last year?	-
	GA007/ GA008	Did [preload household member name] receive any of the following types of individual income in the past year? (check all that apply) How much did he/she receive last year? (1) Pensions (including wages from government institutions and firms, supplemental pension of the firms, and income from such programs as rural pension insurance, Urban residents' pension and commercial pension insurance, and pension subsidy for the oldest old) (2) Unemployment compensation (3) Pension subsidy (4) Workers' compensation from Industrial Accident Compensation Insurance includes wage-replacement benefits, disability benefits, and survivors' benefits (5) elderly family planning subsidies (6) medical aid (7) other government subsidies (8) social assistance (9) other income sources (including alimony, child support) (10) None of the above	-
	GB005	What is the value of all crops and forestry products (including sold lumber) produced in the past year?	-
	GB011	What was the value of all livestock and aquatic life that were sold or consumed in the past year?	-
	GB012	What was the value of all livestock products produced （including the self-consumption value in the past year, including milk, wool (including cashmere, sheep or goat skin), and eggs)	-

Table A-22 Summary of original questions of control variables-5

Variable	Number	Question and options	Recode
Family income	GC005	Not including fixed capital costs, what is your best estimate of the net income earned from this activity by your household members last year?[If the activity was conducted jointly with non-household members, report only the net income earned by household members. Remember to consider the following types of costs: energy, housing or equipment rental, raw materials, transportation, marketing, wages, taxes or fees.]	-
	GD001	How much Dibao assistance did your household receive last year? (if not applicable, fill in 0 yuan).	-
	GD002	Did your household receive any of the following government subsidies in the past year? (check all that apply) How much? (1) Reforestation (2) Agricultural subsidies (3) Wubaohu (targets low-income, blind, disabled, aged persons, and young persons that have no means to support themselves (4) Tekunhu (5) Work injury subsidies to the immediate family members (6) Emergency or disaster relief (jiujikuan, jiuzaikuan) last year? (7) Other (8) None	-
	GD003	Did your household receive any income from the following sources in the past year? (check all that apply) How much? (1) Donations from the society (including cash, and items like food, clothing, etc.) (2) Compensation for land seizure last? (3) Compensation to pulling down your house or apartment last year? (4) None	-

Table A-23 DID Regression result summary of outpatient visit occurrence

Variable	Model					
	logit		xtlogit		probit	
DID	0.13***	(0.05)	0.11*	(0.06)	0.07**	(0.03)
Urban	-0.06	(0.04)	N/A	N/A	-0.03	(0.02)
Age	-0.00	(0.00)	N/A	N/A	-0.00	(0.00)
Sex	-0.11***	(0.03)	N/A	N/A	-0.06***	(0.02)
Education	0.01	(0.02)	-0.08	(0.06)	0.01	(0.01)
Marriage or cohabitation	-0.06	(0.04)	-0.14	(0.10)	-0.03	(0.02)
Self-report health	0.40***	(0.01)	0.28***	(0.02)	0.22***	(0.01)
Number of disabilities	0.08***	(0.02)	0.07**	(0.03)	0.05***	(0.01)
Number of chronic diseases	0.21***	(0.01)	0.08***	(0.02)	0.12***	(0.01)
Number of cancers	0.14*	(0.08)	0.25**	(0.12)	0.09**	(0.05)
Kind of social activities	0.11***	(0.01)	0.06***	(0.02)	0.06***	(0.01)
Smoke	-0.21***	(0.03)	-0.11	(0.08)	-0.12***	(0.02)
Drink	-0.13***	(0.03)	-0.03	(0.06)	-0.07***	(0.02)
Family income	0.00**	(0.00)	0.00**	(0.00)	0.00**	(0.00)
Commercial insurance	0.11	(0.08)	0.06	(0.13)	0.06	(0.05)
Constant	-2.59***	(0.36)	N/A	N/A	-1.52***	(0.21)
Observations	51,745		21,119		51,745	
City FE	YES		NO		YES	
Year FE	YES		YES		YES	
Individual FE	NO		YES		NO	
Pseudo R-squared	0.0798		0.0307		0.0798	
Number of id	N/A		6,334		N/A	
AIC	48,963		15,184		48,962	
BIC	50,220		15,303		50,219	
Final choice					√	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; N/A-Not Applicable; xt-panel model

Table A-24 DID Regression result summary of hospitalisation occurrence

Variable	Model					
	logit		xtlogit		probit	
DID	0.01	(0.06)	-0.02	(0.07)	0.00	(0.03)
Urban	0.08*	(0.04)	N/A	N/A	0.04*	(0.02)
Age	0.02***	(0.00)	N/A	N/A	0.01***	(0.00)
Sex	0.35***	(0.04)	N/A	N/A	0.19***	(0.02)
Education	-0.00	(0.02)	-0.04	(0.07)	-0.00	(0.01)
Marriage or cohabitation	-0.00	(0.04)	-0.44***	(0.13)	-0.01	(0.02)
Self-report health	0.49***	(0.02)	0.32***	(0.02)	0.26***	(0.01)
Number of disabilities	0.16***	(0.02)	0.13***	(0.03)	0.10***	(0.01)
Number of chronic diseases	0.20***	(0.01)	0.08***	(0.02)	0.11***	(0.01)
Number of cancers	0.75***	(0.09)	0.54***	(0.13)	0.45***	(0.05)
Kind of social activities	0.01	(0.02)	-0.03	(0.02)	0.01	(0.01)
Smoke	-0.41***	(0.04)	-0.66***	(0.09)	-0.22***	(0.02)
Drink	-0.30***	(0.04)	-0.30***	(0.07)	-0.16***	(0.02)
Family income	-0.00	(0.00)	-0.00	(0.00)	-0.00	(0.00)
Commercial insurance	0.12	(0.10)	0.31*	(0.16)	0.08	(0.05)
Constant	-6.35***	(0.42)	N/A	N/A	-3.55***	(0.23)
Observations	51,872		15,212		51,872	
City FE	YES		NO		YES	
Year FE	YES		YES		YES	
Individual FE	NO		YES		NO	
Pseudo R-squared	0.104		0.0864		0.103	
Number of id	N/A	N/A	4,577		N/A	N/A
AIC	37,274		10,219		37,292	
BIC	38,531		10,333		38,550	
Final choice	√					

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; N/A-Not Applicable; xt-panel model

Table A-25 DID Regression result summary of Length of Stay (LOS)

Variable	Model							
	poisson		nbreg		xtpoisson		xtnbreg	
DID	-0.01	(0.07)	-0.03	(0.05)	-0.03	(0.03)	-0.02	(0.05)
Urban	0.03	(0.05)	0.04	(0.04)	N/A		N/A	
Age	-0.01***	(0.00)	-0.00***	(0.00)	N/A		N/A	
Sex	0.20***	(0.04)	0.18***	(0.03)	N/A		N/A	
Education	-0.02	(0.02)	-0.01	(0.02)	0.09***	(0.03)	0.04	(0.04)
Marriage or cohabitation	-0.13**	(0.05)	-0.11***	(0.04)	0.55***	(0.05)	0.33***	(0.07)
Self-report health	0.05***	(0.02)	0.05***	(0.01)	-0.03***	(0.01)	0.01	(0.02)
Number of disabilities	0.06**	(0.03)	0.06***	(0.02)	0.03***	(0.01)	0.01	(0.02)
Number of chronic diseases	-0.01	(0.01)	-0.01	(0.01)	0.01**	(0.01)	0.02	(0.01)
Number of cancers	0.36***	(0.05)	0.38***	(0.07)	0.12***	(0.03)	0.08	(0.06)
Kind of social activities	-0.03*	(0.02)	-0.03**	(0.01)	-0.03**	(0.01)	-0.02	(0.02)
Smoke	-0.07*	(0.04)	-0.05	(0.03)	-0.01	(0.04)	0.00	(0.06)
Drink	-0.15***	(0.03)	-0.13***	(0.03)	-0.17***	(0.03)	-0.13**	(0.06)
Family income	-0.00*	(0.00)	-0.00**	(0.00)	-0.00***	(0.00)	-0.00	(0.00)
Commercial insurance	-0.12	(0.08)	-0.07	(0.07)	0.07	(0.07)	0.08	(0.12)
Constant	2.36***	(0.25)	2.25***	(0.22)	N/A		0.98***	(0.13)
Observations	6,676		6,676		2,652		2,652	
City FE	YES		YES		NO		NO	
Year FE	YES		YES		YES		YES	
Individual FE	NO		NO		YES		YES	
Pseudo R-squared	0.0693		0.0173		N/A		N/A	
Number of id	N/A		N/A		1,179		1,179	
LR test	N/A		p=0.000		N/A		N/A	
Final choice			√					

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; N/A-Not Applicable; xt-panel model

Table A-26 DID Regression result summary of out-of-pocket costs

Outcome	Outpatient visit out-of-pocket cost		Hospitalisation out-of-pocket costs	
Model	glm-gamma		glm-gamma	
DID	-0.33***	(0.12)	-0.20**	(0.09)
Urban	0.16**	(0.08)	0.12**	(0.06)
Age	-0.00	(0.00)	-0.01***	(0.00)
Sex	0.33***	(0.07)	0.16***	(0.05)
Education	0.05	(0.04)	0.06*	(0.03)
Marriage or cohabitation	0.20***	(0.08)	-0.05	(0.07)
Self-report health	0.21***	(0.03)	0.02	(0.02)
Number of disabilities	0.10*	(0.05)	0.11***	(0.04)
Number of chronic diseases	-0.01	(0.02)	-0.01	(0.01)
Number of cancers	0.88***	(0.18)	0.91***	(0.12)
Kind of social activities	-0.07***	(0.03)	0.01	(0.02)
Smoke	-0.49***	(0.08)	-0.28***	(0.06)
Drink	-0.10	(0.08)	-0.02	(0.06)
Family income	-0.00*	(0.00)	0.00	(0.00)
Commercial insurance	0.09	(0.18)	-0.09	(0.15)
Constant	3.66***	(0.52)	10.05***	(0.63)
Observations	9,350		5,733	
City FE	YES		YES	
Year FE	YES		YES	
Individual FE	N/A		N/A	
R-squared	N/A		N/A	
Number of id	N/A		N/A	
Final choice	√		√	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; N/A-Not Applicable;

Table A-27 Result summary of parallel trend test of health service use

Outcome	Outpatient visit occurrence		Hospitalisation occurrence		Length of stay	
Variables	probit		logit		nbreg	
7 years ago	-0.11	(0.08)	-0.26	(0.17)	-0.29*	(0.15)
6 years ago	-0.13*	(0.07)	-0.30*	(0.17)	-0.1	(0.17)
5 years ago	**-0.16****	(0.07)	-0.23	(0.15)	-0.18	(0.14)
4 years ago	-0.06	(0.07)	-0.15	(0.15)	-0.12	(0.15)
3 years ago	-0.05	(0.06)	**-0.26****	(0.13)	-0.18	(0.12)
2 years ago	-0.04	(0.06)	-0.19	(0.13)	-0.11	(0.14)
Year of implementation	0.01	(0.06)	-0.24*	(0.12)	-0.13	(0.13)
1 year later	0.02	(0.06)	-0.13	(0.14)	-0.12	(0.15)
2 years later	0.11*	(0.07)	-0.31**	(0.14)	-0.02	(0.14)
3 years later	0.07	(0.07)	-0.29*	(0.15)	-0.07	(0.16)
4 years later	0.08	(0.08)	-0.21	(0.17)	-0.21	(0.16)
5 years later	0.16	(0.11)	-0.48**	(0.23)	-0.19	(0.24)
6 years later	0.01	(0.09)	-0.15	(0.18)	-0.12	(0.17)
7 years later	0.01	(0.11)	-0.32	(0.21)	0.02	(0.20)
Constant	-1.46***	(0.21)	-6.08***	(0.43)	2.36***	(0.25)
Observations	51,745		51,872		6,676	
City FE	YES		YES		YES	
Year FE	YES		YES		YES	
Control Variables	YES		YES		YES	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; probit-probit model; logit-logit model; nbreg-negative binomial model

Table A-28 Result summary of parallel trend test of out-of-pocket costs

Outcome	Outpatient visit out-of-pocket costs		Hospitalisation out-of-pocket costs	
Variables	Glm-gamma		Glm-gamma	
7 years ago	-0.11	(0.31)	-0.18	(0.27)
6 years ago	0.09	(0.34)	0.47	(0.30)
5 years ago	-0.23	(0.29)	-0.00	(0.24)
4 years ago	-0.09	(0.30)	0.17	(0.26)
3 years ago	-0.18	(0.25)	0.07	(0.22)
2 years ago	0.22	(0.28)	0.17	(0.23)
Year of implementation	-0.36	(0.25)	-0.06	(0.21)
1 year later	0.01	(0.32)	-0.08	(0.26)
2 years later	-0.43	(0.27)	0.08	(0.24)
3 years later	-0.31	(0.32)	-0.27	(0.26)
4 years later	-0.45	(0.31)	-0.34	(0.27)
5 years later	-0.02	(0.38)	0.02	(0.41)
6 years later	0.05	(0.32)	-0.22	(0.32)
7 years later	0.07	(0.42)	-0.09	(0.35)
Constant	3.95***	(0.59)	9.84***	(0.69)
Observations	9,350		5,733	
City FE	YES		YES	
Year FE	YES		YES	
Control Variables	YES		YES	

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A-29 Summary of descriptive statistics of sampling for placebo test in Chapter 8

Outcomes	Coefficients		P-value	
	Mean	Standard Deviation	Mean	Standard Deviation
Outpatient occurrence	0.0006	0.0132	0.989	0.0001
Outpatient out-of-pocket costs	-0.0049	0.0771	0.954	0.0015
Hospitalisation occurrence	-0.0007	0.0287	0.977	0.0001
Length of stay	-0.0029	0.0275	0.980	0.0004
Hospitalisation out-of-pocket costs	-0.0005	0.0500	0.966	0.0005

Table A-30 Regression result summary of DDD estimations

Outcome	Outpatient visit occurrence		Hospitalisation occurrence	
Model	probit		logit	
DDD	0.10*	(0.06)	-0.04	(0.13)
Resident	-0.2	(0.22)	-0.06	(0.47)
Urban	-0.02	(0.02)	0.18***	(0.04)
Age	0.00	(0.00)	0.03***	(0.00)
Sex	-0.05***	(0.02)	0.34***	(0.03)
Education	0.02**	(0.01)	0.02	(0.02)
Marriage or cohabitation	-0.03	(0.02)	0.01	(0.04)
Self-report health	0.22***	(0.01)	0.49***	(0.01)
Number of disabilities	0.05***	(0.01)	0.17***	(0.02)
Number of chronic diseases	0.13***	(0.00)	0.21***	(0.01)
Number of cancers	0.09**	(0.04)	0.71***	(0.07)
Kind of social activities	0.06***	(0.01)	0.01	(0.01)
Smoke	-0.13***	(0.02)	-0.40***	(0.04)
Drink	-0.10***	(0.02)	-0.33***	(0.03)
Family income	0.00***	(0.00)	0.00	(0.00)
Commercial insurance	0.01	(0.04)	0.13*	(0.08)
Constant	-1.29***	(0.18)	-7.25***	(0.58)
Observations	69,934		70,018	
City fixed effect	YES		YES	
Year fixed effect	YES		YES	
Resident * City	YES		YES	
Resident * Year	YES		YES	
City * Year	YES		YES	
Pseudo R-squared	0.0927		0.119	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Appendix G: Code case in Chapter 8

***Input data and process**

```
set more off
use "processed\all.dta", clear
```

***Bundle control variables**

```
global convar1 i.urban_nbs age sex edu maliving srh n_disa n_chro n_canc n_social smoke
    drink incomeh_am oinsu2 //for sectional data model
global convar2 edu maliving srh n_disa n_chro n_canc n_social smoke drink incomeh_am
    oinsu2 // for panel data model
```

***DID estimation**

```
*set panel
xtset id year

qui: logit outa i.did i.year i.cit $convar1 if insu==1, vce (cluster id) //for resident group only
    outreg2 using "est_results\outpatient_r.xls", replace bdec(2) sdec(2) ///
    ctitle(outa_logit) keep(i.did $convar1) addtext(City FE, YES, Year FE, Yes) ///
    addstat(Pseudo R-squared, `e(r2_p)')
    estat ic
    margins, dydx(i.did) post
    outreg2 using "est_results\outpatient_m.xls", replace ctitle(outa_logit)
qui: xtlogit outa i.did i.year $convar2 if insu==1,fe
    outreg2 using "est_results\outpatient_r.xls", append bdec(2) sdec(2) ///
    ctitle(outa_xtlogit) keep(i.did $convar2) addtext(Individual FE, YES, Year FE, Yes) ///
    addstat(Pseudo R-squared, `e(r2_p)')
    estat ic
    margins, dydx(i.did) post
    outreg2 using "est_results\outpatient_m.xls", append ctitle(outa_xtlogit)
```

```
qui: probit outa i.did i.year i.cit $convar1 if insu==1, vce (cluster id)
    outreg2 using "est_results\outpatient_r.xls", append bdec(2) sdec(2) ///
    ctitle(outa_probit) keep(i.did $convar1) addtext(City FE, YES, Year FE, Yes) ///
    addstat(Pseudo R-squared, `e(r2_p)')
    estat ic
    margins, dydx(i.did) post
    outreg2 using "est_results\outpatient_m.xls", append ctitle(outa_probit)
```

***Parallel trend test**

```
*generate test variables
gen dist=year-inteyear
order year inteyear dist
tab dist,m //9 year before integration, 11 year after integraiton
forvalues i=1/9 {
    gen b_`i'= (inte==1 & dist==-`i') //inte==1 means individual in integration group
}
forvalues j=1/11 {
    gen a_`j'=(inte==1 & dist==`j')
}
gen im=0
replace im=1 if inte==1&dist==0
replace b_7=1 if b_9==1|b_8==1
replace a_7=1 if a_8==1|a_9==1|a_10==1|a_11==1
*set 1-year before integration as comparator
global cttest7 b_7 b_6 b_5 b_4 b_3 b_2 im a_1 a_2 a_3 a_4 a_5 a_6 a_7

qui:probit outa $cttest7 i.year i.cit $convar1 if insu==1, vce(cluster id) //for resident group only
    outreg2 using "test\ct7.xls", replace bdec(2) sdec(2) ctitle(outa_probit) ///
    keep($cttest7) addtext(City FE, YES, Year FE, YES, Controls,YES)
    coefplot,keep($cttest7) levels(95) vertical coeflabels(b_7=-7 b_6=-6 ///
```

```
    b_5=-5 b_4=-4 b_3=-3 b_2=-2) title("Common Trend Test, Outpatient Visit Occurrence")
///
    yline(0, lpattern(dash) lcolor(black) lwidth(thin)) ///
    ytitle("Coefficients") xtitle("Year") msymbol(O) msize(small) mcolor(gs1) ///
    addplot(line @b @at, lcolor(black) lwidth(medthick) lpattern(solid)) ///
    ciopts(recast(rcap) lcolor(black) lwidth(thin)) scheme(s1mono)
    graph export "test\ct7-outa.png",as(png) replace width(800) height(600)
```

***Placebo test**

```
*Monte Carlo simulation
permute did beta=_b[did] se=_se[did], reps(500) rseed(180135271) ///
    saving("processed\simulations_outa.dta",replace): ///
    probit outa did i.cit i.year $convar1 if insu==1, vce(cluster id)

*Drawing
use "processed\simulations_outa.dta",clear
gen z_value=beta/se
gen p_value=2*(1-normal(se))
sum(beta)
kdensity beta, scheme(s1mono) ///
    title("") subtitle("Sampling Coefficient Distribution",size(5)) note("") ///
    xtitle("Coefficients",size(5)) ytitle ("Density",size(5)) xlabel(-0.04(0.02)0.08) ///
    xline(`r(mean)', lpattern(dash)) xline(0.0689366, lpattern(solid) lc(red)) ///
    saving (pltest_beta,replace)
sum(p_value)
kdensity p_value, scheme(s1mono) ///
    title("") subtitle("Sampling p_value Distribution",size(5)) note("") ///
    xtitle("p_value",size(5)) ytitle ("")    ///
    xline(`r(mean)', lpattern(dash)) ///
    saving (pltest_p,replace)
```

```
graph combine pltest_beta.gph pltest_p.gph, scheme(s1mono) ///
    title("Placebo Test, Outpatient Visit Occurrence",size(5)) ///
    note("Note: Vertical dash line-the mean value of sampling;solid line-the actual value",size(3))
    graph export "test\pltest-1.png",as(png) replace width(800)
```

***DDD estimation**

```
qui: probit outa i.ddd i.insu##i.cit i.cit##i.year i.insu##i.year $convar1, vce (cluster id)
    outreg2 using "est_results\ddd_r.xls", replace bdec(2) sdec(2) ///
    ctitle(outa_probit) keep(i.ddd i.insu $convar1) addtext(City FE, YES, Year FE, YES, Individual FE, YES) ///
    addstat(Pseudo R-squared, `e(r2_p)')
    margins, dydx(i.ddd) post
```

Appendix H: Supplemental Tables for Chapter 9

Table A-31 Comparison of macro characteristics of cities selected in EMR and CHARLS

Database	City	GDP per capita (CNY)	Health institutions per 10,000 people	Urban PCDI (CNY)	Proportion of resident insurance
EMR	A	39,352	10.86	29,996	77.44%
	B	58,744	12.23	27,420	67.54%
	C	35,560	7.86	33,823	92.15%
	D*	119,418	5.04	39,629	61.47%
	E	18,676	12.34	27,448	90.92%
	F	43,999	3.87	31,230	85.10%
CHARLS	Haozhou*	24,547	3.61	29,711	95.23%
	Anqing*	41,088	4.45	31,187	90.42%
	Chaoyang	28,266	13.11	25,462	76.93%
	Yulin#	112,845	11.29	31,317	86.56%
	Anshan	48,810	4.93	35,619	62.32%

Note: GDP-Gross Domestic Product; PCDI-Per Capita Disposable Income
*The number of resident insurance enrolled has not been published, so the number of employee insurance enrolled is used for calculation
#Only the number of people in the new rural cooperative medical scheme is published, so the proportion of the total number of residents insured should be higher

Table A-32 Trend test of total costs for treatment and control groups of selected cities in CHALRS

City	Outpatient total costs			Hospitalisation total costs		
	Coefficients of group difference change		p value	Coefficients of group difference change		p value
Haozhou	2011	Baseline		2011	Baseline	
	2013	-0.2386	0.853	2013	0.6179	0.586
	2015	-2.1422	0.139	2015	-2.2873	0.078
	2018	-2.3015	0.152	2018	0.3021	0.786
Anqing	2011	Baseline		2011	Baseline	
	2013	0.3374	0.912	2013	1.5212	0.241
	2015	1.1371	0.443	2015	1.2587	0.242
	2018	0.6102	0.284	2018	0.9104	0.138
Chaoyang	2011	Baseline		2011	Baseline	
	2013	0.1632	0.638	2013	-1.6309	0.071
	2015	0.3744	0.431	2015	0.1730	0.778
	2018	0.6902	0.120	2018	-0.6477	0.919
Yulin	2011	Baseline		2011	Baseline	
	2013	-1.0789	0.226	2013	0.5600	0.440
	2015	-0.9354	0.341	2015	1.0666	0.217
	2018	-1.9092	0.068	2018	0.9684	0.146
Anshan	2011	Baseline		2011	Baseline	
	2013	-12.6729	0.503	2013	-1.9952	0.084
	2015	-12.2585	0.594	2015	-3.5305	0.075
	2018	-1.7740	0.694	2018	0.4152	0.631

Note: The covariates used are exactly the same as those used in Chapter 8

Table A-33 DID Regression result summary of outpatient total fees

Variable	Hospital							
	A		B		E		F	
DID	-0.30**	(-0.13)	-0.15***	(-0.03)	-0.08	(-0.05)	-0.27***	(-0.03)
Resident	0.36***	(-0.12)	0.12***	(-0.02)	0.01	(-0.03)	-0.15***	(-0.01)
Sex	0.09***	(-0.02)	0.05***	(-0.01)	0.09***	(-0.01)	0.12***	(-0.01)
Age	-0.00***	(0.00)	0.00***	(0.00)	0	(0.00)	0.01***	(0.00)
Angina pectoris	-0.51***	(-0.06)	0.04	(-0.09)	-0.14*	(-0.07)	-0.32***	(-0.09)
Myocardial infarction	0.68***	(-0.06)	0.90***	(-0.09)	1.26***	(-0.06)	0.38***	(-0.10)
Acute coronary syndrome	1.24***	(-0.32)	-0.07	(-0.07)	-0.14	(-0.10)	0.22**	(-0.11)
Ischemic heart disease	-0.34***	(-0.07)	0.02	(-0.07)	-0.47***	(-0.07)	-0.12	(-0.1)
Cancer	0.3	(-0.36)	0.45**	(-0.21)	-0.13	(-0.24)	0.27	(-0.63)
Heart failure	-0.34**	(-0.16)	0.1	(-0.24)	-0.59***	(-0.15)	0.01	(-0.04)
Hypertension	0.27***	(-0.03)	0.28***	(-0.01)	-0.31***	(-0.05)	0.24***	(-0.01)
Cerebral infarction	0.24***	(-0.05)	0.38***	(-0.02)	-0.18**	(-0.08)	0.17***	(-0.02)
COPD	0.66***	(-0.11)	0.22***	(-0.03)	0.59***	(-0.17)	0.28**	(-0.12)
Diabetes	0.24***	(-0.05)	0.35***	(-0.02)	0.52***	(-0.16)	0.07	(-0.09)
Constant	5.96***	(-0.1)	5.29***	(-0.08)	5.81***	(-0.08)	4.77***	(-0.1)
Observations	18,018		53,578		40,973		49,439	
Year FE	YES		YES		YES		YES	
Comorbidity FE	YES		YES		YES		YES	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; COPD-Chronic Obstructive Pulmonary Disease

Table A-34 DID Regression result summary of hospitalisation total fees

Variable	Hospital							
	A		B		C		D	
DID	0.15***	(0.03)	0.16***	(0.03)	-0.08	(0.08)	0.22***	(0.07)
Resident	-0.06***	(0.02)	-0.32***	(0.02)	-0.04	(0.06)	-0.13**	(0.05)
Sex	0.08***	(0.01)	0.12***	(0.01)	0.04	(0.03)	0.22***	(0.04)
Age	-0.00***	(0.00)	-0.00**	(0.00)	-0.01***	(0.00)	0.00	(0.00)
Angina pectoris	0.13***	(0.02)	0.09***	(0.02)	0.07	(0.09)	0.36***	(0.05)
Myocardial infarction	0.61***	(0.02)	0.73***	(0.03)	0.73***	(0.06)	0.42***	(0.05)
Acute coronary syndrome	0.32***	(0.10)	0.16***	(0.03)	0.22***	(0.08)	0.24	(0.15)
Ischemic heart disease	0.31***	(0.10)	0.13	(0.08)	0.08	(0.08)	-0.09	(0.12)
Cancer	0.03	(0.04)	-0.11	(0.08)	-0.00	(0.12)	0.15	(0.13)
Heart failure	-0.05	(0.04)	0.21***	(0.07)	-0.16***	(0.05)	0.30***	(0.09)
Hypertension	0.01	(0.01)	0.05***	(0.02)	0.13***	(0.03)	0.02	(0.04)
Cerebral infarction	0.11***	(0.01)	0.08***	(0.02)	-0.04	(0.04)	0.04	(0.04)
COPD	-0.04*	(0.02)	-0.00	(0.02)	0.00	(0.05)	0.14***	(0.05)
Diabetes	0.02	(0.02)	-0.05*	(0.03)	0.12**	(0.05)	-0.04	(0.04)
Constant	8.61***	(0.12)	8.26***	(0.09)	9.61***	(0.15)	8.50***	(0.17)
Observations	23,298		26,805		5,406		3,587	
Year FE	YES		YES		YES		YES	
Comorbidity FE	YES		YES		YES		YES	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; COPD-Chronic Obstructive Pulmonary Disease

Table A-35 DID Regression result summary of LOS-1

Variable	Hospital							
	A-Poisson		A-Negative Binomial		B-Poisson		B-Negative Binomial	
DID	0.09***	(0.01)	0.09***	(0.01)	0.07***	(0.01)	0.07***	(0.01)
Resident	-0.07***	(0.01)	-0.07***	(0.01)	-0.16***	(0.00)	-0.16***	(0.01)
Sex	0.01	(0.00)	0.01	(0.01)	0.02***	(0.00)	0.02***	(0.01)
Age	-0.00	(0.00)	-0.00	(0.00)	-0.00***	(0.00)	-0.00***	(0.00)
Angina pectoris	-0.07***	(0.01)	-0.06***	(0.01)	0.00	(0.01)	0.00	(0.01)
Myocardial infarction	0.06***	(0.01)	0.06***	(0.01)	0.00	(0.01)	0.01	(0.01)
Acute coronary syndrome	-0.18***	(0.04)	-0.17***	(0.05)	0.06***	(0.01)	0.06***	(0.01)
Ischemic heart disease	0.12***	(0.04)	0.12**	(0.05)	0.01	(0.02)	0.02	(0.03)
Cancer	0.03**	(0.01)	0.03	(0.02)	-0.01	(0.02)	-0.01	(0.03)
Heart failure	0.04***	(0.01)	0.04**	(0.02)	0.05**	(0.02)	0.04*	(0.02)
Hypertension	0.01	(0.00)	0.01	(0.01)	0.04***	(0.00)	0.04***	(0.01)
Cerebral infarction	0.06***	(0.00)	0.06***	(0.01)	0.06***	(0.00)	0.06***	(0.01)
COPD	0.01	(0.01)	0.01	(0.01)	-0.01	(0.01)	-0.01	(0.01)
Diabetes	0.03***	(0.01)	0.03***	(0.01)	0.02***	(0.01)	0.02**	(0.01)
Constant	2.27***	(0.04)	2.25***	(0.06)	2.37***	(0.03)	2.37***	(0.03)
Observations	23,276		23,276		26,805		26,805	
Year FE	YES		YES		YES		YES	
Comorbidity FE	YES		YES		YES		YES	
LR test		p=0.000				p=0.000		
Final choice			√				√	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; COPD-Chronic Obstructive Pulmonary Disease

Table A-36 DID Regression result summary of LOS-2

Variable	Hospital							
	C-Poisson		C-Negative Binomial		D-Poisson		D-Negative Binomial	
DID	-0.07***	(0.02)	-0.07*	(0.04)	0.04**	(0.02)	0.03	(0.04)
Resident	0.04***	(0.02)	0.05	(0.03)	-0.07***	(0.02)	-0.07**	(0.03)
Sex	-0.01	(0.01)	-0.00	(0.02)	0.01	(0.01)	0.01	(0.02)
Age	0.00	(0.00)	0.00	(0.00)	-0.00***	(0.00)	-0.00*	(0.00)
Angina pectoris	-0.12***	(0.03)	-0.10**	(0.05)	0.07***	(0.01)	0.06**	(0.03)
Myocardial infarction	0.01	(0.02)	-0.01	(0.03)	0.06***	(0.01)	0.05**	(0.02)
Acute coronary syndrome	-0.17***	(0.02)	-0.14***	(0.05)	-0.03	(0.04)	-0.01	(0.08)
Ischemic heart disease	-0.03	(0.02)	-0.03	(0.05)	0.05	(0.03)	0.04	(0.06)
Cancer	-0.07**	(0.03)	-0.06	(0.07)	0.08**	(0.03)	0.08	(0.07)
Heart failure	-0.09***	(0.01)	-0.08***	(0.03)	0.07**	(0.03)	0.07	(0.05)
Hypertension	0.12***	(0.01)	0.11***	(0.02)	-0.01	(0.01)	-0.01	(0.02)
Cerebral infarction	0.11***	(0.01)	0.12***	(0.02)	0.06***	(0.01)	0.06***	(0.02)
COPD	0.03**	(0.01)	0.03	(0.02)	0.07***	(0.02)	0.07**	(0.03)
Diabetes	0.17***	(0.01)	0.15***	(0.03)	-0.05***	(0.01)	-0.05**	(0.02)
Constant	2.06***	(0.04)	2.04***	(0.08)	2.56***	(0.05)	2.56***	(0.09)
Observations	5,343		5,343		3,570		3,570	
Year FE	YES		YES		YES		YES	
Comorbidity FE	YES		YES		YES		YES	
LR test	p=0.000				p=0.000			
Final choice			√				√	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; COPD-Chronic Obstructive Pulmonary Disease

Table A-37 Pre-integration parallel trend test for outpatient total fees

Variables	Hospitals					
	B		E		F	
4 years ago	-		-		-0.01	(0.03)
3 years ago	-		-		0.00	(0.03)
2 years ago	0.45***	(0.06)	0.01	(0.07)	-0.09***	(0.03)
Year of implementation	-0.04	(0.04)	-0.08	(0.05)	-0.27***	(0.04)
1 year later	-0.08*	(0.04)	-		-	
Constant	5.27***	(0.08)	5.81***	(0.08)	4.73***	(0.09)
Observations	53,578	0	40,973	0	55,266	0
Year FE	YES		YES		YES	
Control Variables	YES		YES		YES	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A-38 Pre-integration parallel trend test for hospitalisation total fees

Variables	Hospitals					
	A		B		C	
4 years ago	-0.10*	(0.05)	-0.04	(0.05)	-	
3 years ago	-0.04	(0.05)	-0.07	(0.05)	-	
2 years ago	0.01	(0.05)	-0.02	(0.05)	0.01	(0.12)
Year of implementation	0.08	(0.05)	0.14***	(0.04)	-0.16	(0.10)
1 year later	0.12***	(0.04)	0.12**	(0.05)	0.06	(0.11)
2 year later	0.12**	(0.05)	-		-	
Constant	8.80***	(0.12)	8.26***	(0.09)	9.62***	(0.15)
Observations	21,646		26,805		5,406	
Year FE	YES		YES		YES	
Control Variables	YES		YES		YES	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A-39 Pre-integration parallel trend test for LOS

Variables	Hospitals					
	A		B		C	
4 years ago	-0.04*	(0.03)	-0.02	(0.02)	-	
3 years ago	-0.03	(0.02)	-0.01	(0.02)	-	
2 years ago	0.04	(0.02)	0.03*	(0.02)	0.01	(0.07)
Year of implementation	0.06**	(0.02)	0.05***	(0.02)	-0.10*	(0.05)
1 year later	0.09***	(0.02)	0.11***	(0.02)	-0.01	(0.06)
2 year later	0.10***	(0.02)	-		-	
Constant	2.26***	(0.06)	2.37***	(0.04)	2.04***	(0.08)
Observations	21,624	0	26,805	0	5,343	0
Year FE	YES		YES		YES	
Control Variables	YES		YES		YES	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A-40 Summary of descriptive statistics of sampling for placebo test in Chapter 9

Outcomes	Hospitals	Coefficients		P-value	
		Mean	Standard Deviation	Mean	Standard Deviation
Outpatient fees	A	-0.00247	0.0482	0.964	2.56e-05
	B	0.000429	0.0222	0.982	1.87e-05
	E	0.000362	0.0381	0.970	3.57e-05
	F	0.000310	0.0274	0.980	5.37e-06
Hospitalisation fees	A	0.000346	0.0132	0.989	7.82e-06
	B	0.000695	0.0178	0.985	1.18e-05
	C	0.000251	0.0329	0.975	7.51e-05
	D	0.00140	0.0394	0.969	8.59e-05
LOS	A	0.000606	0.00764	0.995	4.87e-06
	B	0.000518	0.00764	0.994	8.37e-06
	C	0.00213	0.0243	0.986	1.79e-05
	D	0.00107	0.0286	0.983	4.31e-05

Table A-41 PSM-DID Regression result summary

Variable	Hospitals and outcomes							
	A outpatient fees		B outpatient fees		D hospitalisation fees		D LOS	
Treated	-0.34**	(0.14)	-0.12***	(0.04)	0.22***	(0.07)	0.03	(0.04)
Resident	0.41***	(0.13)	0.12***	(0.03)	-0.13**	(0.05)	-0.07**	(0.03)
Sex	0.07**	(0.03)	0.07***	(0.01)	0.22***	(0.04)	0.01	(0.02)
Age	-0.01***	(0.00)	-0.00	(0.00)	0.00	(0.00)	-0.00**	(0.00)
Angina pectoris	-0.74***	(0.21)	0.13	(0.24)	0.36***	(0.05)	0.06**	(0.03)
Myocardial infarction	0.86***	(0.27)	0.90***	(0.19)	0.42***	(0.05)	0.05**	(0.02)
Acute coronary syndrome	N/A		0.03	(0.21)	0.24	(0.15)	-0.01	(0.08)
Ischemic heart disease	0.12	(0.31)	0.11	(0.21)	-0.08	(0.12)	0.03	(0.06)
Cancer	0.09	(0.87)	0.75	(0.78)	0.14	(0.13)	0.08	(0.07)
Heart failure	-0.43	(0.40)	0.06	(1.04)	0.30***	(0.10)	0.07	(0.05)
Hypertension	0.19***	(0.05)	0.28***	(0.01)	0.02	(0.04)	-0.01	(0.02)
Cerebral infarction	0.32**	(0.14)	0.44***	(0.06)	0.04	(0.04)	0.06***	(0.02)
COPD	0.28	(0.29)	0.15*	(0.09)	0.14***	(0.05)	0.07**	(0.03)
Diabetes	0.36***	(0.10)	0.45***	(0.12)	-0.04	(0.04)	-0.05**	(0.02)
Constant	5.72***	(0.33)	5.41***	(0.21)	8.50***	(0.17)	2.58***	(0.09)
Observations	10,378		26,066		3,568		3,551	
Year FE	YES		YES		YES		YES	
Comorbidity FE	YES		YES		YES		YES	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; N/A-Not Applicable; COPD-Chronic Obstructive Pulmonary Disease

Table A-42 Dynamic effects regression result summary-1

Variable	Hospitals and outcomes							
	A outpatient fee		B outpatient fee		A hospitalisation fee		B hospitalisation fee	
Treated	-0.43***	(0.17)	-0.11**	(0.05)	0.11***	(0.04)	0.17***	(0.04)
Differential effect in 2nd year	0.14	(0.12)	-0.04	(0.06)	0.03	(0.04)	-0.02	(0.05)
Differential effect in 3rd year	0.08	(0.14)	N/A		0.04	(0.05)	N/A	
Resident	0.41***	(0.13)	0.12***	(0.03)	-0.04**	(0.02)	-0.32***	(0.02)
Sex	0.07**	(0.03)	0.07***	(0.01)	0.08***	(0.01)	0.12***	(0.01)
Age	-0.01***	(0.00)	-0.00	(0.00)	-0.00***	(0.00)	-0.00**	(0.00)
Angina pectoris	-0.74***	(0.21)	0.13	(0.24)	0.10***	(0.02)	0.09***	(0.02)
Myocardial infarction	0.85***	(0.27)	0.91***	(0.19)	0.61***	(0.02)	0.73***	(0.03)
Acute coronary syndrome		N/A	0.03	(0.21)	0.34***	(0.09)	0.16***	(0.03)
Ischemic heart disease	0.12	(0.31)	0.11	(0.21)	0.34***	(0.11)	0.13	(0.08)
Cancer	0.08	(0.87)	0.75	(0.78)	0.02	(0.04)	-0.11	(0.08)
Heart failure	-0.43	(0.40)	0.06	(1.04)	-0.05	(0.04)	0.21***	(0.07)
Hypertension	0.19***	(0.05)	0.28***	(0.01)	0.00	(0.01)	0.05***	(0.02)
Cerebral infarction	0.32**	(0.14)	0.44***	(0.06)	0.12***	(0.01)	0.08***	(0.02)
COPD	0.27	(0.29)	0.15*	(0.09)	-0.04*	(0.02)	-0.00	(0.02)
Diabetes	0.36***	(0.10)	0.44***	(0.12)	0.03	(0.02)	-0.05*	(0.03)
Constant	5.72***	(0.33)	5.41***	(0.21)	8.77***	(0.12)	8.26***	(0.09)
Observations		10,378		26,066		21,646		26,805
Year FE	YES		YES		YES		YES	
Comorbidity FE	YES		YES		YES		YES	

Note: Coefficients are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; N/A-Not Applicable; COPD-Chronic Obstructive Pulmonary Disease

Table A-43 Dynamic effects regression result summary-2

Variable	Hospitals and outcomes							
	C hospitalisation fee		A LOS		B LOS		C LOS	
Treated	-0.17*	(0.09)	0.06***	(0.02)	0.05***	(0.01)	-0.11**	(0.05)
Differential effect in 2nd year	0.22**	(0.09)	0.03	(0.02)	0.06***	(0.02)	0.10*	(0.05)
Differential effect in 3rd year	N/A		0.04	(0.02)	N/A		N/A	
Resident	-0.04	(0.06)	-0.07***	(0.01)	-0.16***	(0.01)	0.05	(0.03)
Sex	0.04	(0.03)	0.01	(0.01)	0.02***	(0.01)	-0.00	(0.02)
Age	-0.01***	(0.00)	-0.00	(0.00)	-0.00***	(0.00)	0.00	(0.00)
Angina pectoris	0.07	(0.09)	-0.06***	(0.01)	0.00	(0.01)	-0.10**	(0.05)
Myocardial infarction	0.73***	(0.06)	0.06***	(0.01)	0.01	(0.01)	-0.01	(0.03)
Acute coronary syndrome	0.22***	(0.08)	-0.17***	(0.05)	0.06***	(0.01)	-0.14***	(0.05)
Ischemic heart disease	0.08	(0.08)	0.12**	(0.05)	0.02	(0.03)	-0.03	(0.05)
Cancer	0.00	(0.12)	0.03	(0.02)	-0.01	(0.03)	-0.06	(0.07)
Heart failure	-0.15***	(0.05)	0.04**	(0.02)	0.04*	(0.02)	-0.08***	(0.03)
Hypertension	0.13***	(0.03)	0.01	(0.01)	0.04***	(0.01)	0.11***	(0.02)
Cerebral infarction	-0.03	(0.04)	0.06***	(0.01)	0.06***	(0.01)	0.12***	(0.02)
COPD	0.00	(0.05)	0.01	(0.01)	-0.01	(0.01)	0.03	(0.02)
Diabetes	0.12**	(0.05)	0.03***	(0.01)	0.02**	(0.01)	0.15***	(0.03)
Constant	9.62***	(0.15)	2.25***	(0.06)	2.37***	(0.03)	2.04***	(0.08)
Observations	5,406		23,276		26,805		5,343	
Year FE	YES		YES		YES		YES	
Comorbidity FE	YES		YES		YES		YES	

Note: Coefficient are reported; Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1; N/A-Not Applicable; COPD-Chronic Obstructive Pulmonary Disease

Appendix I: Code case in Chapter 9

***DID estimation (hospital A outpatient fees)**

```
*Input data
use "processed\out_final.dta",clear

*Bundle control variables
global convar sex age dis1 dis2 dis4 dis5 dis6 dis7 dis8 dis11 dis12 ///
dis13 icd1 icd2 icd3 icd4 icd5 icd6 icd7 icd8 icd10 icd11 icd12 icd13 icd14 ///
icd15 icd16 icd17 icd18 icd19 icd20 icd21 icd22 icd23 icd24
*-------------------------------------------------------------------------------
gen treat=1 if year>=2017&resident==1
replace treat=0 if treat==.
qui: glm fee i.treat i.resident i.year $convar if org_code==62, family(gamma) link(log)
outreg2 using "est_results/outfee_did.xls",replace bdec(2) sdec(2) ctitle(A)
margins, dydx(i.treat) post
```

***parallel trend test (hospital B outpatient fees)**

```
*Input data
use "processed\out_final.dta",clear
bys org_code: tab year resident //A unable to test
*-------------------------------------------------------------------------------
tab year,gen(time)

*Bundle control variables
global convar sex age dis1 dis2 dis4 dis5 dis6 dis7 dis8 dis11 dis12 ///
dis13 icd1 icd2 icd3 icd4 icd5 icd6 icd7 icd8 icd10 icd11 icd12 icd13 icd14 ///
icd15 icd16 icd17 icd18 icd19 icd20 icd21 icd22 icd23 icd24
tab year if org_code==212 //integration 2018
```

```
gen b_2=resident*time3
gen b_1=resident*time4
gen im=resident*time5
gen a_1=resident*time6

*Bundle test variables
global cttest b_2 im a_1

*pre-parallel trend test
glm fee $cttest i.resident i.year $convar if org_code==212, family(gamma) link(log)
      outreg2 using "est_test\outfee.xls", replace bdec(2) sdec(2) ctitle(B)
coefplot, keep(b_2) levels(95) vertical coeflabels(b_2=2016) title("Hospital B") ///
      yline(0, lpattern(dash) lcolor(black) lwidth(thin)) ylabel(-0.1(0.1)0.6) ///
      ytitle("Coefficients") xtitle("Year") msymbol(O) msize(small) mcolor(gs1) ///
      addplot(line @b @at, lcolor(black) lwidth(medthick) lpattern(solid)) ///
      ciopts(recast(rcap) lcolor(black) lwidth(thin)) scheme(s1mono) ///
      note("Note: Integration in 2018, contol year is 2017",size(3)) ///
      saving (cttest_out_b,replace)
```

***placebo test (hospital B outpatient fees)**

```
*Input data
use "processed\out_final.dta",clear
keep if org_code==212 //integration 2018
gen treated=1 if resident==1&year>2017
replace treated=0 if treated==.

*Bundle control variables
global convar sex age dis1 dis2 dis4 dis5 dis6 dis7 dis8 dis11 dis12 ///
dis13 icd1 icd2 icd3 icd4 icd5 icd6 icd7 icd8 icd10 icd11 icd12 icd13 icd14 ///
```

```
icd15 icd16 icd17 icd18 icd19 icd20 icd21 icd22 icd23 icd24


*Monte Carlo simulation
permute treated beta=_b[treated] se=_se[treated], reps(500) rseed(180135271) ///
    saving("processed\simulations_outfee_B.dta",replace): ///
    glm fee treated i.resident i.year $convar, family(gamma) link(log)
use "processed\simulations_outfee_B.dta",clear
gen z_value=beta/se
gen p_value=2*(1-normal(se))
outreg2 using "est_test/pltsum.xls", append sum(detail) keep(beta p_value) eqkeep(mean sd)
title(outfee_b)
sum(beta)
kdensity beta, scheme(s1mono) ///
    title("") subtitle("Sampling Coefficient Distribution",size(5)) note("") ///
    xtitle("Coefficients",size(5)) ytitle ("Density",size(5)) xlabel(-0.18(0.05)0.08) ///
    xline(`r(mean)', lpattern(dash)) xline(-0.15, lpattern(solid) lc(red)) ///
    saving (pltest_beta,replace)
sum(p_value)
kdensity p_value, scheme(s1mono) ///
    title("") subtitle("Sampling p_value Distribution",size(5)) note("") ///
    xtitle("p_value",size(5)) ytitle ("") ylabel("") ///
    xline(`r(mean)', lpattern(dash)) ///
    saving (pltest_p,replace)
graph combine pltest_beta.gph pltest_p.gph, scheme(s1mono) ///
    title("Hospital B") saving(pltest_outfee_b,replace)
```

***PSM-DID (hospital B outpatient fees)**

```
forvalues i=2016/2019 {
```

```
    use "processed\out_final.dta",clear
    keep if org_code==212&year==`i'
    global convar sex age dis1 dis2 dis4 dis5 dis6 dis7 dis8 dis10 dis11 ///
    dis12 dis13 icd1 icd2 icd3 icd4 icd5 icd6 icd7 icd8 icd10 icd11 icd12 ///
    icd13 icd14 icd15 icd16 icd17 icd18 icd19 icd20 icd21 icd22 icd23 icd24
    psmatch2 resident $convar, out(fee) logit neighbor(1) common ties
    pstest, both graph scheme(s1mono)
    drop if _weight==.
    save "processed\outfee_212_`i'.dta",replace
}
use "processed\outfee_212_2016.dta",clear
append using "processed\outfee_212_2017.dta"
append using "processed\outfee_212_2018.dta"
append using "processed\outfee_212_2019.dta"
save "processed\outfee_212_psm.dta",replace

*PSM-DID
use "processed\outfee_212_psm.dta",clear
global convar1 sex age dis1 dis2 dis4 dis5 dis6 dis7 dis8 dis11 dis12 ///
dis13 icd1 icd2 icd3 icd4 icd5 icd6 icd7 icd8 icd10 icd11 icd12 icd13 icd14 ///
icd15 icd16 icd17 icd18 icd19 icd20 icd21 icd22 icd23 icd24
gen treat=1 if year>=2018&resident==1
replace treat=0 if treat==.
qui: glm fee i.treat i.resident i.year $convar1, family(gamma) link(log)
outreg2 using "est_results/psmdid.xls",append bdec(2) sdec(2) ctitle(B)
margins, dydx(i.treat) post

*dynamic effect (hospital B outpatient fees)
use "processed\outfee_212_psm.dta",clear
keep if org_code==212
```

```
gen treat=1 if year>=2018&resident==1
replace treat=0 if treat==.
tab year,gen(time)
gen next1=treat*time4 //
global convar sex age dis1 dis2 dis4 dis5 dis6 dis7 dis8 dis11 dis12 ///
dis13 icd1 icd2 icd3 icd4 icd5 icd6 icd7 icd8 icd10 icd11 icd12 icd13 icd14 ///
icd15 icd16 icd17 icd18 icd19 icd20 icd21 icd22 icd23 icd24
qui: glm fee i.treat next1 i.year i.resident $convar, family(gamma) link(log)
outreg2 using "est_results/lageffect.xls",append bdec(2) sdec(2) ctitle(outfee_B)
margins, dydx(i.treat next1) post
```

***quantile effect (hospital B hospitalisation fees)**

```
use "processed\in_final.dta",clear
keep if org_code==212
gen treat=1 if year>=2018&resident==1
replace treat=0 if treat==.
global convar sex age dis1 dis2 dis4 dis5 dis6 dis7 dis8 dis11 dis12 ///
dis13 icd1 icd2 icd3 icd4 icd5 icd6 icd7 icd8 icd10 icd11 icd12 icd13 icd14 ///
icd15 icd16 icd17 icd18 icd19 icd20 icd21 icd22 icd23 icd24
*----------------------------
qui:sqreg fee i.treat i.resident i.year $convar, q(.05 .1 .25 .5 .75 .9 .95)
outreg2 using "est_results\infee_QR_B.xls", replace keep(i.treat) bdec(0) sdec(0)
```

References

Abadie, A., Diamond, A. and Hainmueller, J. (2010) 'Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program', *Journal of the American Statistical Association*, 105(490), pp. 493–505. doi: 10.1198/jasa.2009.ap08746.

Aho, K., Derryberry, D. and Peterson, T. (2014) 'Model selection for ecologists: the worldviews of AIC and BIC', *Ecology*, 95(3), pp. 631–636.

Ai, C. and Norton, E. C. (2003) 'Interaction terms in logit and probit models', *Economics Letters*, 80(1), pp. 123–129. doi: 10.1016/S0165-1765(03)00032-6.

Aldrich, J. H. and Nelson, F. D. (1984) *Linear probability, logit, and probit models*.

Andersen, R. and Newman, J. F. (1973) 'Societal and Individual Determinants of Medical Care Utilization in the United States', *The Milbank Memorial Fund Quarterly. Health and Society*, 51(1), p. 95. doi: 10.2307/3349613.

Angrist, J. D. and Pischke, J.-S. (2009) *Mostly harmless econometrics : an empiricist's companion*. Princeton University Press.

Arrow, K. J. (1963) 'Uncertainty and the Welfare Economics of Medical Care', *The American Economic Review*, 53(5), pp. 941–973.

Arrow, K. J. (1971) *Essays in the Theory of Risk-Bearing*. Chicago: Markham Publishing Co.

Arrow, K. J. (1976) 'Welfare analysis of changes in health coinsurance rates', in Rossett, R. (ed.) *The Role of Health Insurance in the Health Services Sector*. New York: National Bureau of Economic Research. Available at: https://www.nber.org/chapters/c3815.pdf (Accessed: 16 May 2019).

Ashenfelter, O. (1978) 'Estimating the Effect of Training Programs on Earnings', *The Review of Economics and Statistics*, 60(1), p. 47. doi: 10.2307/1924332.

Athey, S. and Imbens, G. W. (2006) 'Identification and Inference in Nonlinear Difference-in-Differences Models', *Econometrica*, 74(2), pp. 431–497. doi: 10.1111/j.1468-0262.2006.00668.x.

Atun, R. *et al.* (2013) 'Universal health coverage in Turkey: enhancement of equity', *The Lancet*, 382(9886), pp. 6–12.

Babiarz, K. S. *et al.* (2010) 'New evidence on the impact of China's New Rural Cooperative Medical Scheme and its implications for rural primary healthcare: multivariate difference-in-difference analysis', *BMJ*, 341, p. c5617. doi: 10.1136/BMJ.C5617.

Baird, K. (2016) 'High Out-of-Pocket Medical Spending among the Poor and Elderly in Nine Developed Countries', *Health Services Research*, 51(4), pp. 1467–1488. doi: 10.1111/1475-6773.12444.

Baird, K. E. (2016) 'The incidence of high medical expenses by health status in seven developed countries', *Health Policy*, 120(1), pp. 26–34. doi: 10.1016/j.healthpol.2015.10.004.

Bazyar, M. *et al.* (2021) 'The experiences of merging health insurance funds in South Korea, Turkey, Thailand, and Indonesia: a cross-country comparative study', *International journal for equity in health*, 20(1), p. 66.

Beck, T., Levine, R. and Levkov, A. (2010) 'Big Bad Banks? The Winners and Losers from Bank Deregulation in the United States', *The Journal of Finance*, 65(5), pp. 1637–1667. doi: 10.1111/j.1540-6261.2010.01589.x.

Berck, P. and Villas-Boas, S. B. (2016) 'A note on the triple difference in economic models', *Applied Economics Letters*, 23(4), pp. 239–242. doi: 10.1080/13504851.2015.1068912.

Besley, T. J. (1988) 'Optimal reimbursement health insurance and the theory of Ramsey taxation', *Journal of Health Economics*, 7(4), pp. 321–336. doi: 10.1016/0167-6296(88)90019-7.

Bilinski, A. and Hatfield, L. A. (2020) *Nothing to See Here? Non-inferiority Approaches to Parallel Trends and Other Model Assumptions*.

Blomqvist, Å. (1997) 'Optimal non-linear health insurance', *Journal of Health Economics*, 16(3), pp. 303–321. doi: 10.1016/S0167-6296(96)00529-2.

Blumenthal, D. and Hsiao, W. (2005) 'Privatization and Its Discontents — The Evolving Chinese Health Care System', *New England Journal of Medicine*, 353(11), pp. 1165–1170. doi: 10.1056/NEJMhpr051133.

Blundell, R. and Dias, M. C. (2009) 'Alternative Approaches to Evaluation in Empirical Microeconomics', *Journal of Human Resources*, 44(3), pp. 565–640. doi: 10.3368/jhr.44.3.565.

Bonhomme, S. and Sauder, U. (2011) 'Recovering Distributions in Difference-in-Differences Models: A Comparison of Selective and Comprehensive Schooling', *The Review of Economics and Statistics*, 93(2), pp. 479–494. doi: 10.1162/REST_a_00164.

Borusyak, K., Jaravel, X. and Spiess, J. (2021) *Revisiting Event Study Designs: Robust and Efficient Estimation*.

Bronzini, R. and Piselli, P. (2016) 'The impact of R&D subsidies on firm innovation', *Research Policy*, 45(2), pp. 442–457. doi: 10.1016/J.RESPOL.2015.10.008.

Buntin, M. B. and Zaslavsky, A. M. (2004) 'Too much ado about two-part models and transformation?: Comparing methods of modeling Medicare expenditures', *Journal of Health Economics*, 23(3), pp. 525–542. doi: 10.1016/J.JHEALECO.2003.10.005.

Burnham, K. P. and Anderson, D. R. (2004) 'Multimodel Inference: Understanding AIC and BIC in Model Selection', *Sociological methods & research*, 33(2), pp. 261–304.

Caldwell, J. C. (1986) 'Routes to low mortality in poor countries.', *Population and development review*, pp. 171–220.

Caliendo, M. and Kopeinig, S. (2008) 'Some practical guidance for the implementation of propensity score matching', *Journal of economic surveys*, 22(1), pp. 31–72.

Callaway, B. and Li, T. (2017) 'Quantile Treatment Effects in Difference in Differences Models with Panel Data', *SSRN Electronic Journal*. doi: 10.2139/ssrn.3013341.

Callaway, B. and Sant'Anna, P. H. C. (2021) 'Difference-in-Differences with multiple time periods', *Journal of Econometrics*, 225(2), pp. 200–230. doi: 10.1016/J.JECONOM.2020.12.001.

Cao, J. (2020) 'From Policy Announcement to Effectiveness: Analysis of the Time Difference in Policy Implementation', *Journal of Public Administration*, (5), pp. 144–159.

Cao, X. *et al.* (2020) 'The epidemic trend and prediction of chronic diseases with high incidence in China from 1990 to 2025', *Chronic disease prevention and control in China*, 28(1), pp. 14–19.

Carrin, G. *et al.* (1999) 'The reform of the rural cooperative medical system in the People's Republic of China: interim experience in 14 pilot counties', *Social Science & Medicine*, 48(7), pp. 961–972. doi: 10.1016/S0277-9536(98)00396-7.

Chaisemartin, C. de, D'Haultfoeuille, X. and Guyonvarch, Y. (2023) 'DID_MULTIPLEGT: Stata module to estimate sharp Difference-in-Difference designs with multiple groups and periods'. Statistical Software Components,Boston College Department of Economics.

Chang, N. *et al.* (2019) 'Analysis on Equity of Residents' Hospitalization Service Utilization Under The Background of Medical Insurance System Integration', *Medicine and Society*, 32(02), pp. 1–4.

Chang, X., Su, Q. and Zhou, C. (2018) 'The Effects of Urban-rural Integrated Medical Insurance System on Health', *Chinese Journal of Population Science*, (6), pp. 103–114.

Changxue, S. *et al.* (2016) *Research on the integration and optimization design of universal heatlh care system in China*. Beijing: China Planning Press.

Chen, C. (2016) *Health Economic Analysis of China's Health Insurance System*. University of York. Available at: http://etheses.whiterose.ac.uk/17451/ (Accessed: 10 March 2019).

Chen, G., Liu, G. G. and Xu, F. (2014) 'The Impact of the Urban Resident Basic Medical Insurance on Health Services Utilisation in China', *PharmacoEconomics*, 32(3), pp. 277–292. doi: 10.1007/s40273-013-0097-7.

Chen, J., Xu, S. and Gao, J. (2020) 'The Mixed Effect of China's New Health Care Reform on Health Insurance Coverage and the Efficiency of Health Service Utilisation: A Longitudinal Approach', *International Journal of Environmental Research and Public Health*, 17(5).

Chen, L. and Wu, H. (2015) 'Research Status and Potential Problems of Differences-in-differences Method in China', *The Journal of Quantitative & Technical Economics*, 32(7), pp. 133–148.

Chen, Y. *et al.* (2020) 'Arrival of Young Talent: The Send-Down Movement and Rural Education in China', *The American Economic Review*, 110(11), pp. 3393–3430.

Chen, Y. and Zhou, L.-A. (2007) 'The long-term health and economic consequences of the 1959–1961 famine in China', *Journal of Health Economics*, 26(5), pp. 659–681.

Chen, Z. *et al.* (2015) 'Analyzing the factors to influence the rural residents' participating in insurance after the integration of basic medical insurance system sampled with Gaoqing County of Shandong Province', *Chinese Health Service Management*, 32(07), pp. 507–509.

Cheng, L. *et al.* (2015) 'The Impact of Health Insurance on Health Outcomes and Spending of the Elderly: Evidence from China's New Cooperative Medical Scheme', *Health Economics*, 24(6), pp. 672–691. doi: 10.1002/hec.3053.

Chernozhukov, V. *et al.* (2013) 'Average and Quantile Effects in Nonseparable Panel Models', *Econometrica*, 81(2), pp. 535–580. doi: 10.3982/ECTA8405.

Chetty, R., Looney, A. and Kroft, K. (2009) 'Salience and Taxation: Theory and Evidence', *The American Economic Review*, 99(4), pp. 1145–1177.

CMS (2019) *Design and development of the Diagnosis Related Group (DRG)*.

Cojocaru, Ștefan (2009) 'Clarifying the theory-based evaluation', *Revista de Cercetare şi Intervenţie Socială*, (26), pp. 76–86.

Crepon, B. *et al.* (2010) *Analyzing the Anticipation of Treatments Using Data on Notification Dates*.

Cunningham, P. J. (2009) 'Chronic burdens: the persistently high out-of-pocket health care expenses faced by many Americans with chronic conditions.', *Issue brief (Commonwealth Fund)*, 63, pp. 1–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19626725 (Accessed: 10 March 2020).

Cutler, D. M. and Zeckhauser, R. J. (2000) 'The Anatomy of Health Insurance', in *Handbook of Health Economics*. Elsevier, pp. 563–643. doi: 10.1016/S1574-0064(00)80170-5.

Dai, J. *et al.* (2018) 'The effect of urban and rural medical insruance integration on the hospitalization expenses of the insurance in a third grade hospital', *Journal of Anhui Health Vocational & Technical College*, 17(06), pp. 1-2,23.

Doolan, D. M. and Froelicher, E. S. (2009) 'Using an existing data set to answer new research questions: A methodological review', *Research and Theory for Nursing Practice: An International Journal*, 23(3), pp. 203–215.

Dor, A., Gertler, P. and van der Gaag, J. (1987) 'Non-price rationing and the choice of medical care providers in rural Cote d'Ivoire', *Journal of Health Economics*, 6(4), pp. 291–304. doi: 10.1016/0167-6296(87)90017-8.

Du, J. (2009) 'Economic reforms and health insurance in China', *Social Science & Medicine*, 69(3), pp. 387–395. doi: 10.1016/J.SOCSCIMED.2009.05.014.

Eaddy, M. T. *et al.* (2012) 'How patient cost-sharing trends affect adherence and outcomes: A literature review', *Pharmacy and Therapeutics*. MediMedia, USA, pp. 45–55.

Eduardo, C. *et al.* (2021) 'Socioeconomic inequalities in low birth weight risk before and during the COVID-19 pandemic in Argentina: A cross-sectional study', *The Lancet Regional Health - Americas*, 2.

Eggleston, K. (2012) 'Health Care for 1.3 Billion: An Overview of China's Health System', *SSRN Electronic Journal*. doi: 10.2139/ssrn.2029952.

Fan, Y. and Yu, Z. (2012) 'Partial identification of distributional and quantile treatment effects in difference-in-differences models', *Economics Letters*, 115(3), pp. 511–515. doi: 10.1016/J.ECONLET.2012.01.001.

Feldman, R. and Dowd, B. (1991) 'A new estimate of the welfare loss of excess health insurance', *The American Economic Review*, 81(1), pp. 297–301. Available at: https://www.jstor.org/stable/2006802 (Accessed: 16 May 2019).

Feldstein, M. and Friedman, B. (1977) 'Tax subsidies, the rational demand for insurance and the health care crisis', *Journal of Public Economics*, 7(2), pp. 155–178. doi: 10.1016/0047-2727(77)90044-5.

La Ferrara, E., Chong, A. and Duryea, S. (2012) 'Soap Operas and Fertility: Evidence from Brazil', *American Economic Journal: Applied Economics*, 4(4), pp. 1–31.

Flintermann, B. (2014) *The quality of market research reports - The case of Marketline Advantage and the automobile industry*. University of Twente.

Freeman, C. and Boynton, X. (2011) *Implementing health care reform policies in China: Challenges and opportunities, USA:Centre for Strategic & International Studies*. Available at: http://csis-prod.s3.amazonaws.com/s3fs-public/legacy_files/files/publication/111202_Freeman_ImplementingChinaHealthReform_Web.pdf (Accessed: 26 March 2019).

Gan, B. (2013) 'China's 2013 Report on the development of migrant Populations released', *Health Daily*, 11 September.

Gao, C. (2018) 'Satisfaction Survey on the Integration of Basic Medical Insurance System of Urban and Rural Residents in Nanjing', *China Collective Economy*, 32, pp. 167–168.

Gao, J. *et al.* (2002) 'Health equity in transition from planned to market economy in China', *Health Policy and Planning*, 17(suppl_1), pp. 20–29. doi: 10.1093/HEAPOL/17.SUPPL_1.20.

GBD (2018) 'Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017:a systematic analysis for the Global Burden of Disease Study 2017', *Lancet*, 392, pp. 1736–1788. doi: 10.1016/S0140-6736(18)32203-7.

Ge, Y. and Gong, S. (2007) *Chinese health care reform: problems, reasons and solutions*. Beijing: China Development Press. Available at: https://scholar.google.co.uk/scholar?hl=en&as_sdt=0%2C5&q=Ge+Y%2C+Gong+S.+China's+health+care+reform%3A+Problems%2C+causes%2C+and+solutions.&btnG= (Accessed: 26 March 2019).

Gebel, M. and Voßemer, J. (2014) 'The impact of employment transitions on health in Germany. A difference-in-differences propensity score matching approach', *Social Science & Medicine*, 108, pp. 128–136. doi: 10.1016/J.SOCSCIMED.2014.02.039.

Gertler, P. J. *et al.* (2016) *Impact Evaluation in Practice*. Second Edi. Washington: World Bank. doi: 10.1596/978-1-4648-0779-4.

Geyndt, W. De, Zhao, X. and Liu, S. (1992) 'From barefoot doctor to village doctor in rural China.' Washington D.C. World Bank 1992. Available at: https://www.popline.org/node/324436 (Accessed: 11 March 2019).

Goodman-Bacon, A. (2021) 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics*, 225(2), pp. 254–277. doi: 10.1016/J.JECONOM.2021.03.014.

Gotsadze, G. *et al.* (2015) 'An impact evaluation of medical insurance for poor in Georgia: preliminary results and policy implications', *Health Policy and Planning*, 30(1), pp. 2–13.

Grossman, M. (1972) 'On the Concept of Health Capital and the Demand for Health', *Journal of Political Economy*, 80(2), pp. 223–255. doi: 10.1086/259880.

Gruber, J. (1994) 'The Incidence of Mandated Maternity Benefits', *The American Economic Review*, 84(3), pp. 622–641.

Hamermesh, D. S. and Trejo, S. J. (2000) 'The Demand for Hours of Labor: Direct Evidence from California', *Review of Economics and Statistics*, 82(1), pp. 38–47. doi: 10.1162/003465300558614.

Hao, G. (2019) *Research on the satisfaction of Shijiazhuang basic medical insurance for urban and rural residents*. Hebei Normal University.

Hao, Y. *et al.* (1998) 'Financing Health Care in Poor Rural Counties in China: Experience from a Township-Based Co-Operative Medical Scheme'. Available at: https://opendocs.ids.ac.uk/opendocs/handle/123456789/3370 (Accessed: 26 March 2019).

Hardin, J. W., Schmiediche, H. and Carroll, R. J. (2003) 'Instrumental Variables, Bootstrapping, and Generalized Linear Models', *The Stata Journal: Promoting communications on statistics and Stata*, 3(4), pp. 351–360. doi: 10.1177/1536867X0300300405.

Havnes, T. and Mogstad, M. (2015) 'Is universal child care leveling the playing field?', *Journal of Public Economics*, 127, pp. 100–114. doi: 10.1016/J.JPUBECO.2014.04.007.

He, H. and Nolen, P. J. (2019) 'The effect of health insurance reform: Evidence from China', *China Economic Review*, 53, pp. 168–179.

Heckman, J. J. and Vytlacil, E. J. (2007) *Chapter 70 Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation*, *Handbook of Econometrics*. Elsevier. doi: 10.1016/S1573-4412(07)06070-9.

Hirano, K., Imbens, G. W. and Ridder, G. (2003) 'Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score', *Econometrica*, 71(4), pp. 1161–1189.

Hoetker, G. (2007) 'The use of logit and probit models in strategic management research: Critical issues', *Strategic management journal*, 28(4), pp. 331–343.

Hsiao, W. C. L. (1995) 'The Chinese health care system: Lessons for other nations', *Social Science & Medicine*, 41(8), pp. 1047–1055. doi: 10.1016/0277-9536(94)00421-O.

Hu, Q. *et al.* (2018) 'Comparison of the Residents'Health Service Needs and Utilization in Different Medical Systems in Ningxia', *Medicine & Philosophy*, 30(9), pp. 57–60.

Huang, F. and Gao, L. (2010) 'Excess Demand or Appropriate Demand?-Helath Insurance, Medical CAre and Mortality of the Elderly in Urban China', *Economic Research Journal*, 6, pp. 105–119.

Huang, W., Zhang, Z. and Liu, A. (2022) 'From difference-in-difference to event study', *Review of Industrial Economics*, 02. Available at: https://doi.org/10.19313/j.cnki.cn10-1223/f.20211227.002.

Huang, Y. and Zhang, J. (2017) 'The Impact of Urban and Rural Medical Insurance Integration on the Utilization of Rural Residents' Medical Services: A Case Study of Guangzhou City', *Chinese Public Policy Review*, 12(01), pp. 34–52.

Ikegami, N. *et al.* (2011) 'Japanese universal health coverage: evolution, achievements, and challenges', *The Lancet*, 378(9796), pp. 1106–1115.

Imbens, G. W. and Lemieux, T. (2008) 'Regression discontinuity designs: A guide to practice', *Journal of Econometrics*, 142(4), pp. 615–635.

Imbens, G. W. and Wooldridge, J. M. (2009) 'Recent Developments in the Econometrics of Program Evaluation', *Journal of Economic Literature*, 47(1), pp. 5–86. doi: 10.1257/jel.47.1.5.

Jacobs, R. and Goddard, M. (2000) *Social Health Insurance Systems in European Countries The Role of the Insurer in the Health Care System: A Comparative Study of Four European Countries*. York, UK. Available at: https://www.york.ac.uk/che/pdf/op39.pdf (Accessed: 2 May 2019).

Jacobson, L. (2000) 'The family as producer of health — an extended grossman model', *Journal of Health Economics*, 19(5), pp. 611–637. doi: 10.1016/S0167-6296(99)00041-7.

Jiang, C. *et al.* (2012) 'Measuring financial protection for health in families with chronic conditions in Rural China', *BMC Public Health*, 12(1), p. 988. doi: 10.1186/1471-2458-12-988.

Jiang, C., Luo, C. and Zeng, G. (2018) 'Medical care cost changes under the background of medical insurance integration for urban and rural residents: trends and characteristics', *Journal of Jiangxi University of Science and Technology*, 39(04), pp. 115–120.

Jing, T. (2007) 'Discussion on the development of long-term nursing insurance for the elderly in China', *Chinese Journal of Gerontology*, (03), pp. 295–298. Available at: http://kns.cnki.net/kcms/detail/detail.aspx?filename=ZLXZ200703044&dbcode=CJFQ&dbname=cjfd2007&v=.

Johnston, M. P. (2014) 'Secondary Data Analysis: A Method of which the Time Has Come', *Qualitative and Quantitative Methods in Libraries*, 3(3), pp. 619–626.

Jones, A. M. and Rice, N. (2011) 'Econometric evaluation of health policies', in *The Oxford Handbook of Helath Economics*. Oxford: Oxford University Press.

Kellogg, R. and Wolff, H. (2008) 'Daylight time and energy: Evidence from an Australian experiment', *Journal of Environmental Economics and Management*, 56(3), pp. 207–220. doi: 10.1016/J.JEEM.2008.02.003.

Kim, S. (2017) *Solidarity, labour, and institution: The politics of health insurance reform in Japan and South Korea*. University of Sheffield.

Koch, S. and Alaba, O. (2010) 'On health insurance and household decisions: A treatment effect analysis', *Social Science & Medicine*, 70(2), pp. 175–182.

Kuha, J. (2004) 'AIC and BIC: Comparisons of Assumptions and Performance', *Sociological methods & research*, 33(2), pp. 188–229.

Kwon, S. (2009) 'Thirty years of national health insurance in South Korea: lessons for achieving universal health care coverage', *Health Policy and Planning*, 24(1), pp. 63–71.

Lagomarsino, G. *et al.* (2012) 'Moving towards universal health coverage: health insurance reforms in nine developing countries in Africa and Asia', *The Lancet*, 380(9845), pp. 933–943.

Law, M. R. *et al.* (2013) 'Growth in private payments for health care by Canadian households', *Health Policy*, 110(2–3), pp. 141–146. doi: 10.1016/j.healthpol.2013.01.014.

Lechner, M. (2010) 'The Estimation of Causal Effects by Difference-in-Difference MethodsEstimation of Spatial Panels', *Foundations and Trends® in Econometrics*, 4(3), pp. 165–224. doi: 10.1561/0800000014.

Lechner, M. (2011) *The estimation of causal effects by difference-in-difference methods*. Available at: https://faculty.smu.edu/millimet/classes/eco7377/papers/lechner 2010a.pdf (Accessed: 22 October 2021).

Lee, J. and McKibbin, W. (2004) *Estimating the global economic costs of SARS, Washington DC: National Academies Press, US*. Available at: https://books.google.co.uk/books?hl=en&lr=&id=GSY0Tjq_M74C&oi=fnd&pg=PA92&dq=Estimating+the+global+economic+costs+of+SARS.+&ots=Wc8KWPvjgL&sig=PfAqGFloGX2ClAa48NPFv1BqZ_I (Accessed: 26 March 2019).

Lei, X. and Lin, W. (2009) 'The New Cooperative Medical Scheme in rural China: does more coverage mean more service and better health?', *Health Economics*, 18(S2), pp. S25–S46. doi: 10.1002/hec.1501.

Levine, D., Polimeni, R. and Ramage, I. (2016) 'Insuring health or insuring wealth? An experimental evaluation of health insurance in rural Cambodia', *Journal of Development Economics*, 119, pp. 1–15.

Li, C. (2017) 'Study on the Effects of the Integration of Social Medical Insurance for Urban and Rural Residents on Residents' Medical Payment Burden', *China Health Standard Management*, 8(6), pp. 11–13.

Li, P., Lu, Y. and Wang, J. (2016) 'Does flattening government improve economic performance? Evidence from China', *Journal of Development Economics*, 123, pp. 18–37.

Li, X. and Tang, R. (2018) 'Research on the equity of health services utilization overall urban and rural residents ——Based on the survey of Shizuishan City, Ningxia', *Soft Science of Health*, 32(9), pp. 16–19.

Li, X. and Zhang, W. (2013) 'The impacts of health insurance on health care utilization among the older people in China', *Social Science & Medicine*, 85, pp. 59–65. doi: 10.1016/J.SOCSCIMED.2013.02.037.

Lin, Y., Shen, Y. and Sun, A. (2022) 'Evaluating the Stimulus Effect of Consumption Vouchers in China', *Economic Research Journal*, 55(7), pp. 4–20.

Liu, G. *et al.* (2004) 'Health Human Capital and Income Growth in China', *China Economic Quarterly*, 4(1), pp. 101–118.

Liu, H. and Zhao, Z. (2014) 'Does health insurance matter? Evidence from China's urban resident basic medical insurance', *Journal of Comparative Economics*, 42(4), pp. 1007–1020. doi: 10.1016/J.JCE.2014.02.003.

Liu, L. and Lin, H. (2018) 'Does Integration of Medical Insurance Reduce the Medical Burden of Poor Urban and Rural Residents?-Analysis Based on the Difference in Difference Method', *Collected Essays on Finance and Economics*, (08), pp. 22–31.

Liu, P., Wang, H., *et al.* (2013) 'Analyzing the Structures of Hospitalization Costs of Peasant Based on the Integration of Urban and Rural Areas', *Chinese Health Economics*, (10), pp. 41–43.

Liu, P., Xiong, L., *et al.* (2013) 'Empirical Analysis on Outpatient Costs of Township Hospitals in a Certain Country Based on Urban-Rural Integration', *Chinese Health Economics*, (7), pp. 50–52.

Liu, P. *et al.* (2018) 'The integration of urban and rural medical insurance to reduce the rural medical burden in China: a case study of a county in Baoji City.', *BMC health services research*, 18(1), p. 796. doi: 10.1186/s12913-018-3611-y.

Liu, X. (2017) 'The Integrated Urban and Rural Medical Insurance and the Utilization of Medical Care: Evidence from China', *The Journal of World Economy*, 40(03), pp. 169–192.

Liu, Y. (2002) 'Reforming China's urban health insurance system', *Health Policy*, 60(2), pp. 133–150. doi: 10.1016/S0168-8510(01)00207-X.

Liu, Y. and Zhang, X. (2017) 'Carbon emissions trading system and enterprise R & D innovation -- An Empirical Study Based on triple difference model', *Economic Science*, 03, pp. 102–114. Available at: doi:10.19523/j.jjkx.2017.03.008.

Liu, Y., Zhang, X. and Lan, X. (2016) 'The Influence of the Mixed Ownership Reform on the Total Factor Productivity of the State Owned Enterprises —— An Empirical Study Based on PSM-DID Method', *Public Finance Research*, 10(7), pp. 63–75.

Long, Q. *et al.* (2013) 'Changes in health expenditures in China in 2000s: Has the health system reform improved affordability', *International Journal for Equity in Health*, 12(1), pp. 40–47. doi: 10.1186/1475-9276-12-40.

Lu, S., Dong, R. and Ye, C. (2021) 'Does"the Belt and Road Initiative"Promote High-quality Exports——Evidence from Firms in China', *China Industrial Economics*, (3), pp. 80–98.

Luan, L., Hu, H. and Li, S. (2018) *Cost of Ischemic Heart Disease In China*.

Łuczak, J. and García-Gómez, P. (2012) 'Financial burden of drug expenditures in Poland', *Health Policy*, 105(2–3), pp. 256–264. doi: 10.1016/j.healthpol.2012.01.004.

Ma, C., Gu, H. and Song, Z. (2018) 'Inequality of Opportunity of Floating Population's Health Care under the Integrated Medical Insurance System-Paradox of the Ex-ante or Ex-post Compensation Priciple', *China Industrial Economics*, (02), pp. 100–117.

Ma, C., Gu, H. and Sun, X. (2015) 'Does Higher Grade of Medical Insurance Lead to Better Health: Evidence from Nature Experiment of Urban-Rural Integratied Medical Insurance', *Journal of Public Management*, 12(02), pp. 106–118.

Ma, C., Gu, H. and Sun, X. (2017) 'The Effects of Urban-Rural Integrated Medical Insurance System on Substantial Equity of Health Care and Health-Based on Equality of Opportunity', *Journal of Public Management*, 14(2), pp. 97–109.

Ma, C., Song, Z. and Gu, H. (2016) 'Research on Policy Effects of Integrated Medical Insurance System on Urban-rural Equality of Opportunity', *Chinese Journal of Population Science*, (1), pp. 108–117.

Ma, C., Zhao, G. and Gu, H. (2016) 'The Effect of Urban-Rural Integrated Medical Insurance System on Rural Residents' Health Care Behavior', *Statistical Research*, 33(4), pp. 78–85.

Ma, J., Lu, M. and Quan, H. (2008) 'From A National, Centrally Planned Health System To A System Based On The Market: Lessons From China', *Health Affairs*, 27(4), pp. 937–948. doi: 10.1377/hlthaff.27.4.937.

Ma, J. and Xu, A. (2018) 'Utilization of Health Services in Hypertensive Patients Enjoying 4 Different Types of Basic Medical Insurance:a Comparative Analysis', *Chinese General Practice*, 21(28), pp. 3518–3522.

Ma, Z. and Li, Y. (2018) 'A Study on the Impact of Urban and Rural Resident Medical Insurance on the Utilization of Medical Service for Residents', *China Journal of Pharmaceutical Economics*, 13(1), pp. 26-31,39.

Makhloufi, Kh., Ventelou, B. and Abu-Zaineh, M. (2015) 'Have health insurance reforms in Tunisia attained their intended objectives?', *International Journal of Health Economics and Management*, 15, pp. 29–51.

Mankiw, N. G. (2020) *Principles of Economics*.

Manning, W. G. and Marquis, M. S. (1996) 'Health insurance: The tradeoff between risk pooling and moral hazard', *Journal of Health Economics*, 15(5), pp. 609–639. doi: 10.1016/S0167-6296(96)00497-3.

Marshall, S., McGarry, K. and Skinner, J. (2010) *The Risk of Out-of-Pocket Health Care Expenditure at End of Life*. Cambridge, MA. doi: 10.3386/w16170.

Martineau, T., Gong, Y. and Tang, S. (2004) 'Changing medical doctor productivity and its affecting factors in rural China', *The International Journal of Health Planning and Management*, 19(2), pp. 101–111. doi: 10.1002/hpm.748.

Meng, Q. *et al.* (2012) 'Trends in access to health services and financial protection in China between 2003 and 2011: a cross-sectional study', *The Lancet*, 379(9818), pp. 805–814. doi: 10.1016/S0140-6736(12)60278-5.

Meng, Q. *et al.* (2015) 'Consolidating the social health insurance schemes in China: towards an equitable and efficient health system', *The Lancet*, 386(10002), pp. 1484–1492. doi: 10.1016/S0140-6736(15)00342-6.

Mettler, S. and Sorelle, M. (2018) 'Policy Feedback Theory', in *Theories of the Policy Process*. 4th edn. New York: Routledge, p. 32. doi: https://doi.org/10.4324/9780429494284.

Meyer-Sahling, J.-H. (2007) *Time and European Governance: An Inventory*. Nottingham. Available at: http://aei.pitt.edu/8017/1/sahling-j-04h.pdf.

Meyer, B. D. (1995) 'Natural and Quasi-Experiments in Economics', *Journal of Business & Economic Statistics*, 13(2), pp. 151–161.

Migiro, S. O. and Magangi, B. A. (2011) 'Mixed methods: a review of literature and the future of the ew research paradigm', *African Hournal of Business Management*, 5(11), pp. 3757–3764.

Miller, N. ., Eggleston, K. . and Zhang, Q. (2011) 'Understanding China's Mortality Decline under Mao: A Provincial Analysis, 1950–1980', *Unpublished manuscript.[PMC free article][PubMed]*.

Ministry of Finance of the People's Republic of China (2014) *Analysis of Repeated Participation of New Rural Cooperative Medical Scheme and Urban Resident Basic Medical Insurance*. Available at: http://www.mof.gov.cn/mofhome/shanxig/lanmudaohang/dcyj/201406/t20140612_1098399.html (Accessed: 2 May 2019).

Mu, R. (2014) 'Regional disparities in self-reported health: evidence from Chinese older adults', *Health Economics*, 23(5), pp. 529–549.

Mullahy, J. (1997) 'Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior', *Review of Economics and Statistics*, 79(4), pp. 586–593. doi: 10.1162/003465397557169.

Murray, C. J. L. and Frenk, J. (2000) 'A framework for assessing the performance of health systems', *Bulletin of the World Health Organization*, 78, pp. 717–731. doi: 10.1590/S0042-96862000000600004.

Mushkin, S. J. (1962) 'Health as an Investment', *Journal of Political Economy*, 70(5, Part 2), pp. 129–157. doi: 10.1086/258730.

Muurinen, J.-M. (1982) ‘Demand for health: A generalised Grossman model’, *Journal of Health Economics*, 1(1), pp. 5–28. doi: 10.1016/0167-6296(82)90019-4.

Mwabu, G., Ainsworth, M. and Nyamete, A. (1993) ‘Quality of Medical Care and Choice of Medical Treatment in Kenya: An Empirical Analysis’, *The Journal of Human Resources*, 28(4), p. 838. doi: 10.2307/146295.

National Health and Family Planning Commission (2015) *Report on Nutrition and Chronic Diseases of Chinese Residents*. Beijing, China: People’s Medical Publishing House.

National Healthcare Security Administration (2018) *Notice on implementation of integration of urban and rural resident basic medical insurance in 2018, National Healthcare Security Administration*. Available at: http://www.nhsa.gov.cn/art/2018/8/14/art_37_316.html (Accessed: 31 March 2019).

Newhouse, J. . and Phelps, C. . (1974) ‘Price and income elasticities for medical care services’, in *The economics of health and medical*. Available at: https://link.springer.com/chapter/10.1007/978-1-349-63660-0_9 (Accessed: 25 July 2022).

Nguyen, C. V. (2012) ‘The impact of voluntary health insurance on health care utilization and out-of-pocket payments: New evidence for Vietnam’, *Health Economics*, 21(8), pp. 946–966.

Nocera, S. and Zweifel, P. (1998) ‘The Demand for Health: An Empirical test of the Grossman Model Using Panel Data’, in. Springer, Boston, MA, pp. 35–49. doi: 10.1007/978-1-4615-5681-7_2.

Nyman, J. A. and Maude-Griffin, R. (2001) ‘The Welfare Economics of Moral Hazard’, *International Journal of Health Care Finance and Economics*, 1(1), pp. 23–42. doi: 10.1023/A:1011547904553.

Olabode, S. O., Olateju, O. I. and Bakare, A. A. (2019) ‘An Assessment of the Reliability of Secondary Data in Management Science Research’, *International Journal of Business and Management Review*, 7(3), pp. 27–43.

Olden, A. and Moen, J. (2022) ‘The triple difference estimator’, *The Econometrics Journal*, 25(3), pp. 531–553.

Or, Z. *et al.* (2010) ‘Are health problems systemic? Politics of access and choice under Beveridge and Bismarck systems’, *Health Economics, Policy and Law*, 5(3), pp. 269–293.

Pandey, A. *et al.* (2018) 'Inequity in out-of-pocket payments for hospitalisation in India: Evidence from the National Sample Surveys, 1995–2014', *Social Science and Medicine*, 201, pp. 136–147. doi: 10.1016/j.socscimed.2018.01.031.

Pendergrass, S. A. and Crawford, D. C. (2018) 'Using Electronic Health Records To Generate Phenotypes For Research', *Current Protocols in Human Genetics*, 100(1).

Peng, F. and Fan, Z. (2016) 'Tax Incentives, Donation Costs and Enterprises' Charitable Donations', *The Journal of World Economy*, 39(7), pp. 144–167.

Peng, X. (2019) *Research on the Integration of Basic Medical Insurance for Urban and Rural Residents in Xiangyin County*. Central South University of Forestry and Technology.

Pierson, P. (2000) 'Increasing Returns, Path Dependence, and the Study of Politics', *American Political Science Review*, 94(2), pp. 251–267. doi: https://doi.org/10.2307/2586011.

Qi, L. and Zhao, J. (2012) 'Nutritional Intervention and Human Capital Development of Boarding Students in Poor Areas——Based on the Study of Controlled Experimental Projects', *Management World*, (2), pp. 52–61.

Qin, J. (2014) 'Trends in the prevalence of chronic diseases and related risk factors in China, problems faced and countermeasures', *Chinese Journal of Public Health*, 30(1), pp. 1–4. doi: 10.11847/zgggws2014-30-01-01.

Qiu, Y., Zhai, S. and Hao, J. (2011) 'Theory, Empirical Study and Recommendations for the Integrated Development of Urban and Rural Health Care', *China Soft Science*, (04), pp. 75–87.

Rambachan, A. and Roth, J. (2023) 'A More Credible Approach to Parallel Trends', *The Review of Economic Studies*.

Regnault, A., Willgoss, T. and Barbic, S. (2017) 'Towards the use of mixed methods inquiry as best practice in health outcomes research', *Journal of patient-reorted outcome*, 2(1), p. 19.

Ren, X. *et al.* (2020) 'Burden of depression in China, 1990–2017: Findings from the global burden of disease study 2017', *Journal of Affective Disorders*, 268, pp. 95–101. doi: 10.1016/j.jad.2020.03.011.

Ried, W. (1998) 'Comparative dynamic analysis of the full Grossman model', *Journal of Health Economics*, 17(4), pp. 383–425. doi: 10.1016/S0167-6296(97)00038-6.

Roth, J. (2022) ‘Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends’, *American Economic Review: Insights*, 4, pp. 305–322.

Rubin, D. B. (1974) ‘Estimating causal effects of treatments in randomized and nonrandomized studies.’, *Journal of Educational Psychology*, 66(5), pp. 688–701. doi: 10.1037/h0037350.

Sari, N. and Osman, M. (2015) ‘The effects of patient education programs on medication use among asthma and COPD patients: a propensity score matching with a difference-in-difference regression approach’, *BMC Health Services Research*, 15(1), p. 332. doi: 10.1186/s12913-015-0998-6.

Schultz, T. W. (1961) ‘Investment in Human Capital’, *The American Economic Review*, 51(1), pp. 1–17.

Shan, L. *et al.* (2018) ‘Dissatisfaction with current integration reforms of health insurance schemes in China: are they a success and what matters?’, *Health Policy and Planning*, 33(3), pp. 345–354. doi: 10.1093/heapol/czx173.

Shang, X. *et al.* (2016) ‘Study on the Selection of Insurance Premium Level of Residents and the Influencing Factors after the Integration of Urban and Rural Medical Insurance’, *Chinese Health Economics*, 35(10), pp. 29–31.

Smith, A. K. *et al.* (2011) ‘Conducting High-Value Secondary Dataset Analysis: An Introductory Guide and Resources’, *Journal of General Internal Medicine*, (26), pp. 920–929.

Song, L. and Zuo, D. M. (2010) ‘Gender Differences in Medical Expenditure of Rural Elderly and Its Influencing Factors: A Case Study of Chaohu District’, *Chinese Rural Economy*, (5), pp. 74–83.

Stuart, E. A. (2010) ‘Matching methods for causal inference: A review and a look forward’, *Statistical science*, 25(1), pp. 1–25.

Su, D. *et al.* (2019) ‘Effect of integrated urban and rural residents medical insurance on the utilisation of medical services by residents in China: a propensity score matching with difference-in-differences regression approach.’, *BMJ open*, 9(2), p. e026408. doi: 10.1136/bmjopen-2018-026408.

Su, M. *et al.* (2018) 'Comparing the effects of China's three basic health insurance schemes on the equity of health-related quality of life: Using the method of coarsened exact matching', *Health and Quality of Life Outcomes*, 16(1), p. 41. doi: 10.1186/s12955-018-0868-0.

Sulieman, L. *et al.* (2022) 'Comparing medical history data derived from electronic health records and survey answers in the All of Us Research Program', *Journal of the American Medical Informatics*, 29(7), pp. 1131–1141.

Sun, Q. *et al.* (2009) 'Evaluating the financial protection of patients with chronic disease by health insurance in rural China', *International Journal for Equity in Health*, 8(1), p. 42. doi: 10.1186/1475-9276-8-42.

Sun, Q., Zhao, M. and Yin, A. (2017) 'Willingness-to-participate in medical insurance and its influencing factors among residents in Shandong province after integration of urban and rural medical insurance systems', *Chinese Journal of Public Health*, 33(11), pp. 1639–1642.

Sun, X. *et al.* (2009) 'Catastrophic medical payment and financial protection in rural China: evidence from the New Cooperative Medical Scheme in Shandong Province', *Health Economics*, 18(1), pp. 103–119. doi: 10.1002/hec.1346.

Sun, Y. *et al.* (2014) 'Studying on the residents' satisfaction on integrated medical insurance system in Ningxia', *Chinese Health Service Management*, 31(9), pp. 672–674.

Tao, D. *et al.* (2010) *Fixing the public hospital system in China*. Washington DC.

Tariq, S. and Woodman, J. (2013) 'Using mixed methods in health research', *The Journal of the Royal Society of Medicine*, pp. 1–8.

Tashakori, A. and Creswell, J. W. (2007) 'The New Era of Mixed Methods', *Journal of Mixed Methods Research*, 1(1), pp. 3–7.

The State Council of China (1998) *The Decision of the State Council on Setting up Basic Medical Insurance System for Staff Members and Workers in Cities and Towns, The central people's government of the people's republic of China.* Available at: http://www.gov.cn/banshi/2005-08/04/content_20256.htm (Accessed: 10 March 2019).

The State Council of China (2009) *The opinions of the Central Committee of the Communist Party of China and the State Council on Deepening the Reform of Healthcare System, The State Council of China*. Available at: http://www.gov.cn/gongbao/content/2009/content_1284372.htm (Accessed: 2 May 2019).

The State Council of China (2012) *White Paper: Medical and Health Services in China, The state Council Information Office of the People's Republic of China*. Available at: http://www.scio.gov.cn/tt/Document/1261745/1261745.htm (Accessed: 28 February 2019).

The State Council of China (2016) *The Opinions of the State Council on the Integration of the Basic Medical Insurance System for Urban and Rural Residents, Central People's Government of the People's Republic of China*. Available at: http://www.gov.cn/zhengce/content/2016-01/12/content_10582.htm (Accessed: 28 February 2019).

Trujillo, A. J., Portillo, J. E. and Vernon, J. A. (2005) 'The impact of subsidized health insurance for the poor: Evaluating the colombian experience using propensity score matching', *International Journal of Health Care Finance and Economics*, 5(3), pp. 211–239. doi: 10.1007/s10754-005-1792-5.

Vinha, K. (2006) *A primer on propensity score matching estimators*.

Viroj, T. *et al.* (2007) 'Achieving Universal Coverage in Thailand: What Lessons Do We Learn?', *SSRN Electronic Journal*.

Wagstaff, A. (1986) 'The demand for health: Some new empirical evidence', *Journal of Health Economics*, 5(3), pp. 195–233. doi: 10.1016/0167-6296(86)90015-9.

Wagstaff, A. (1993) 'The demand for health: An empirical reformulation of the Grossman model', *Health Economics*, 2(2), pp. 189–198. doi: 10.1002/hec.4730020211.

Wagstaff, A. *et al.* (2009) 'Extending health insurance to the rural population: An impact evaluation of China's new cooperative medical scheme', *Journal of Health Economics*, 28(1), pp. 1–19. doi: 10.1016/J.JHEALECO.2008.10.007.

Wagstaff, A. and Lindelow, M. (2008) 'Can insurance increase financial risk?: The curious case of health insurance in China', *Journal of Health Economics*, 27(4), pp. 990–1005. doi: 10.1016/J.JHEALECO.2008.02.002.

Wang, H. (2009) 'A dilemma of Chinese healthcare reform: How to re-define government roles?', *China Economic Review*, 20(4), pp. 598–604. doi: 10.1016/J.CHIECO.2009.04.001.

Wang, X. *et al.* (2014) 'Integration of rural and urban healthcare insurance schemes in China: an empirical research', *BMC Health Services Research*, 14(1), p. 142. doi: 10.1186/1472-6963-14-142.

Wang, X. and Liu, Y. (2015) 'An Empirical Analysis of the Influence of Urban and Rural Medical Insurance Coordination on Medical Service Differences', *Finance and Accounting Monthly*, (32), pp. 104–107.

Wang, Y. and Wang, S. (2010) 'Research progress on the demand for medical care and affecting factors', *Chinese Journal of Health Policy*, 3(8), pp. 55–62.

Wang, Y. and Wu, X. (2019) 'How Unconventional Monetary Policy Works in China:The Collateral Channel', *Economic Research Journal*, 54(12), pp. 86–101.

Wang, Z. *et al.* (2019) 'The comparison of healthcare utilization inequity between URRBMI and NCMS in rural China', *International Journal for Equity in Health*, 18(1), p. 90. doi: 10.1186/s12939-019-0987-1.

Watts, J. (2006) 'China's rural health reforms tackle entrenched inequalities', *The Lancet*, 367(9522), pp. 1564–1565. doi: 10.1016/S0140-6736(06)68675-3.

Weiyuan, C. (2008) 'China's village doctors take great strides.', *Bulletin of the World Health Organization*, 86(12), pp. 914–5. doi: 10.2471/BLT.08.021208.

WHO (2000) *The World Health Report 2000 Health Systems: Improving Performance*. World Health Organization. Available at: https://books.google.co.uk/books?hl=en&lr=&id=luqgKK2euxoC&oi=fnd&pg=PR7&dq=world+health+report+2000&ots=sMm42daIQa&sig=tRD__F_r6K2VwzHOR33pLgZ-45w#v=onepage&q=world health report 2000&f=false.

WHO (2003) *Social Health Insurance*. New Delhi. Available at: http://apps.searo.who.int/PDS_DOCS/B3457.pdf (Accessed: 2 May 2019).

WHO (2005) *Sustainable health financing, universal coverage and social health Insurance*. Geneva.

Wirtz, V. J. *et al.* (2012) 'Heterogeneous Effects of Health Insurance on Out-of-Pocket Expenditure on Medicines in Mexico', *Value in Health*, 15(5), pp. 593–603.

Wooldridge, J. M. (2012) *Introductory Econometrics: A Modern Approach*. Fifth. Available at: https://economics.ut.ac.ir/documents/3030266/14100645/Jeffrey_M._Wooldridge_Introductory_Econometrics_A_Modern_Approach__2012.pdf.

World Bank (1997) *Financing health care: Issues and options for China*. Washington DC.

Wu, Y. (2020) *The implementation effect of urban and rural basic medical insruance in Zhongmou county*. Henan University of Economics and Law.

Xi, B. *et al.* (2014) 'Maternal and child mortality in China', *The Lancet*, 383(9921), pp. 953–954. doi: 10.1016/S0140-6736(14)60481-5.

Xie, D., Wang, Q. and Zhu, C. (2017) 'Analysis of hospitalization of children and students before and after integration of medical insurance system', *China Social Security*, (11), pp. 86–87.

Xie, E. (2009) 'Income-related inequality of health and health care utilization', *Economic Research*, 2(92–105).

Xie, H., Mu, J. and LI, Y. (2017) 'The Medium That Added-value Tax Reform Influences Labor Share——An empirical study based on DDD model of the panel', *Science Technology and Industry*, 17(10), pp. 113–118.

Xie, Z., Hu, M. and Chen, W. (2017) 'The choices of medical institutions and the satisfaction of rural residents in Ningxia before and after the integration of urban and rural residents' basic health insurance schemes', *Chinese Health Resources*, 20(4), pp. 308-312,335.

Xingzhu, L. and Huaijie, C. (1992) 'China's Cooperative Medical System: Its Historical Transformations and the Trend of Development', *Journal of Public Health Policy*, 13(4), p. 501. doi: 10.2307/3342538.

Xu, A. and Zhang, Z. (2014) 'The real rate of reimbursement on urban-rural integration medical insurance:Empirical anal-ysis of Tianjin city', *Chinese Journal of Health Policy*, (9), pp. 4–11.

Xu, K. *et al.* (2019) *Global Spending on Health: A World in Transition*.

Xueshan, F. *et al.* (1995) 'Cooperative medical schemes in contemporary rural China', *Social Science & Medicine*, 41(8), pp. 1111–1118. doi: 10.1016/0277-9536(94)00417-R.

Yan, F. *et al.* (2018) 'Investigation on satisfaction of overall urban and rural residents after participating basic medical insurance', *Soft Science of Health*, 32(1), pp. 39–41.

Yang, L. *et al.* (2008) 'Economic burden of cardiovascular diseases in China', *Expert Review of Pharmacoeconomics & Outcomes Research*, 8(4), pp. 349–356.

Yang, X. *et al.* (2018) 'The inequality of inpatient care net benefit under integration of urban-rural medical insurance systems in China', *International Journal for Equity in Health*, 17(1), p. 173. doi: 10.1186/s12939-018-0891-0.

Yardley, L. and Bishop, F. L. (2015) 'Using mixed methods in health research: benefits and challenges', *British journal of health psychology*, 20(1), pp. 1–4.

Ye, H. and Liu, B. (2020) 'Study on Benefit Equity of Basic Medical Insurance for Rural Residents with Different Income Levels-Based on theSurvey of Poverty-striken Ethnic Minority Counties in Hubei Province', *Social Security Studies*, (01), pp. 23–32.

Yi, P. and Zhang, W. (2018) 'Integration of Social Health Insurance Program-"Single System, Multi-Standard" or "Single System, Single-Standard"', *Journal of Public Management*, 15(04), pp. 80–90.

Yip, W. C.-M. *et al.* (2012) 'Early appraisal of China's huge and complex health-care reforms', *The Lancet*, 379(9818), pp. 833–842. doi: 10.1016/S0140-6736(11)61880-1.

Yip, W. and Hsiao, W. (2009) 'China's health care reform: A tentative assessment', *China Economic Review*, 20(4), pp. 613–619. doi: 10.1016/j.chieco.2009.08.003.

Yip, W. and Hsiao, W. C. (2015) 'What Drove the Cycles of Chinese Health System Reforms?', *Health Systems & Reform*, 1(1), pp. 52–61. doi: 10.4161/23288604.2014.995005.

Yu, B. *et al.* (2010) 'How does the New Cooperative Medical Scheme influence health service utilization? A study in two provinces in rural China', *BMC Health Services Research*, 10(1), p. 116. doi: 10.1186/1472-6963-10-116.

Yu, H. (2015) 'Universal health insurance coverage for 1.3 billion people: What accounts for China's success?', *Health Policy*, 119(9), pp. 1145–1152. doi: 10.1016/J.HEALTHPOL.2015.07.008.

Yu, M. *et al.* (2021) 'The Impact of Catastrophic Medical Insurance in China: A five-year patient-level panel study', *The Lancet Regional Health - Western Pacific*, 13.

Zanutto, E. L. (2006) 'A comparison of propensity score and linear regression analysis of complex survey data', *Journal of data Science*, 4(3), pp. 67–91.

Zeckhauser, R. J. (1970) 'Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives', *Journal of Economic theory*, 2(1), pp. 10–26.

Available at: https://www.sciencedirect.com/science/article/pii/0022053170900104 (Accessed: 16 May 2019).

Zhai, H. and Liu, J. (2016) 'Rethinking of Ischemic Heart Disease', *Progress in cardiovascular disease*, 37(04), pp. 395–400.

Zhang, A., Nikoloski, Z. and Mossialos, E. (2017) 'Does health insurance reduce out-of-pocket expenditure? Heterogeneity among China's middle-aged and elderly', *Social Science & Medicine*, 190, pp. 11–19. doi: 10.1016/J.SOCSCIMED.2017.08.005.

Zhang, C. (2003) 'Nutrition, health and efficiency: Evidence from poor rural areas in China', *Economic Research*, 1, pp. 3–12.

Zhang, D. and Unschuld, P. U. (2008) 'China's barefoot doctor: past, present, and future.', *Lancet (London, England)*, 372(9653), pp. 1865–1867. doi: 10.1016/S0140-6736(08)61355-0.

Zhang, D. and Zheng, X. (2013) 'Is House Purchase Limit an Effective Policy to Control Housing Price?', *The Journal of Quantitative & Technical Economics*, 30(11), pp. 56–72.

Zhang, H., Huang, H. and Yan, Q. (2016) 'Mixed Ownership Reform, Policy Burden and State-owned Enterprise Performance——Based on Industrial Enterprises Database from 1999-2007', *Economist*, (9), pp. 32–41.

Zhang, Z. (1994) 'A retrospective study on Chinese rural cooperative medical system', *Chinese Rural Health Care Management*, 14(14), pp. 4–9.

Zhao, D. *et al.* (2019) 'Epidemiology of cardiovascular disease in China: current features and implications', *Nature Reviews Cardiology*, 16, pp. 203–212.

Zhao, Z. and Hou, Z. G. (2005) 'Health Demand of Urban Residents in China and Grossman Model: Evidence from Cross-sectional Data', *Economic Research Journal*, (10), pp. 79–90.

Zheng, G. (2014) 'Analysis and recommendations for integrating the urban and rural health insurance schemes', *China Health Insurance*, 2, pp. 8–11.

Zheng, X. *et al.* (2017) 'Elasticity of the demand for health care services in China: a systematic review', *Chinese Journal of Evidence-Based Medicine*, 17(5), pp. 564–572.

Zhou, L. and Chen, Y. (2005) 'Policy Effect of Rural Tax and Fee Reform in China: Estimation Based on Difference in Difference Model', *Economic Research Journal*, (8), pp. 44–53. Available at:

http://137.189.161.113/PaperCollection/webmanager/wkfiles/6686_1_paper.pdf (Accessed: 7 August 2019).

Zhou, X., Wang, D. and Li, J. (2011) 'The Effect of the New Cooperative Medical System on Household Savings in Rural China', *ECONOMIC SCIENCE*, (2), pp. 63–76.

Zhou, Zhongliang *et al.* (2011) 'Estimates of Elasticity of Rural Residents' Demand for Health Care', *Chinese Health Economics*, 30(12), pp. 14–16.

Zhou, Zhongliang, Zhou, Zhiying, *et al.* (2014) 'The Effect of Urban Basic Medical Insurance on Health Service Utilisation in Shaanxi Province, China: A Comparison of Two Schemes', *PLoS ONE*. Edited by A. Mahal, 9(4), p. e94909. doi: 10.1371/journal.pone.0094909.

Zhou, Zhongliang, Zhu, L., *et al.* (2014) 'The effects of China's urban basic medical insurance schemes on the equity of health service utilisation: evidence from Shaanxi Province', *International Journal for Equity in Health*, 13(1), p. 23. doi: 10.1186/1475-9276-13-23.

Zhu, H. (2019) *Research on medical reform from the perspective of Finance*.

Zhu, N. S. *et al.* (1989) 'Factors associated with the decline of the Cooperative Medical System and barefoot doctors in rural China.', *Bulletin of the World Health Organization*, 67(4), pp. 431–41. Available at: http://www.ncbi.nlm.nih.gov/pubmed/2805221 (Accessed: 11 March 2019).

Zweifel, P. (2012) 'The Grossman model after 40 years', *The European Journal of Health Economics*, 13(6), pp. 677–682. doi: 10.1007/s10198-012-0420-9.