INTRA-PROVINCIAL FISCAL DECENTRALISATION AND HEALTHCARE SERVICE PERFORMANCE: EVIDENCE FROM CHINA

BY

SHUO SUN

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy of Cardiff University

Management, Employment, and Organisation Section of Cardiff Business School, Cardiff University

January 2023

ABSTRACT

Over the past decades, fiscal decentralisation (FD), which refers to the downward transfer of spending responsibilities and revenue sources within the multi-level public sector, has become a global trend. Theories of FD (i.e., fiscal federalism, FF) argue that local governments are more familiar with local demands and have greater advantages in mobilising local resources. Therefore, FD improves public service performance (PSP). However, there is still a lack of empirical evidence for this argument, particularly regarding FD's impact on healthcare services within Chinese provinces.

Using secondary data from China's 26 provinces from 2006 to 2017, this research addresses the above gap. Moreover, the role of local government capacity (measured by relative wealth) in the relationships between FD and healthcare performance, which is also a key issue highlighted in FD theories, is investigated. This study focuses on two key concepts of PSP: efficiency and effectiveness, and three dimensions of FD: healthcare expenditure decentralisation (HED), total expenditure decentralisation (TED), and revenue decentralisation (RD). It is found that FD from the provincial to the sub-provincial level of government improves healthcare efficiency. Additionally, the benefit of HED for healthcare efficiency may be greater with the increase in government capacity. However, apart from HED, TED and RD are negatively associated with healthcare effectiveness, and such negative relationships fail to be addressed by increasing government capacity.

Literature-based discussions suggest that the divergent effects of FD can be attributed to key features of the Chinese-style FD. That is, despite a high level of decentralisation, the performance management system (PMS) in China still follows a top-down pattern. Therefore, to gain tangible benefits such as promotion opportunities, decentralised sub-provincial cadres (e.g. bureaucrats of the local government and the healthcare department) tend to be more accountable for better healthcare efficiency – a key target highlighted by upper-level government leaders. Accordingly, cadres may pay less attention to healthcare effectiveness which is relatively ignored by their superiors.

Key words: fiscal decentralisation, healthcare service performance, efficiency, effectiveness, Chinese provinces

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3Es Economy, Efficiency, and Effectiveness
BCC Banker, Chames and Cooper DEA Model
CCR Chames, Cooper and Rhodes DEA Model
CPC Communist Party of China
CRS Constant Returns to Scale (adopted in the CCR DEA Model)
DMU Decision-Making Unit
ED Expenditure Decentralisation
FD Fiscal Decentralisation
FF Fiscal Federalism
HC Health Commission
HED Healthcare Expenditure Decentralisation
IMR Infant Mortality Rate
IOO Input, Output, and Outcome
NPM New Public Management
OECD Organisation for Economic Co-operation and Development
PMS Performance Management System
PMR Perinatal Mortality Rate
PSR Perinatal Survival Rate
PSP Public Service Performance
PRC People's Republic of China
RD Revenue Decentralisation
TED Total Expenditure Decentralisation
TRS Target Responsibility System
VRS Variable Return to Scale (adopted in the BCC DEA Model)

ACKNOWLEDGEMENTS

I would like to express my gratitude to many people who give me great and endless support in the past years of pursuing my doctoral study. First of all, my appreciation goes to my first supervisor Professor Rhys Andrews, for his patient guidance, academic advice, and continuous encouragement throughout the past years of my PhD study. Moreover, it is also a great honour for me to benefit from his personality and diligence as well as his ethical attitude, which is of great value for my future career as a researcher and a human being.

Also, I would like to thank my second and third supervisors Professor James Downe and Dr Heike Doering. Without their valuable suggestions and generous assistance, my thesis could not have reached to the destination. Particularly, I would like to express my appreciation to Dr Dennis De Widt, not only for his insightful advice about my PhD study but also for his selfless support in terms of research in public finance/accounting, teaching, and career development within and beyond CARBS.

I would also like to thank my friends and PhD colleagues at Cardiff Business School. I am grateful for all your valuable comments about my research and help regarding my personal life. I wish you all success in your life and academic career.

Finally, my heartfelt gratitude goes to my family members especially my beloved parents for their endless love and support over the past years.

CHAPTER 1 INTRODUCTION

1.1 Research Background

Over the past fifty years, FD, as a policy initiative to promote government performance, has spread widely around the world. FD usually refers to the downward transfer of spending responsibilities and revenue sources within the multi-level public sector (Oates 1972; Wang and Ma 2014; Reingewertz 2014). Starting from the 1970s, FD was introduced in industrialised countries such as Italy, Spain, and Australia to promote localised decision-making (Stegarescu 2005; Blöchliger 2006; Bodman and Hodge 2010). Then, supported by international institutions such as the World Bank and the Organisation for Economic Co-operation and Development (OECD) (World Bank 1999; OECD 2016), FD spread to Latin America and East European countries with the aim of building modern intergovernmental relations and addressing the failures of centralised planning, so as to achieve sustainable economic growth (Rezk 1999; Wiesner 2003; Oates 2005). Moreover, with sub-national governments' advantages in understanding local demands and utilising local resources, FD has been increasingly considered as a way to improve public service performance (PSP) (Oates 1999; Dwicaksono and Fox 2018). Nowadays, FD-related practices are being implemented globally, including in emerging countries such as China (Wang and Ma 2014).

Starting from the late 1970s, China's economic reform, known as "Reform and Opening-up" (*Gaige Kaifang* in Chinese), has substantially changed many socioeconomic aspects of the country, including its inter-governmental relations. As highlighted by Shen et al. (2012), decentralisation from the top leadership of the central government and the Communist Party of China (CPC) to local governments and the private sector is a key feature of China's reform. Although top-down political control still exists, sub-national bureaucracies have been authorised to play a key role in formulating and implementing local policies, leading to a high level of FD – particularly on the expenditure side (Dollar and Hofman 2006; Shen et al. 2012). With the current arrangements of inter-governmental relations, China has become a de facto federal state (Zheng 2007; Boadway and Shah 2009), providing a perfect research field to look into the impact of FD on various socioeconomic outcomes, including PSP. Promoting PSP is an important concern for citizens, as it directly addresses their demands for better public services and concerns about the usage of their tax payments (Brignall and Modell 2000). It also benefits the government by reducing political risks and building trust between the public sector and the general public (de Walle and Bouckaert 2003).

The remarkable spread of FD around the world has brought multi-disciplinary attention (Wang and Ma 2014) and has been conceptualised by researchers as fiscal federalism (FF) (Ter-Minassian 1997). As explained by the seminal studies of Musgrave (1959) and Oates (1972), FF revolves around the assignment of responsibilities for providing public services and the fiscal instruments for implementing these responsibilities between different levels of governments. One of the core responsibilities of the government is to provide public services. For "national" services, such as maintaining macroeconomic stabilisation, fiscal federalists recognise the dominant role of the central or federal government. However, most public services are consumed within a relatively small jurisdiction, having strong local features, and the role of service provision should therefore be undertaken by local governments (Oates 2005; Boadway and Shah 2012). This is because local governments, compared with the central or federal government, have a greater capacity to collect information about local demands, resources, and circumstances (Oates 1999). Thus, decentralised local governments can perform better in providing local services, leading to greater social welfare. To measure the performance of service provision, various indicators have been proposed, and many of them are concerned with three key dimensions: economy, efficiency, and effectiveness. Economy refers to the costs or inputs needed to provide a service. However, this indicator is criticised because it provides limited reflections on service standards or the success/failure of the government (Bouckaert 1993). Efficiency, in most empirical studies, refers to the costs or inputs for a given level of output (Boyne 2002). Finally, effectiveness measures the achievement of a service's formal objectives (Andrews and Entwistle 2010). The latter two performance dimensions, i.e., efficiency and effectiveness, are usually considered as essential elements of a well-structured performance framework for public organisations (Boyne 2002), and thus, they are the key concerns of this doctoral research.

Among all public services, healthcare is recognised as one with especially strong local preferences for which local governments have great advantages in mobilising local information and resources (Verbeeten and Spekle 2015). Healthcare services have some externalities, i.e., some policies and outputs can be accessed by people from other jurisdictions, which may make local bureaucrats reluctant to invest in service performance (Ahmad and Craig 1997). However, the externalities of healthcare are considered to be limited and can be effectively controlled with top-down subsidies (Jiménez-Rubio 2011; Abimbola et al. 2019). Even with the losses from externalities being counted, the welfare gain of decentralising healthcare responsibilities might still be greater than that from centralised provision (Oates 2005). Given the above reasons, there has been increasing support for decentralising healthcare responsibilities and finance to subnational governments in many countries (Vrangbaek 2007), such as EU nations (Prieto and Saez 2006); Canada (Jiménez-Rubio 2011b), Switzerland (Arends 2017), India (Asfaw et al. 2007), and China (Uchimura and Jutting 2009) – the research background of this study.

Despite the growing practice of FD around the world, there are surprisingly few academic works looking into the impact of FD on healthcare performance. Existing FD studies in China and elsewhere – mostly in the field of public economics – have intensively discussed FD's economic effects but pay relatively little attention to the relationship between FD and PSP (Wang and Ma 2014), not to mention that only a few of them focus specifically on the healthcare sector (e.g. Asfaw et al. 2007; Jiménez-Rubio 2011). Moreover, hardly any previous studies simultaneously investigate the impact of FD on the efficiency and effectiveness of healthcare services and probe into the similarities and differences in the FD-efficiency and FD-effectiveness relationships (with the rare exception of Arends 2017). Also, the impact of FD on healthcare performance in China, an emerging country with a unique socioeconomic environment different from the western setting, is largely ignored. Within Chinese provinces, since the 1980s, the power of spending and formulating policies for local healthcare services have been largely decentralised to sub-provincial governments[1]. However, academic research on the relationship between FD and

[1] Sub-provincial governments refer to the prefecture-level and county-level governments which are located at the third and fourth tiers of China's administrative hierarchy. China has a

healthcare service performance is still in its infancy. Even though there are limited studies, most of them focus on decentralisation from the central government to provinces, rather than looking within the provinces where a high level of FD[2] for most public services (e.g. education and healthcare) also exists (e.g. Jin and Sun 2011; Hao et al. 2021).

To fill the above research gaps, this doctoral study aims to explore the impact of intra-provincial FD on healthcare service performance in China. As mentioned above, two key performance elements namely efficiency and effectiveness are considered, so as to identify and discuss the potentially diversified healthcare performance effects of FD. Moreover, one of the key pre-assumptions in theoretical and empirical FD studies, i.e., local governments' greater capacity in understanding local information and managing local services, is tested as well. Such "local advantages" should not be taken for granted (Azfar et al. 2001; Brinkerhoff and Wetterberg 2013). As suggested by Prud'homme (1995), if these advantages diminish, FD might contribute less to and even jeopardise PSP. With the research gaps being addressed, this research will provide theoretical and empirical contributions to the multi-disciplinary discussion of FD and PSP, and also, practitioners in Chinese provinces will be informed.

5-tier administrative hierarchy, in which the central government is located at the first tier, followed by provincial governments, prefecture-level governments, county-level governments, and township governments. Each level of the government has a series of functional agencies such as the Educational Bureau. Township governments are directly controlled by the counties above them and thus could be considered as 'branches' of the county government (Donaldson 2017b). Detailed introductions are given in the Context chapter.

[2] FD within provinces is hereafter referred to as 'intra-provincial FD'.

1.2 Research Context: China

China has the largest population and the third largest territory of any country around the world. Throughout its history, rulers of China have always aimed to build a unified country and to maintain its stability – the Qin dynasty made the first successful attempt in 220 BC. However, with the territory and population being continuously expanded, the top rulers of ancient China were increasingly faced with more complicated local situations, and they recognised that their regime could not be sustained without the active participation of local governments. Thus, autonomy in managing local affairs was authorised to local governments which were controlled by local clans (Wang 2006), as summarised by a Chinese proverb, "the imperial power does not go down to the counties ('*Huangquan buxia xian'*)".

Due to its unique geographical context, over the past 2000 years, Chinese history has largely revolved around the interactions and conflicts between the central and local rulers (Donaldson 2017a). To ensure the legitimacy of the governing regimes in China, Confucianism was formally recognised as the official philosophy of Chinese dynasties (Weatherley and Magee 2018). Confucian ideas indicate that the top rulers, i.e., emperors, are selected by heaven and govern the people on behalf of heaven. Such a description confers a sacred status on rulers and dynasties. However, it also requires the rulers to satisfy people's varying needs, and if not, the rulers will lose the 'authorisation' of the heaven and people will overturn their regime. The above key themes of Confucianism have influenced almost all dynasties of China, including the contemporary communist regime. In 1949, the Communist Party led by Mao Zedong established the People's Republic of China (PRC). However, the attempt to rule the country with Communist ideologies failed to satisfy people's demand for a better life.

After 1976, the second generation of CPC leaders (re)acknowledged the importance of satisfying people's enduring needs and started China's economic reform (i.e., Reform and Opening-up, *Gaige Kaifang*) in 1978 (Central Committee of the CPC 1978).

Since then, the state has gradually released its control over the economy. The private sector was rebuilt, and foreign capital was allowed to be invested in various industries. Also, local governments at all levels were authorised with discretionary power in local socioeconomic policies. The economic achievements of the reform are remarkable. Since 2011, China has overtaken Japan as the second-largest economy in the world (IMF 2022). On the other hand, the political system is largely unreformed. Different from a Weberian administrative-political pattern, the governments are by no means politically impartial. The CPC is always the single ruling party with the power of controlling the government and key policies (Chu 2013; Zhou et al. 2021), and there is no clear division between party and state (government) bureaucracies at each administrative level (Snape and Wang 2020). In other words, all government and the party agencies belong to an integrated organisation, that is, the "party-state[3]" or the "party-government" (Edin 2003; Jing et al. 2015). Particularly, when looking into the vertical relationship between the authorities at different administrative levels (e.g. the intra-provincial FD – the concern of this research), both the government and the party should be taken into consideration from an integrated perspective (Zhou et al. 2021).

[3] In a broad sense, 'state' refers to the state organs at all administrative levels, including governmental, legislative, and judicial organs. All of the state organs in China are under the leadership of the CPC. However, to narrow the research focus, the term 'party-state', in most social science studies – particularly those focusing on the economy and public management, usually refers to the integrated entity which includes the party and governmental agencies at a certain administrative level (Snape and Wang 2020). For the sake of clarity, unless otherwise specified, this thesis uses "government" to refer to the integrated party-state bureaucracy in China.

Although top-down political control has been maintained and even consolidated, as a key part of the reform, local governments at all administrative levels (e.g. the provincial, prefectural, and county levels) have gained power and responsibilities in local socioeconomic affairs – as long as they follow the political leadership of the higher-level government (Zheng 2007). This is because the enormous land and population size of China makes it impossible for the central government to extend its power down to lower levels of jurisdictions (Donaldson 2017a). Also, in this way, subnational governments are encouraged to achieve better performance with local initiatives (Chien 2010; Bulman and Jaros 2020). Such arrangements between the higher and lower level of governments make China a perfect context to investigate the impact of decentralisation on various socioeconomic dimensions such as healthcare performance.

1.3 Contributions of this research

This research is expected to bring academic and social contributions by addressing significant gaps, advancing theoretical debates, and informing non-academic readers. Firstly, this study benefits empirical discussions by filling various research gaps. Most existing empirical studies focus on FD's economic impact (Wang and Ma 2014), while the impact of FD on PSP is relatively ignored. This research directly addresses this gap by focusing specifically on the FD-PSP relationships. Also, by examining healthcare efficiency and effectiveness, this doctoral study overcomes the following two limitations: firstly, most studies on the FD-PSP relationship only focus on one performance aspect (e.g. Barankay and Lockwood 2007; Balaguer-Coll et al. 2010; Jiménez-Rubio 2011), thereby missing the opportunity to identify potentially divergent effects of FD on different PSP aspects. By discussing both efficiency and effectiveness, this doctoral research offers a deeper understanding of the impact of FD. Secondly, previous studies, particularly those focusing on the FD-efficiency relationship, are usually concerned with the level of efficiency for all major public services rather than one specific service (e.g. Alonso and Andrews 2019). However, local citizens and bureaucrats, as public service users and providers, have varying preferences and requirements for different public services (Oates 2005; Gao et al. 2014). Therefore, the impact of FD on the performance of different services may vary. By specially exploring the impact of FD on healthcare service performance, this gap can be addressed as well.

Apart from addressing above gaps in general settings, this research also responds to the lack of academic discussions about FD's impact on healthcare performance in the Chinese context, particularly in Chinese provinces. Traditional FD research in China, as in elsewhere, has predominantly focused on FD's economic benefits, arguing that FD initiates local innovations to promote economic development (Qian and Weingast 1997; Smoke 2001). However, the social impact of FD has not yet been fully investigated (Wang and Ma 2014), with only rare exceptions such as Zhang et al. (2019), who explored the impact of expenditure-side decentralisation on the efficiency of environmental protection with problematic decentralisation indicators (see Chapter 3 Literature Review). Among the limited studies about FD's impact on public services, many of them are concerned about the quantity or cost of input which is of limited value for gauging performance and can be politically manipulated (Bouckaert 1993; Walker et al. 2010). This doctoral research, recognising the above research gaps, looks into FD's impact on healthcare efficiency and effectiveness using reliable measures drawn from trustworthy sources. Moreover, previous Chinese and global-based studies usually focus on FD between the central government and the provinces, rather than looking into FD within provinces (i.e., FD to the “third-tier” of the government) where most responsibilities for public service provision are allocated and performed (Niu 2013; Reingewertz 2014). Furthermore, many of these studies (e.g. Brock et al. 2015; Zhang et al. 2019; Hao et al. 2021) measure a province’s level of FD by the ratio of central government fiscal expenditure (per capita) to the province’s fiscal expenditure (per capita). However, the central government per capita expenditure on all provinces is by no means equal. This PhD research, by collecting provincial and sub-provincial fiscal data from 26 of the 27 Chinese provinces, establishes three indicators that directly measure the extent to which fiscal responsibilities and revenues are decentralised from the provincial to sub-provincial levels. Thus, the above two gaps are also addressed.

Secondly, the analysis and empirical findings of this research are expected to contribute to the ongoing debates in relation to the New Public Management (NPM), FF, and PSP theories. Regarding NPM, the results of this research provide real-world reflection that FD, as an NPM-inspired reform, could indeed address the efficiency concerns emphasised by NPM theories and practices. Furthermore, the negative FD-effectiveness relationship identified in this study aligns with the growing criticism that the traditional NPM theories focused too much on input and direct output indicators (e.g. quantity and productive efficiency) but ignored formal and long-term objectives like public service effectiveness (O’Flynn 2007; Çolak 2019).

Regarding contributions to FF and PSP theories, this research offers a comprehensive description of the key arrangements of FD in China, which leads to an insightful analysis of the behaviours of local cadres in a decentralised setting and the FD-PSP relationships. This study identifies three key arrangements of FD in the unique socio-political context of Chinese provinces: upward accountability, the target-setting system (TRS) and the top-down performance management system (PMS). With these institutions, cadres at each administrative level are held accountable to the government located one administrative level above. The higher-level government can deliver their key agendas and requirements to the local governments (via TRS) and regulates them with top-down performance evaluation (via PMS). Consequently, with FD, sub-provincial cadres have more autonomy to work on the key concern of the higher-level government, which, for the case of healthcare services, pertain to efficiency. However, for issues ignored by the higher-level government (for healthcare is effectiveness), local cadres may be less motivated to dedicate greater efforts. The diverse behaviours of local cadres also help explain the influence of "local advantages" on the FD-performance relationships. FF theories assume that compared with the higher-level government, local governments have informational, resource, and managerial advantages in carrying out local responsibilities (Oates 1972, 2005). While most empirical studies take this assumption for granted, this doctoral research empirically confirms that in Chinese provinces, as local advantages increase, FD contributes more to healthcare efficiency. However, such advantages fail to incentivise local cadres to work better on healthcare effectiveness at a certain degree of FD, as effectiveness is ignored in the top-down performance evaluation.

Finally, non-academic readers, particularly policymakers in Chinese provinces who are in charge of formulating FD policies, will also be informed by this research. Over the past decades, the CPC has increasingly claimed its legitimacy from better governance (Holbig and Gilley 2010; Wang 2013). As argued by Zeng (2014), optimising the arrangement of inter-governmental relations is an effective way to promote governance capacity and thus consolidates the legitimacy base for the CPC's ruling position. The performance in providing healthcare services as well as other social services, undoubtedly, is a key reflection of the governance capacity (Jing et al. 2015). In recent years, with the stagnation of economic development, fiscal tightening, and people's growing demand for better healthcare services, promoting healthcare services performance as a way to strengthen legitimacy has been receiving greater attention from the central and local governments and CPC organisations (Duckett and Munro 2022; Ratigan 2022). In this case, it is of timely necessity to perform this doctoral study and clarify whether or not a greater level of intra-provincial FD could benefit the efficiency and effectiveness of healthcare services. Thus, policy-makers in Chinese provinces can be informed to design better inter-governmental fiscal relations, and the general public, as users of healthcare services, can potentially be benefited as well.

1.4 Research objectives and aims

This doctoral research, focusing on the impact of intra-provincial FD on healthcare service efficiency and effectiveness, has the following specific research objectives:

1) To review the major theories on FD (essentially FF) and PSP which clarify the rationale for pursuing better performance in healthcare and other services.

2) To explain theoretical arguments regarding the mechanisms that drive FD to impact on the performance of healthcare and other major public services.

3) To critically review previous empirical findings about FD's impact on the efficiency and effectiveness of public services (i.e., healthcare and others).

4) To draw upon appropriate indicators to measure FD and the efficiency and effectiveness of healthcare services.

5) To construct appropriate statistical models for analysing the impact of FD on healthcare efficiency and effectiveness within Chinese provinces, along with the effects of sub-provincial government capacity on the relationships between FD and efficiency/effectiveness.

6) To explore the findings by re-examining the institutional contexts and further reviewing the literature.

In line with the research objectives, these are the research questions:

1) What is the relationship between FD and healthcare service efficiency in Chinese provinces?

2) What is the relationship between FD and healthcare service effectiveness in Chinese provinces?

3) Does government capacity (measured by a province's relative wealth) have a moderating effect on the relationship between FD and healthcare service efficiency/effectiveness?

1.5 Methodology of this research

First, an objectivist and positivist philosophical stance is chosen to guide the research design. This is in line with the aforementioned research aims and questions as well as the intention of this study to identify generalisable knowledge under a large research context, i.e., Chinese provinces (Saunders et al. 2007). Second, this doctoral study follows a quantitative research strategy and a correlational research design (Creswell 2009; Stangor 2011). This research strategy ensures that appropriate methods for data collection and estimation can be adopted to develop robust findings Specifically, panel data from 2006 to 2017 were collected from statistical yearbooks and reports that are openly published by the national and provincial Bureaus of Statistics as well as the central and provincial governments (e.g. the Finance Yearbooks of China). The data will cover all provinces of Mainland China apart from Tibet – due to data accessibility[4]. Then, the quantitative data were analysed using regression models. Detailed explanations of the sample, data, and methods are given in the Methodology chapter.

1.6 Structure of the thesis

This thesis includes eight chapters. Chapter 2 reviews the concepts, dimensions, and measurements of PSP at the beginning. Then, global practices to improve PSP are briefly introduced. Finally, this chapter elaborates on the theories of PSP, including public choice theory, goal-setting theory, stakeholder theory, and accountability theory. Chapter 3 provides a detailed introduction to the first and second-generation theories of FF – the theoretical underpinning of FD. Specifically, the mechanisms that

---

[4] Mainland China does not include Hong Kong and Macau - two quasi-independent territories of the PRC, and Taiwan. Four provincial metropolitan cities, including Beijing, Shanghai, Tianjin, and Chongqing, are politically located at the same tier as other ordinary provinces and provincial ethnical autonomous regions (e.g. Inner Mongolia. For the sake of clarity, they are also referred to as provinces). However, they are usually considered as cities rather than provinces. Also, their key socio-economic features such as the percentage of urban residents, level of economic development, and inter-governmental relations are significantly different from other provinces. Thus, they are not considered in this research. For sample selection and the units of analysis, see Section 5.2 of Chapter 5.

drive FD to promote PSP are demonstrated in detail. After that, empirical studies in regard to the impact of FD on the efficiency and effectiveness of healthcare services and other public services are reviewed.

Chapter 4, is a context chapter, incorporating two major sections. The first section introduces China's administrative hierarchy where FD is arranged in a top-down manner. Then, this section explains the evolution and arrangements of FD in China, which is followed by the conceptualisation of the Chinese-style FD. The second section of Chapter 4 details the evolution and current arrangements of the healthcare system in China.

Chapter 5 reports the methodology. It introduces the research paradigm and strategy at first. Then, Section 5.2 presents the units of analysis, the period of research, and data sources, which leads to the demonstrations of measurements for FD, healthcare efficiency, and healthcare effectiveness in Section 5.3 and 5.4. After that, quantitative models, variables, and main hypotheses are explained in Section 5.5, followed by the estimation strategies in Section 5.6.

Chapter 6 reports research findings, in which the impact of FD on healthcare efficiency and effectiveness are separately analysed in two sections which begin with summary statistics. Then, regression results depicting FD's effects on healthcare efficiency/effectiveness are reported, followed by the findings of relative wealth's (the indicator for government capacity) moderating effect on the FD-efficiency/effectiveness relationships.

Chapter 7 discusses the quantitative findings. This chapter starts by re-examining the arrangements of the institutional background, before considering the impact of FD on healthcare efficiency and effectiveness, respectively. Finally, the effect of relative wealth on the relationship between FD and healthcare service efficiency/effectiveness is discussed. Finally, the conclusion chapter (chapter 8) describes the academic contributions, practical contributions, limitations, and directions for further research.

CHAPTER 2 UNDERSTANDING PUBLIC SERVICE PERFORMANCE: AN INTRODUCTION TO ACADEMIC AND POLICY DEBATES

This chapter introduces the academic and policy debates regarding the concept of PSP. In Section 2.1 and Section 2.2, the concept, major dimensions of PSP, and the selection of performance indicators for this study are introduced. Then, Section 2.3 briefly reviews the practices of PSP improvement in a global context. Finally, drawing on major social science theories, Section 2.4 explains the motivations behind and rationales for the efforts of the government to improve service performance.

2.1 Pinning down public service performance: concept and dimensions

Given the importance of PSP for this thesis, there is a need to clarify what PSP refers to and to consider its key dimensions. Groundbreaking studies in PSP (e.g. Meier and O'Toole 2002; Meier *et al*. 2004) conceptualise the performance of public organisations as the result of service production. In other words, PSP is about "various inputs, organisational processes or management practices, outputs and longer-term impacts or outcomes, and the organisational environment (Walker et al. 2010, p.8)". Moreover, some researchers, using the 'logic of governance' framework, add more elements about political achievements such as participation, democratic processes, and accountability into the concept of PSP (Heinrich and Lynn 2000).

The aforementioned definitions and elements of PSP reveal substantial differences between the performance of private and public organisations. That is, private sector performance is mostly about creating financial benefits for stakeholders – a greater level of performance can therefore be directly measured by financial indicators, such as costs and profits (Andrews et al. 2006). Although the importance of external stakeholders and corporate social responsibility has been increasingly emphasised in private sector management, in essence, practices revolving around external stakeholders and corporate social responsibility are still about maximising shareholders' benefits in the long run (Nguyen et al. 2020). By contrast, public sector performance is more complicated, as it concerns the long-term impacts and the final outcomes of public services, and it cannot be measured solely by financial indicators such as outputs and profits (Boyne 2002; Andrews et al. 2006). Thus, organisational performance in the public sector cannot be viewed and addressed in the same way as performance in the private sector, although fruitful lessons and experiences from private entities could be (and have been) adopted to enhance the management of public organisations (O'Mahony and Stevens 2006; Ghobadian et al. 2007).

Within the public sector performance literature, various dimensions of performance have been proposed, and most of them could be summarised into the Economy-Efficiency-Effectiveness (3Es) framework and the Input-Output-Outcome (IOO) model (Palmer 1993; Boyne 2002; Andrews et al. 2006). The 3Es framework includes three major aspects of performance: economy, efficiency, and effectiveness. First, economy refers to the costs of procuring a certain level of inputs of a given quantity (e.g. equipment and staff costs). This dimension can be easily measured in financial terms and is also a key concern in private sector performance measurement (Palmer

1993). However, for measuring PSP, this indicator is criticised because it hardly reflects service quality, standards, or the success/failure of the government (Bouckaert 1993). For example, in most countries, it is illegal to unlimitedly squeeze the wages of the public sector workforce simply to control costs, and by no means would citizens benefit from low-cost services provided by underpaid staff (Boyne 2002).

The second aspect of the 3Es framework, efficiency, has four sub-dimensions: productive (or technical) efficiency, allocative efficiency, distributive efficiency, and dynamic efficiency (Andrews and Entwistle 2013). The first and dominant sub-dimension (i.e., productive efficiency) measures the direct outputs vis-à-vis the input of resources. Allocative efficiency is about the level of responsiveness to people's specific service demands. Distributive efficiency focuses on the distribution and redistribution of resources between social groups (Mercier Ythier 2010). Finally, dynamic efficiency is concerned with the allocation of resources over a long period of time, that is, the present consumption of resources should not sacrifice the potential for future consumption, and vice versa. Researchers indicate that apart from productive efficiency, the latter three types of efficiency may be difficult to measure due to their more subjective nature and a lack of objective, numerical indicators that can fully capture responsiveness, redistribution, and future consumption needs (Andrews et al. 2006; Andrews and Entwistle 2013).

Finally, effectiveness refers to the achievement of a service's formal objectives (Boyne 2002). In other words, effectiveness represents to what extent a defined task has been accomplished (Palmer 1993; Jackson and Palmer 1988). For example, the effectiveness of education services could be measured as the percentage of students passing exams, and the effectiveness of healthcare can be represented by mortality rates.

Apart from the 3Es framework, researchers propose the IOO model which covers the three aforementioned dimensions of 3Es but includes more aspects of PSP (Boyne 2002). “IOO” refers to three basic elements of any service production function: inputs, outputs, and outcomes. In this model, outputs not only measure the quantity of a service’s direct product but also include the quality of that output. For outcomes, effectiveness – the achievement of a service’s formal objective – is a key component, along with the impact of the service and the equity of service distribution among different groups (e.g. groups by gender, income, race, and regions). As explained by Le Grand (1982), public organisations should take equity into account, because the allocation of public services should be determined by the criteria of need, rather than the ability to pay. Also, within the IOO model, input-output efficiency as well as the level of outcomes relative to inputs (or cost-effectiveness i.e., cost per unit of outcome) are considered.

In addition to the above indicators, it is argued that the IOO model should include more indicators of PSP, especially responsiveness (Pollitt 1988), which is similar to the concept of allocative efficiency. Boyne (2002) suggests that public service providers should be responsive to not only the preferences of a service’s direct users or their representatives (e.g. parents of school students) but also the requirements of the wider community (e.g. taxpayers who pay for but hardly access public services).

Furthermore, researchers (e.g. Boyne 2002; Rixon 2010; Harrison et al. 2012) argue that for the public sector, performance evaluation should fully consider democratic values, such as accountability (being answerable for the actions of the public organisations and having redress measures), participation (involving citizens and other external stakeholders in decision-making and implementation processes), and probity (avoiding fraud when using public funds and ensuring the integrity of politicians and officials). The aforementioned indicators are summarised by Boyne (2002) in a table (Table 2.1) for measuring the performance of local governments, as shown below:

Table 2.1 Dimensions and indicators of local government performance

Dimensions	**Indicators**
Outputs	Quantity (in an appropriate level of quality)
Efficiency	Cost per unit of output
Outcomes	Effectiveness, impact, equity, cost per unit of service outcome
Responsiveness	Consumer satisfaction, citizen satisfaction, staff satisfaction, cost per unit of responsiveness
Democratic outcomes	Probity, participation, accountability, cost per unit of democratic outcomes

Source: Boyne (2002)

2.2 Measuring PSP in Chinese provinces

Since the 1980s, a top-down PMS has been used in China as a tool to regulate local governments and cadres (Burns and Zhou 2010). Economic and social benefits for bureaucrats and their departments (e.g. promotion and extra funding) are directly linked to their achievements on key performance indicators (Jing et al. 2015; Chan and Gao 2008). The performance indicators for public services – adopted by both the government and scholars for their research – are covered by the 3Es framework and the IOO model.

Among these indicators, economy (see 3Es framework) or inputs and outputs indicators (see IOO model) are located at the centre of the Chinese government's PSP evaluation framework. There are numerous empirical studies focusing on them, in which the impact of various socioeconomic policies on the size of inputs for healthcare and other services is the key concern (e.g. Heng and Hong 2012; Xu and Lin 2022; Wang et al. 2011; Liu et al. 2020b). However, data is lacking for measuring equity or distributive efficiency at the provincial level, especially in terms of the equity of access to healthcare services among major social groups (i.e., groups by gender, income, and regions) over the selected period for this research (i.e., 2006 to 2017). Similarly, there is no longitudinal publicly available data relating to dynamic efficiency and impact, and it might be difficult to isolate the "real" impact of a service (e.g. healthcare) from other socioeconomic factors (e.g. an unhealthy lifestyle and other socioeconomic influences, such as unemployment) (Hasenfeld 1983).

Furthermore, measuring responsiveness or allocative efficiency requires a precise estimation of all people's subjective feelings on the services provided in an area, for which surveys about consumer, staff, and citizen satisfaction may be suitable (Boyne 2002, Dowding and Mergoupis 2003; Andrews and Entwistle 2013). But there is no such survey for all provincial jurisdictions of China over a long period, although a few studies discuss this issue on a single-provincial or a one-year basis (Chao et al. 2017; Kowal et al. 2011). Finally, regarding democratic outcomes, admittedly, there are more opportunities nowadays for Chinese citizens and external stakeholders (e.g. non-governmental organisations and research institutions) to participate in policy-making and policy evaluation (He 2011; Yu and Ma 2015) under a non-democratic setting (So 2014). However, such democratic attempts mainly occur at grassroots administrative levels (e.g. at the county level or within urban/rural communities) and cannot substantially influence policy-making and policy implementation at higher administrative levels (e.g. prefecture-level cities and provinces), not to mention that there is a lack of nationwide data for measuring such democratic outcomes.

Thus, due to the lack of suitable measurements and data for measuring the full range of PSP dimensions and indicators within Chinese provinces, this doctoral research focuses on two key performance dimensions covered by both the 3Es framework and IOO model, that is, efficiency (i.e., productive efficiency) and effectiveness. These two indicators can be measured with nationwide numerical data – details are given in the Methodology chapter. However, this does not mean that other performance indicators are less important, instead, they should be investigated in further studies focusing on a smaller research background or at a time when data capturing all performance dimensions of healthcare and other services are fully accessible.

2.3 Performance as a global concern: an overview of practices

The performance of public service provision is a timeless topic, as every human society is concerned with 'what commodities are produced, how these goods are made, and for whom they are produced' (Samuelson and Nordhaus 2005, p.7). For citizens in most countries, enjoying public services is now a part of their everyday life

(Walker et al. 2010): national and public safety is guaranteed by the army, local policemen, and fire services department; household waste is collected by the sanitisation sector; most children access the education service provided by public schools; people visit public healthcare institutions to treat illness and seek job opportunities from government employment agencies. At the same time, supporting the operation of public sectors by paying taxes is also an inevitable responsibility. Thus, citizens, as taxpayers, undoubtedly want to pay less on taxes but access public services with a higher level of quantity, efficiency, and effectiveness (Brignall and Modell 2000).

Recognising people's expectations for better healthcare and other public services, in past decades, an increasing number of reforms have been conducted, many inspired by the NPM movement (Ma 2016a; Andrews et al. 2019). As argued by O'Flynn (2007), NPM targeted perceived problems of traditional PM patterns such as bureaucratism, monopoly, and public service inefficiency (O'Flynn 2007). Thus, performance improvement is always one of the key targets in the NPM movement (Hood 1991; Ferlie et al., 1996). Major NPM-inspired reforms in Europe can be classified into five types (Hammerschmid et al. 2019), each of which is aimed at achieving improvements in specific dimensions of performance as shown in Table 2.2:

Table 2.2 Types of NPM-inspired reforms

Type of NPM-inspired reforms	Specific measures	Performance dimensions to be improved by the reforms
Downsizing	Shrink the size of the government.	Economy and efficiency
Agencification	Disaggregate large governmental departments into smaller agencies.	Efficiency and effectiveness
Contracting out	Introduce private organisations in public service provision through contractual mechanisms.	Efficiency
Customer orientation	Treat citizens as customers and be more responsive to their ideas about public service.	Efficiency, effectiveness, and responsiveness
Flexible employment	Draw on experiences of the private sector; break the lifelong employment; introduce competition and performance-based reward system.	Efficiency, responsiveness, and service quality

Source: Hammerschmid et al. (2019)

Apart from the above five reforms, FD has also been increasingly recognised as an NPM-inspired and mobilised initiative. While it is true that FD predates NPM and should not be considered a direct outcome of it, the principles and requirements of NPM have greatly fostered the implementation and generalisation of FD around the world (Alonso et al. 2015). For example, FD plays important roles in facilitating NPM's key agendas such as reducing central controls, enhancing managerial autonomy, and mobilising local professionals (Hope and Chikulo 2000). It also enhances competitions between local governments in different jurisdictions (Borins 1994; Crowley and Sobel 2011). Furthermore, in democratic settings, FD strengthens the accountability between local governments and citizens, encouraging greater citizen participation in local decision-making processes (Hope 1998a). As a result, FD contributes to better governance by enhancing the efficiency, effectiveness and quality of public services, which are all fundamental elements of the NPM framework (Silverman 1992; Borge et al. 2008; Hankla 2009).

Outside of Europe, NPM practices in other OECD countries have also demonstrated increasing attention to PSP improvement. The National Performance Review, a public sector reform initiated by the U.S. federal government in the 1990s, aims to achieve working better with less cost and enhance public service quality (Thompson 2000). The Australian Public Service reforms in the 1990s also intended to enhance productive efficiency and effectiveness – two significant aspects of PSP (Shaw 2012). Similar reforms focusing on PSP promotion such as outsourcing and decentralisation were also conducted by Japanese local governments in the early 2000s, with the inter-regional spread of NPM experience and practices from Europe and the USA (Muramatsu and Matsunami 2003).

In the developing world, NPM initiatives which target performance improvement have also become increasingly popular (OECD 2005; Salman 2021). For example, in Indonesia, NPM-inspired reforms such as decentralisation, result-based management, and performance-based payment have been initiated since the 2000s. Accordingly, policy tools such as the citizen charter – a document clearly specifying government commitments and performance standards, satisfaction surveys, and co-production have been adopted (Brinkerhoff and Wetterberg 2013). In South Africa, in response to the commitment to better service quality, in 2009, the Zuma administration established a resource-based management system which clarified policy goals and targets for governmental ministries (Friedman 2011). In Pakistan, privatisation was initiated in the 1990s. Then, in recent years, outsourcing and public-private partnerships have become increasingly popular in the construction, education, and healthcare sector (Salman 2021; Irfan and Nutley 2016). Similarly, NPM-inspired tools such as privatisation, mergers, and contracting out are initiated in the infrastructure sector of Jamaica (Caribbean Policy Research Institute 2011).

- **Reviewing the impact of NPM-inspired reforms on PSP**

A growing body of empirical studies have examined the effects of NPM on PSP. Many of these studies have focused on specific types of NPM or NPM-inspired reforms, with agencification and private sector involvement being particularly popular, which possibility because they could be easily quantified by accessible secondary data. For instance, Yamamoto (2006) found that agencification significantly improved public service efficiency, quality, and effectiveness in Japan. Similarly, Cingolani and Fazekas's (2019) conducted quasi-experimental research and identified that agencified public organisations demonstrated better value-for-money performance in France, Germany, Spain, and the UK. Furthermore, they found that older agencies exhibited greater efficiency improvements compared with those recently agencified organisations. However, the study of Kim and Cho's (2014) on forty-four executive agencies in South Korea showed a negative relationship between agencification and selected performance indices (e.g., profit ratio, customer satisfaction, and quality of service).

The involvement of the private sector, as another significant empirical concern, has also yielded diverse outcomes. Angrist et al. (2006) examined the school voucher programme in Columbia, which provides financial support for families and students who wish to choose schools by themselves. The empirical results identified an overall improvement in high-school graduation rates and achievement, demonstrating the benefits of fostering public-private competitions. However, similar programmes implemented in Chile failed to enhance average educational outcomes (Hsieh and Urguila 2006). Regarding elderly caring services, evidence from Sweden suggests that the establishment of a quasi-market did not significantly contribute to service performance, as measured by indicators such as citizen satisfaction and the educational level of care sector staff (Broms et al. 2019). Similarly, Alonso et al. (2015) found no significant impact of outsourcing on public sector size in 15 EU countries. Even a negative impact between contracting out and prison healthcare performance was identified by Bedard and Frech (2009) in their US-based study. Privatization practices also resulted in varied outcomes. For example, Bergman et al. (2016) argued that privatisation improved the effectiveness of elderly care service in Sweden, as indicated by lower mortality rates. However, Grabowski and Stevenson (2008) found a negative relationship between privatisation and the quantity and efficiency of resource inputs in US nursing homes.

In recent years, supported by large-scale survey data, few studies have attempted to investigate the impact of multiple NPM reforms across different settings. For example, Andrews and Van de Walle (2013) examined the impact of six types of NPM reforms on citizens' perceptions of service efficiency, responsiveness, equity and effectiveness, supported by survey data collected from residents of all English local governments. The empirical results showed that public-private-partnerships had a negative impact on all four selected PSP dimensions, while an entrepreneurial strategic orientation had a positive relationship with all four. Additionally, performance management significantly improved citizens perceptions on three of the four performance dimensions: efficiency, responsiveness, and effectiveness. However, other three NPM reforms, namely a strong customer focus, the use temporary staff, and the introduction of capital charging, did not improve citizens' perceptions on multiple performance dimensions. Similarly, Hammerschmid et al. (2019) investigated the impact of five key NPM reforms (downsizing, agencification, customer orientation, flexible employment, and contracting out) on four PSP dimensions: efficiency, quality, policy coherence, and coordination. They collected survey data from top executives in central government ministries/agencies of 20 European countries. The structural equation modelling results revealed the diverse impact of NPM reforms on different PSP dimensions. Contracting out and downsizing were found to enhance service efficiency, but downsizing exacerbated service quality. Moreover, agencification did not have a significant impact on service performance.

To provide clarity on the effects of NPM reforms on different performance dimensions, Pollitt and Dan (2011) conducted a comprehensive review of 518 journal articles and reports covering 15 types of NPM tools and mechanisms in 26 European countries and European Commission since 1980. The effects of NPM reforms on organisational processes or inputs, service outputs (including cost efficiency), and outcomes (including cost effectiveness) were all considered. Meta-analysis results indicated that among studies focused on inputs, 57.9% reported positive findings, 18.5% reported negative findings, and 23.6% had uncertain results. However, for studies focused on outputs and outcomes, higher proportions reported negative or uncertain results. Among output-focused studies, the ratios were 53.4% (positive), 19.6% (negative), and 27.9% (uncertain), while for outcome-focused studies, the percentage of positive findings reduced to 43.9%, with 22.8% reporting negative findings and 33.3% uncertain findings. Similarly, another meta-analysis research of (Funck and Karlsson 2019) covering 299 peer-reviewed articles under global backgrounds between 1991 and 2016 also revealed that among the papers specifically discussing NPM’s effects, 50% of them had a negative research tone.

The aforementioned empirical and reviewing studies demonstrate the varied outcomes of NPM on PSP dimensions, which, to some extent, aligns with people’s growing criticism towards NPM in recent years (Reiter and Klenk 2019). However, as explained by Wollmann (2003), public management reforms inevitably face paradoxes, trade-offs, and conflicts between different performance dimensions. In this case, NPM reforms typically focus on addressing the most enduring issues in the public sector rather than solving all performance dilemmas. Thus, NPM should not be viewed as a universal solution for all performance challenges. Furthermore, the outcomes of NPM reforms are related to various contextual factors. For example, Lapuente and Van de Walle (2020) indicate that NPM reforms may be better suited to Anglo-Saxon and Scandinavian administrative cultures. Singh (2007) and Sulle (201) argue that in developing countries, considering local managerial and institutional cultures is crucial for the success of NPM-inspired reforms. Ware et al. (2007) and Bauhr et al. (2020) demonstrate the importance of information transparency in public-private competitions. Additionally, managerial factors such as public managers’ previous experience (Van de Walle et al. 2020) and personal attitudes (Kudo 2016) towards management innovations also play a role in determining reform outcomes.

In summary, there is no one-size-fit-all solution when implementing NPM-based initiatives, and a careful consideration of local situations should always be a guiding principle when analysing an NPM-inspired reform under any settings, including China. As explained in the Introduction chapter, the aim of this study is to discuss the impact of intra-provincial FD, an NPM-inspired reform, on the efficiency and effectiveness of healthcare in Chinese provinces. Thus, before delving into the specific effects of FD, it is essential to provide an overview of the broader institutional settings and discuss the rationale and attitudes of public service providers towards PSP enhancement within these settings. To achieve this, the study applies four social science theories: public choice theory, legitimacy theory, accountability theory, and stakeholder theory.

2.4 Theoretical underpinnings of the pursuit for better PSP

- **Public choice theory**

As mentioned above, the NPM movement has dominated public sector reforms in most countries across the world during the past three decades. To a large extent, the arguments for PSP improvement in public management are underpinned by the public choice theory which emerged around the 1960s and became prevalent at the same time as NPM. As Gruening (2001) explained, many solutions for PSP promotion in NPM-inspired reforms, such as using market mechanisms and reducing government size, are closely associated with the arguments of the public choice theory that bureaucracies are inherently inefficient and that in large democratic units of government exploitation of minorities by the majority is common.

Based on the assumption of 'rational egoistic maximisation', the public choice theory argues that governments and bureaucrats are concerned with three major types of competition: 1) competition between different regions in a country (e.g. Tiebout 1956), 2) competition between political parties within a region (e.g. Downs 1957), and 3) competition between public organisations and private organisations regarding public service provision. For bureaucrats, their benefits could be maximised by winning elections, attracting population (and tax) inflow, and undertaking greater responsibilities in public service provision. Thus, in a setting with the above three types of competitions, it is of great importance for bureaucrats to improve PSP, so as to win elections, attract population (and capital) inflow, and undertake more local responsibilities (Tiebout 1956; Ostrom et al. 1961; Mueller 1979; Boyne 1998).

As one of the earliest theories explaining political behaviour under competition, the public choice theory has been widely supported by international practices in the real world – a 'natural laboratory' (Boyne 1996). However, in China, competition between political parties does not exist, as the long-term ruling position of the Chinese Communist Party (CCP) has been confirmed by the Constitution. Also, the private sector provides very little competition to public service provision in China. Between the 1980s and early 2000s, many Chinese local governments transferred their responsibilities for providing basic public services (e.g. medical service and education) to private enterprises, as a marketisation attempt (Dong et al. 2015). However, widespread social problems such as regional inequality resulted, leading to an end of the public service privatisation trend (Li and Zhu 2004). Nowadays, the private sector usually performs only as a partner of public organisations in public service provision through mechanisms such as the Public-Private-Partnership (Tan and Zhao 2019), rather than being a competitor which provides the same public products with a competitive price or quality. Thus, discussions about public choice in the Chinese context usually revolve around the aforementioned first type of competition, i.e., inter-jurisdictional competition for population, which is also described as "residential sorting" (Brouhle et al. 2005).

As explained by Tiebout (1956), in a setting with a high level of information transparency, sufficient mobility, enough jurisdictions, and low transaction costs, citizens could 'vote with their feet' by moving to the jurisdictions where their preferred package of local public services are provided (Gill and Rodríguez-Pose 2012). In China, people's mobility and transaction costs are determined by the household registration system (the "Hukou" system). As explained by Song (2014), each Chinese citizen has a Hukou, which, in past decades, imposed strict limitations on migration between different jurisdictions and between urban and rural areas. However, nowadays, such limitations have been partially relaxed, creating greater opportunities for population inflow and outflow (Chan and Buckingham 2008). By attracting migrants, the local economy could be benefited, which supports local cadres to win the "GDP tournament"[5] (Li and Zhou 2005; Zhao and Li 2021). The

---

[5] Because the local economy – measured by GDP is also an important indicator in China's

performance of healthcare and other wellbeing-related services is viewed as a key consideration for migrants - particularly those from better socioeconomic backgrounds (Zhang et al. 2017). Thus, it is of great importance for the government and cadres to promote PSP – so as to facilitate socioeconomic growth with the support of those newcomers (Cai and Wang 2008; Chan 2012; Shen and Li 2020).

- **Performance legitimacy**

Legitimacy is a central topic in the theory and practice of politics. For people in power, controlling and strengthening legitimacy is always a key concern. The legitimacy of a state, broadly speaking, refers to citizens' acceptance and support of the state's right to rule – in other words, state legitimacy is about the "rightfulness" of the regime (Mcloughlin 2015; Levi et al. 2009). As explained by Beetham (2012), different types of regimes have different sources of state legitimacy. Regarding liberal democratic states, it is usually argued that their political authority derives from the people, and their legitimacy to rule mainly comes from free and competitive elections. However, in recent years, the role of another type of legitimacy, namely performance legitimacy, has been increasingly investigated.

Despite the lack a unified definition, previous studies explain that performance legitimacy is determined by the state's outputs (OECD 2011). In other words, performance legitimacy is guaranteed by satisfying people's basic demands which, obviously, include the need for public services (Mcloughlin 2015). From a Weberian

---

top-down PMS, ambitious local cadres have to participate in the GDP tournament, that is, to perform better than their counterparts in other jurisdictions and win opportunities for promotion and other high-powered incentives (Li and Zhou 2005).

perspective, a political system has an input side where legitimacy is gained through elections and a high level of representation and an output side where legitimacy is determined by "good governance" – i.e., the performance of policy implementation (Rothstein 2009). One of the key responsibilities of the government is to provide public services, thus, delivering better services – particularly those "tangible" services such as healthcare, childcare, and education – is of vital importance to performance legitimacy (Rothstein 2012; Whaites 2008). By showing people, as taxpayers, that their money has been used efficiently and effectively (OECD 2011), performance legitimacy is promoted so that the society is "bound" with the state (Milliken and Krause 2002). In this case, citizens are willing to comply with the laws and rules of the state, leading to less conflict (Mcloughlin 2015; Dagher 2018).

Some researchers argue that performance is potentially an even more important source of legitimacy than democracy, and that any "deficit" in democratic legitimacy could be remedied by promoting performance legitimacy (Gilley 2006). This argument seems to be particularly generalisable to the Chinese context. As mentioned in the first chapter, the CPC is the only legal ruling party of all levels of governments in China, and competitive elections hardly exist (Chu 2013). However, it does not mean that the legitimacy of the CPC and the CPC-led government (i.e., the "party-state") is beyond challenge and that performance is irrelevant. Instead, better performance in fulfilling responsibilities such as delivering public services is now considered the dominant source of legitimacy in contemporary China (Duckett and Munro 2022).

Originally, the legitimacy of the CPC's long-term rule came from its success in the Chinese revolution as well as through Communist ideology (Zheng 1999; Beetham

2012). However, without satisfying people's emerging demands for basic needs, Communist ideology was challenged, leading to a greater appeal for new sources of legitimacy (Wang 2013). In response to the legitimacy risk, the CPC initiated economic reform in the late 1970s. Since then, the performance of Chinese governments in "doing the right thing" (i.e., promoting economic development which delivered tangible benefits for the people) has effectively consolidated the legitimacy of the Communist regime, which can be illustrated by the results of public surveys during the 1990s (see for example Kennedy 2009; Shi 2001; Duckett and Munro 2022). In other words, without abolishing Communist principles (Holbig and Gilley 2010), performance has become the key source of state legitimacy.

However, an excessive focus on the economy brought a series of problems, such as corruption, regional inequality, and the degradation of public services such as healthcare, education, and environmental protection. Particularly after the late 2000s, with the weakening of economic growth, the performance legitimacy of the Chinese regime started to be challenged by the above social risks. Facing greater difficulty in directly intervening in the economy after the marketisation reforms and concerns about regime change (Zhu 2011; Ratigan 2022), Chinese governments opted to regain performance legitimacy by satisfying people's emerging needs in not only better economic performance but also better public services such as healthcare. This approach is summarised by the CPC as the achievement of "good governance" (Castells 1992; Duckett and Wang 2017), and it shows how better PSP can potentially be deployed to strengthen the legitimacy basis of the CPC and the Chinese government (Zeng 2014). In fact, top leaders of the CPC and the Chinese government initiated various political agendas to achieve this aim, such as Hu Jintao's "Scientific

Outlook on Development" and "Harmonious Society". Also, in the contemporary era of Xi Jinping, locating better PSP at the centre of governmental activities is still a core theme of "Modernising Governance System and Capacity" – one of Xi's top political agendas (Zhao 2009; Li et al. 2012; Jing et al. 2015). Thus, in the foreseeable future, better PSP will be an increasingly important source of performance legitimacy (Nathan 2003; Zhu 2011), which, to a large extent, explains the reasons for Chinese governments to work harder on promoting the performance of healthcare and other public services.

- **Accountability and stakeholder theory**

The theory of accountability lies at the centre of liberal democratic thought and can be usefully applied to the promotion of PSP (Kluvers 2003). In the private sector, accountability is about providing an account or information in terms of the resources of an organisation (Stewart 1984). While in the public sector, the concept of accountability goes beyond the report of financial information and pays greater attention to whether organisational decisions and actions meet performance criteria, such as efficiency, effectiveness, democracy, and transparency (Sinclair 1995; Bovens 2005). The theory of accountability has two sides: the side "giving an account" is called the accountor, while the side "being held to account" is the accountee (Pollitt 2003, p. 89). For example, in today's practices, citizens and politicians, politicians and bureaucrats, and senior bureaucrats and their subordinates could be seen as pairs of accountors and accountees (Stone 1995; Mulgan 1997; Aucoin and Heintzman 2000). More detailed classifications of the forms of accountability are shown in Table 2.3.

Table 2.3 Forms of accountability and explanations

Forms of accountability	**Explanations**
Political accountability	A hierarchical/vertical accountability system (Christensen and Laegreid 2012; Bovens 2007) which is influenced by principal-agent relationships (Strom 2000), such as the accountability between voters and their representatives in parliament between representatives and cabinet ministers, and between ministers and civil servants.
Legal accountability	Accountability to the court, dealing with questions of rule of law, rights, and equal treatment (Christensen and Laegreid 2012)
Managerial accountability	The answerability of managers for the attainment of organisational goals and results of organisational activities in comparison with the agreed performance criteria (Day and Klein 1987)
Professional accountability	Accountability to professional organisations which set codes and practical criteria for all members. It is rather important for public organisations with professional functions such as hospitals (Bovens 2007).
Social accountability	Accountability to external stakeholders such as public service users, citizens, non-governmental organisations, and media. It requires public agencies to be obliged to account for their performance in conducting daily functions (McCandless 2001).

Based on the different relationships between accountors and accountees, accountability can be classified into three categories: vertical accountability, horizontal accountability, and diagonal accountability (Bovens 2005, 2007; So 2014). Vertical accountability includes those forms of accountability in which accountors and accountees are located in different layers of a hierarchical system, such as political accountability and legal accountability within the government and the court. Horizontal accountability refers to those relationships in which accountors and accountees are not located in a formal hierarchy, such as social accountability between public service users and institutions which provide public services (Joshi

2017). Finally, the accountability relationships that are in an intermediary form are called diagonal accountability. For example, ombudsmen, audit offices, and supervisory authorities have no direct relationship with public organisations and have limited enforcement power. However, they could hold public organisations accountable by reporting to ministers and/or the parliament which has a direct accountability relationship with public organisations. Such an accountability relationship established by two steps is described as diagonal accountability (Bovens 2007).

Within political systems, accountees can be regarded as stakeholders (Sinclair 1995; Barberis 1998; Barrett 2001; Andre 2010; So 2014). Freeman (1984, p.46), the founder of stakeholder theory, explains that a stakeholder refers to "… any group or individual, who can affect or is affected by the achievement of organisation's objectives". In the context of public services, researchers (e.g. Walker and Wu 2010; Zheng et al. 2019) have listed various typical internal stakeholders (individuals and entities within the public organisation) and external stakeholders (individuals/entities outside the public organisation) to whom public organisations are accountable, as shown in Table 2.4.

Table 2.4 Internal and external stakeholders of public organisations

Internal stakeholders	• Bureaucrats or managers and employees within the public organisation • Professional associations and trade unions
External stakeholders	• Other governmental institutions such as the central government and local governments in other jurisdictions • Elected representatives in legislatures • Internal and external inspection and audit agencies • Local residents and their representative groups • Private sector entities such as foreign investors • Other non-profit organisations • Mass media

Ideally, public organisations should be accountable to all stakeholders (Yuesti et al. 2016). However, this cannot be achieved in the real life, which is not only because stakeholders themselves might be irrational or unreliable (Schalk 2011; Yu and Ma 2015) but also due to the difficulty in satisfying all the requirements of all stakeholders (Chun and Rainey 2005; Andrews et al. 2010). Thus, in real-world practice, being accountable to key stakeholders, i.e., those having a decisive influence on the accountees or those "legitimacy-conferring" stakeholders, is always the primary concern (Farrell and Jones 2000; Harrison et al. 2012). For example, in a democratic setting, political economists argue that politicians are often more accountable to the largest group of voters, i.e., "median voters" (Cho and Duggan 2004). While in recent years, being accountable to minorities who have high stakes in the stability of the regime is also emphasised (Hänni 2017).

So why should public sector organisations be accountable to their stakeholders? Normative perspectives explain that it is an ethical choice and forms the basis of trust in society (Barrett 2001; Kluvers 2003). While in contrast with the normative perspectives, the institutional approach suggests that being accountable is primarily due to its contribution to mutual benefits (Barrett 2001). As highlighted by Campbell (1997) and Farrell and Jones (2000), a company or a public organisation cannot survive without being accountable to its chosen stakeholders, particularly their key stakeholders. A greater level of accountability, in return, strengthens key stakeholders' support for policies and reduces the unnecessary costs and uncertainties in policy implementation, which underpins the basis of regime legitimacy and benefits the interests of both accountors and accountees (Provan and Kenis 2008; Harrison et al. 2012; Wang et al. 2015). While in recent years, with better education, stronger interest groups, and more aggressive and intrusive mass media, citizens – key stakeholders of the government and politicians in a democratic setting, have been paying greater attention to the performance of public services (Aucoin and Heintzman 2000). In summary, the above explanations about accountability, stakeholder theory, and legitimacy demonstrate the importance for politicians and bureaucrats to improve the performance of healthcare and other services. That is, better performance ensures the accountability for key stakeholders' requirements, which, in turn, underpins the basis of state legitimacy.

- **Understanding the rationale behind Chinese local governments' promotion of PSP: a Chinese-based interpretation of accountability, stakeholder theory, and legitimacy**

The above integrated analysis of accountability, stakeholders, and legitimacy are also implicative for interpreting the importance of PSP for Chinese local governments. In China, despite the growing influence of external stakeholders such as citizens, non-governmental organisations, and public media over the past years (So 2014; Yu and Ma 2015), without powerful mechanisms such as multi-party competitions and external oversight, the primary stakeholders of lower-level governments (e.g. the prefecture-level city's government) are still the higher-level governments (e.g. the provincial government) (Deng 2018). Consequently, local governments and leaders are mainly accountable to the upper-level government and superiors. Chinese researchers have conceptualized this phenomenon as "upward accountability" (Lam 2010; Chien 2010).

Since the 1980s, key arrangements of upward accountability in China have exhibited clearer features of NPM and increasingly relied on well-designed PMS (Chan and Rosenbloom 2010). Through the process of 'Reform and Opening-up' (*Gaige Kaifang*), Chinese top leadership became exposed to western concepts and incorporated NPM inspirations into their design of inter-governmental fiscal and administrative relations (Caulfield 2006; Chan and Chow 2007; Christensen et al. 2008). Consequently, higher-level government gradually withdrew from local affairs and decentralised substantial power to bureaucrats and professionals in terms of local services (Wu et al. 2017). This enabled lower-level governments (i.e., 'accountees', in relative to higher-level governments as ‘accountors’) to take charge of formulating and implementing local policies based on their understanding of local conditions. Simultaneously, in order to maintain political control over decentralized local bureaucrats, Chinese top party-state leaders strengthened the accountability link with a more robust top-down PMS (see page 21 and Section 4.1.4) that, to a large extent, still in line with NPM’s emphasis on performance/result-based management and incentive-based motivation (Common 1998; Chan and Rosenbloom 2010; Walker and Wu 2010). Within the PMS framework, the promotion, demotion, and other high-powered incentives (e.g. bonuses and additional grants) for lower-level cadres are all determined by the higher-level government (Gao 2009; Liang and Langbein 2015).

In summary, the system of upward accountability, facilitated by decentralization and an effective PMS, ensures that the agendas and expectations of the central government are successfully communicated to lower-level governments and functional agencies. This allows for the formulation and implementation of detailed policies that take into account local circumstances, thereby meeting the expectations and agendas of higher-level authorities (Chien 2008). As mentioned in the above descriptions of legitimacy theories, apart from economic performance, the central government and the CPC have increasingly prioritised PSP for the sake of performance legitimacy (Duckett and Wang 2017). Accordingly, under the upward accountability system, in order to perform better in the top-down performance evaluations and seize more opportunities for promotion and other benefits, it is of vital importance for local governments to demonstrate their commitment to meeting the higher-level governments’ expectations of improved healthcare and other services (Zhang 2020).

2.5 Summary

This chapter started with an introduction to the concepts and key dimensions of PSP. This was followed by explanations of PSP measurements and an overview of practices for PSP improvement in global and Chinese contexts. After that, this chapter drew on four social science theories, i.e., public choice theory, legitimacy theory, accountability, and stakeholder theory to explain why governments in China (and elsewhere) focus on improving the performance of healthcare as well as other public services. Recognising the importance and necessity of PSP improvement, in past decades, various policies have been initiated around the world, especially FD. Thus, in the next chapter, the theoretical underpinnings and empirical evidence about the FD-PSP (i.e., efficiency and effectiveness) relationships will be reviewed in detail.

CHAPTER 3 FD AND PUBLIC SERVICE PERFORMANCE: A REVIEW OF THEORETICAL AND EMPIRICAL STUDIES

In the past decades, people have witnessed the spread of decentralisation initiatives around the world (Rodríguez-Pose and Gill 2003), many of which are concerned with the decentralisation of spending and revenue-raising responsibilities in the public sector (Ter-Minassian 1997; Rodríguez-Pose and Gill 2003; Reingewertz 2014). Along with its theoretical development, FD was introduced to countries such as Italy, Spain, and Australia in the 1970s (Rodden 2003; Stegarescu 2005; Blöchliger 2006; Bodman and Hodge 2010). Promoted by international institutions such as the World Bank and the OECD (World Bank 1999; OECD 2016), FD was then spread to Latin American and Eastern European countries from the late 1980s, with the aim of building modern intergovernmental relations, promoting good governance, and achieving long-term socioeconomic development (Rezk 1999; Wiesner 2003). Starting from the 1990s, developing countries such as India and Indonesia also adopted FD to replace their traditional centralised fiscal systems and motivate market-oriented reforms (Fukasaku 1999). In many countries, FD in the healthcare sector is perceived as a central issue in their decentralisation reforms (Mosca 2006). Although the practices in various contexts are different, a shared theoretical underpinning can still be identified from the FD practices, namely fiscal federalism (FF) (Ter-Minassian 1997).

As a subfield of public finance, FF is concerned with how spending and revenue power and the supporting instruments for such fiscal assignments are arranged vertically across different levels of government (Musgrave 1959; Oates 1999; Boadway and Shah 2009). With a well-designed arrangement, government responsibilities can be effectively and efficiently conducted to satisfy the preferences of citizens in different jurisdictions, thus maximising public welfare (Oates 2005; Boadway and Shah 2009). As Musgrave (1959) explains, the responsibilities of the government include socioeconomic stabilisation, redistributing of resources to groups and individuals, and providing public services. To achieve the goal of welfare maximisation, the theories of FF acknowledge that fiscal authority related to macroeconomic stabilisation and redistribution should be kept at the national level. This is due to the nationwide impact of these two responsibilities: the central/federal government can better represent the interests of all citizens in a country, and the fiscal and regulatory capability of local governments to stabilise the macroeconomy and redistribute resources is usually constrained (Oates 1972). However, the central argument of FF is that fiscal authority and responsibilities related to public service provision should be decentralised to local governments, apart from the power for providing nationwide services such as national defence (Oates 1999). Over the past decades, extensive research has been conducted on theories of FF. Furthermore, FF has been developed to two generations based on the different perspectives on the altruistic or egoistic nature of local bureaucrats (Oates 1972, 1999; Qian and Weingast 1997). Although both two generations of FF have shed lights on the relationship between FD and public service efficiency (and effectiveness), there remains a dearth of empirically discussions on this issue in various contexts. Particularly, the impact of FD on the efficiency and effectiveness of healthcare provision is largely ignored in the empirical literature. Recognising this issue, this doctoral study specifically focuses on the healthcare sector and, in the following sections, existing theoretical and empirical studies regarding FD's effects on public service efficiency and effectiveness will be presented.

3.1 The impact of FD on public service efficiency and effectiveness: arguments from theoretical studies

Since the 1950s, scholars in the field of public economics have claimed that FD is beneficial for social welfare because governments closer to local residents are more willing and able to undertake actions likely to promote PSP (Samuelson 1954; Musgrave 1959; Arrow 1971). The above argument (which is called the Arrow-Musgrave-Samuelson perspective) was refined to underpin the theories of first-generation fiscal federalism (FF 1.0) by Wallace Oates (Oates 1972). However, from the 1980s, theoretical assumptions of FF 1.0, especially the altruistic nature of government, started to be challenged by second-generation fiscal federalism (FF 2.0, see Qian and Weingast 1997 and Weingast 2009). Notwithstanding these differences, the potential benefits of FD on PSP have remained a common focus point of the FF 1.0 and FF 2.0 theories. Therefore, in this section, theoretical arguments from the two generations of FF will be presented respectively.

3.1.1 FD and public service efficiency: arguments from FF 1.0

The first generalisation of FF claims that public service efficiency increases with the decentralisation of fiscal responsibilities (and revenues) to lower-level governments

(Oates 1972). Consistent with the Arrow-Musgrave-Samuelson perspective, theories of FF 1.0 have three basic presumptions: first, governments at all administrative levels always pursue the maximisation of public interests within their jurisdictions. Second, vertical information asymmetry exists within the public sector, which makes it harder for higher levels of government to know what is happening at the local level. Third, if the central government is responsible for a specific public service, the scale of service output should be uniform in all jurisdictions (Oates 1972, 1999, 2005). Underpinned by the above presumptions, FF 1.0 proposes a positive relationship between FD and allocative efficiency (i.e., the extent to which the supply of public services matches the demand, see Andrews and Entwistle 2013) and between FD and productive efficiency (i.e., the extent to which the outputs are produced for a minimised level of inputs, see Andrews and Entwistle 2013).

FD can enhance allocative efficiency because local governments, compared with the central/federal governments, are more familiar with diversified preferences for local public services (Oates 1972; Balaguer-Coll et al. 2010). In other words, FD, no matter the decentralisation for local responsibilities or revenues, can support local governments to fully apply their information advantages to public service provision. Hence, local demands and preferences for public services can be better satisfied with the decentralisation of fiscal power, and so FD increases allocative efficiency (Andrews and Entwistle 2013) and social welfare (Oates 1999; Barankay and Lockwood 2007).

Productive efficiency can also be promoted under FD. As Oates (1999) suggests, local governments have information advantages about not only the localised demands but also the ways in which local resources can be fully utilised to produce local services in the most cost-efficient way. That is to say, even for some public services with similar demands in different jurisdictions, decentralising fiscal responsibilities or revenues also brings better productive efficiency. Local bureaucrats can reduce the costs of input at a given output or increase outputs at a given input (e.g. by locally purchasing materials with cheaper prices and lower logistic costs) with the support of their information advantages (Andrews and Entwistle 2013).

3.1.2 FD and public service efficiency: arguments from FF 2.0

It is widely acknowledged that FF 1.0 theories clarified the positive relationship between FD and public service efficiency for the first time (Oates 2005). This has not only influenced academic studies on the impact of FD but also strongly inspired FD practices around the world (Qian and Roland 1998). However, from the 1980s, the theoretical basis and presumptions of the early-stage FD theorem were increasingly challenged by the emerging FF 2.0 theories. Similar to FF 1.0 theories, FF 2.0 still suggests that information and technical advantages provide local governments with greater capacity to improve allocative and productive efficiency. However, new perspectives in regard to the motivations of local governments to improve efficiency were developed. Moreover, circumstances where FD might undermine public service efficiency were also identified. Accordingly, both supporting arguments and negative concerns regarding FD’s impact on efficiency in FF 2.0 theories are presented in the following paragraphs.

3.1.2.1 Theoretical basis of FF 2.0

Denying the presumption of altruistic government within FF 1.0 theories, FF 2.0 suggests that the local government is a "Leviathan" composed of politicians and bureaucrats aiming to extract as much rent as possible from citizens and the economy (Brennan and Buchanan 1980; Oates 2005). However, the pursuit of self-interest, as explained by the public choice and FF theories, is still likely to benefit public service efficiency and social welfare (Qian and Weingast 1997; Baskaran 2010). According to the public choice theory, citizens can vote out bad incumbents by using their ballots or "voting with their feet" – moving to other jurisdictions (Tiebout 1956). The above two behaviours can therefore harness the private interest of politicians and bureaucrats to the interests of local citizens. Besley and Case (1995) suggest that with FD, local politicians have to make more efforts to avoid being voted out through the election system or by population outflow with the support of their local advantages. Such efforts, as indicated by Qian and Weingast (1997), can not only protect the private interests in the public sector but also maximise public welfare by promoting public service efficiency. In the following paragraphs, the associations between FD, efforts made by local governments, and public service efficiency are explained in detail.

3.1.2.2 FD and public service efficiency: theoretical contributions of FF 2.0

As the level of FD grows, the local government becomes increasingly accountable for people's public service demands and preferences, which consequently leads to an increase in allocative and productive efficiency (Salmon 1987; Oates 2005). As suggested by Seabright (1996), one of the core considerations of FD research is to justify the connections between FD and the greater accountability of local government, in which the following three mechanisms are particularly significant (Boadway and Tremblay 2012):

First, FD can strengthen the link between local public services and locally elected politicians. Under a fiscally decentralised setting, the local government controls more financial resources and undertakes greater responsibilities of providing local services. In this case, PSP is expected to become a decisive consideration for local citizens to vote in local elections or “vote with their feet” (i.e., moving out to other jurisdiction with better levels of service performance) (Tiebout 1956; Boadway and Tremblay 2012). Clearly, local politicians’ benefits may directly suffer from losing elections. Also, population outflow is associated with economic depression which may damage the interests of local politicians (and bureaucrats) in the long run. Therefore, under a fiscally decentralised setting, to avoid being voted out or coping with dwindling tax revenues from population decline, self-interested politicians have to be more responsive to, and accountable for citizens’ preferences in local services (Persson and Tabellini 2000; Lindaman and Thurmaier 2002; Besley and Coate 2003; Alonso and Andrews 2019). In this case, local politicians tend to make greater efforts, with the support of their greater capacity over the central government, to promote the productive and allocative efficiency of local services (Oates 2005).

Second, even for those local services with interjurisdictionally similar preferences, under FD, local governments’ productive efficiency can also be promoted by fiscal migration through the increasingly fierce yardstick competition (Salmon 1987;

Baicker 2001; Boadway and Tremblay 2012). When fiscal power regarding public services with interjurisdictionally similar preferences is decentralised, citizens (voters) tend to compare the inputs and outputs of services in their jurisdiction with those in other nearby jurisdictions (Tiebout 1956; Barankay and Lockwood 2007; Sow and Razafimahefa 2015). As a result, greater pressure might be given by citizens to require their local government to catch up with its "competitors" – other local governments which provide the desired level of outputs with lower inputs (Bordignon et al. 2004). To win the yardstick competition, decentralised local governments need to be more accountable for citizens' expectations of the input and output size, thus improving efficiency (Ahmad et al. 2008; Cavalieri and Ferrante 2016).

Apart from the above two mechanisms, FD can support local governments to build channels for citizens to monitor and participate in policy processes related to local public services, thereby further strengthening government accountability and improving efficiency (De Mello 2011). In the beginning, such channels (e.g. involving citizens to serve on local advisory bodies and participatory budgeting events) are mainly used for eliciting citizens' preferences regarding local services, but they can subsequently be used by citizens to directly participate in policy-making processes and scrutinize the usage of decentralised public money (Seabright 1996; Schaltegger and Torgler 2007; Boadway and Tremblay 2012). In this case, greater public participation means that a local government has to be even more accountable for the decisions regarding local public services and money, leading to an increase in both allocative and productive public service efficiency (Putnam 1993; Rodríguez-Pose and Bwire 2004; Schaltegger and Torgler 2007).

The above theoretical explanations, in summary, suggest that FD strengthens the link of accountability between local governments and local residents in terms of public service provision. According to empirical studies, under a decentralised setting, such greater accountability is mobilised by policy changes to finally contribute to a higher level of public service efficiency. First, FD motivates local governments to initiate more cost-reducing innovations and experimentations for public service provision (Oates 1999; Rodríguez-Pose and Gill 2005). Those innovations can increase the levels of allocative and productive efficiency (Boadway and Tremblay 2012). Moreover, Balaguer-Coll et al. (2010) indicate that the public money saved by those innovations can be retained in the public sector as it is private "revenue", which also corroborates that the pursuit of innovations can be incentivised under FD (Rodríguez-Pose and Gill 2005). In other words, the interests of both citizens and the government can be promoted.

Second, under FD, more efforts need to be taken to strengthen fiscal discipline and reduce corruption, leading to better public service efficiency (Putnam 1993; Ebel and Yilmaz 2002; Balaguer-Coll et al. 2010; Alonso and Andrews 2019). In order to be more accountable for citizens' demand for cost-efficiency, fiscal discipline is essential to reduce the waste and misuse of public service inputs (Seabright 1996; Boadway and Tremblay 2012). Moreover, channels of communication and participation between citizens and the local government established and consolidated under FD can also enable citizens to monitor the daily operations of the local government (Putnam 1993; Ebel and Yilmaz 2002). In response to more intensive citizen supervision under

FD, a local government has to put more effort to reduce corruption and rent-seeking behaviours in public service provision (Seabright 1996; Persson and Tabellini 2000), which increases public service efficiency.

- **Negative effects of FD**

The above arguments from the FF 2.0 theories illustrate the contributions of FD to public service efficiency. The circumstances where FF 2.0 suggests FD might be detrimental to public service efficiency are analysed in the following paragraphs.

For local services with inter-jurisdictionally homogenous demand and few mechanisms preventing citizens from accessing services in multiple jurisdictions, decentralisation of fiscal responsibilities might generate interjurisdictional spillover effects[6] (Tanzi 1995; Boadway and Tremblay 2012). This raises questions about the appropriate level of service provision for more efficient local governments (Oates 2005; Boadway and Shah 2009). Local governments, whether private-interest maximisers or entities pursuing the interests of local citizens, do not aim to benefit the residents from other jurisdictions (Oates 2005; Besley and Coate 2003). To minimise the use of services by non-residents, more efficient jurisdictions might deliberately decrease the overall scale of output. Such a negative change makes it harder for local residents to access those services that satisfy their preferences – due to the excessive shrink of output scale, and moreover, the productivity of the service might also be reduced (Boadway and Shah 2009).

[6] i.e., people access public services in neighbouring regions

Regarding the solutions to spillover effects under FD, central government provision of matching grants for local governments is suggested as a feasible option. Inspired by the usage of Pigovian taxes to internalise negative externalities (Johansson 1997), Oates (1972, 2005) claims that the benefits from matching grants can effectively offset local governments' concerns regarding spillovers, which will subsequently lead to an increase in service efficiency. The above argument is also supported by FF 2.0 researchers (Besley and Coate 2003; Oates 2005), since matching grants can be beneficial to both the public and the private interests of the local government.

However, when the spillover effect of one local service becomes too substantial, recentralisation might become a better solution (Oates 2006). A strong spillover effect regarding one public service is likely to occur in a small area with many jurisdictions and an identical preference for this service (Besley and Coate 2003). In this case, the increase of allocative efficiency, which derives from the better supply-demand match under the localised provision, tends to be limited, while the losses from spillovers become larger (even larger than the efficiency gains) at the same time (Oates 2006). Thus, centralisation of the fiscal power to a higher-level government which rules all the regions with the same preference might be a better option than using matching grants to eliminate spillovers (Besley and Coate 2003). Moreover, under this circumstance, fiscal recentralisation might also improve productive efficiency. This could be achieved by increasing output under a given input – as spillover is no longer a concern in a centralised setting or by reducing the administrative costs of those lower-level governments – as now they have fewer responsibilities (Hughes and

Edwards 2000; Balaguer-Coll et al. 2010). Accordingly, many researchers interpret the improvement of productive efficiency as the achievement of economies of scale (Balaguer-Coll et al. 2010; Sow and Razafimahefa 2015). Alternatively stated, in a small area with many jurisdictions, if there is a public service for which people share a common preference, decentralising the relevant fiscal responsibilities will sacrifice the efficiency gains from economies of scale (Prud'homme 1995).

Another situation in which both the decentralisation of responsibilities and revenue can have negative effects is when there is a lack of well-established political and regulation systems at the local level (Sow and Razafimahefa 2015). Analysis of FD takes the ideal political and regulation system for granted (Barankay and Lockwood 2007), but in the FF 2.0 research, adverse effects of political and regulation dysfunctions on the relationship between FD and efficiency have been identified. For example, in underdeveloped countries without a transparent electoral system and an effective system for citizen supervision, decentralising responsibilities and revenue-raising power may encourage local bureaucrats to grasp personal interests by distorting tax structures and conducting rent-seeking behaviours. This would lead to lower levels of allocative and productive efficiency (Treisman 2002; Gong 2006; Martinez-Vazquez and McNab 2003). Similarly, Alesina (1999) indicates that without an effective political and bureaucratic supervision system at the local level, FD can enable local elites to control policies related to public services, which is likely to undermine allocative efficiency.

In the above paragraphs, theoretical contributions regarding the impact of FD on public service efficiency are reviewed, in which a positive relationship between FD and public service efficiency is frequently claimed. Although the decentralisation of fiscal power may not necessarily lead to an increase in public service efficiency, to a large extent the emergence of a potentially negative relationship between FD and efficiency can be attributable to the weakness in local institutions in the implementation of FD policies, rather than the inherent problems of FD itself. These arguments provide potential insights into the key research question of this thesis. That is to say, FD may have a positive effect on healthcare service efficiency. Considering that the aim of improving the impact of both inputs and outputs is to achieve the formal objectives of public services – to maximise public service effectiveness (Boyne 2002), in the following paragraphs, the relationship between FD and public service effectiveness is discussed in detail.

3.1.3 FD and public service effectiveness

Different from research on FD's impact on public service efficiency, there is a lack of theoretical analysis regarding the relationship between FD and public service effectiveness (Freinkman and Plekhanov 2009). Nevertheless, researchers indicate that the theoretical links between FD and efficiency also connect FD and effectiveness together, especially the voting mechanisms (Freinkman and Plekhanov 2009; Arends 2017). That is to say, with FD, citizens are able to make voting decisions (voting with their feet or voting with ballots) by fully considering the local government's performance in public service provision. In this case, to avoid population outflow and win elections, the local government, with their better capacity in managing local information and resources, makes greater efforts to promote PSP. Considering that formal service objectives are undoubtedly a key element in PSP, a positive relationship between FD and greater public service effectiveness is therefore expected.

Other mechanisms enhancing the connections between FD and effectiveness are underpinned by the voting systems (ballot or foot voting). The first mechanism is the channel for citizens to communicate, monitor, and participate in the processes of public service provision. As explained above, in a decentralised context, citizens' voting behaviour is often directed by their evaluation of local public services. From the perspective of local politicians and bureaucrats, such voting behaviour motivates them to build an effective channel for citizens to fully express their ideas regarding local public services (e.g. what should be provided and what objectives should be achieved by service provision), monitor daily operations (Asfaw et al. 2007), and even participate in the processes of service provision (Channa and Faguet 2012). Through the channel of communication, participation, and supervision, public service effectiveness can also be promoted in line with the will of the people.

The second mechanism is yardstick competition, which can be intensified by the voting system in those decentralised jurisdictions sharing homogenous demands and preferences with others. As explained above (Salmon 1987), citizens use the PSP in an analogical jurisdiction as a "yardstick" to evaluate the service performance achieved by their own government (Kang et al. 2012), and their voting decisions (vote with feet or vote with ballots) are directed by the evaluation results (Salmon 1987;

Fredriksen 2013). To avoid being voted out or suffering the bad effects of fiscal migration, local bureaucrats need to improve service performance towards (and beyond) the yardstick level and ensure that the formal objectives of public services can be better attained. In other words, FD and the formal (voting through ballot) and informal (voting with their feet) voting systems finally lead to an increase in public service effectiveness (Arends 2017).

However, as for efficiency, FD might jeopardise public service effectiveness if there are significant weaknesses in local institutions. For example, a lack of well-established democratic processes and regulations at the local level might create favourable conditions for interest groups and local elites to control local public services, privileging their private interest at the expense of the formal objectives for those services (Collins and Green 1994; Prud'homme 1995; Azfar et al. 2001; Bardhan and Mookherjee 2005). Moreover, the difficulties of local governments, particularly in underdeveloped regions, to attract highly skilled officials might also undermine the positive effects of FD on effectiveness (Prud’homme 1995; Arends 2017). Accordingly, similar as the FD-efficiency relationship, it is possible that FD has a positive effect on healthcare effectiveness as well.

3.1.4 Conclusions

In this section, theoretical arguments with regard to the effects of FD on public service efficiency and public service effectiveness were reviewed. Although efficiency and effectiveness represent two separate aspects of PSP, it appears that most of the links between FD and efficiency also connect FD and effectiveness together. As identified in this section, the most significant link between FD and efficiency is accountability, in which three specific mechanisms are at work: initially, FD ensures that citizens can vote with their feet or through the ballot box based on public service efficiency, thus local governments are obliged to be more accountable for the requirements of local residents, leading to greater public service efficiency. Secondly, regarding those public services with interjurisdictionally similar preferences, FD can intensify yardstick competitions among peer governments and subsequently promote accountability. Thirdly, a channel of communication, supervision, and participation can be established and consolidated with the process of FD, leading to stronger accountability and subsequently greater efficiency. These three mechanisms, as suggested by previous studies, are also able to promote public service effectiveness. With the above three mechanisms, decentralised governments are motivated to fully utilise their greater capacity in understanding local demands and managing local resources to initiate policy innovations and/or policy learning, which finally contributes to better public service efficiency and effectiveness. Additionally, theoretical studies also argue that the positive relationship between FD and efficiency/effectiveness is connected with FD's benefits to anti-corruption and reducing rent-seeking activities. However, critics of these arguments claim that FD might be negatively related to efficiency and effectiveness due to faults in the implementation of FD policies, institutional failings, and external socio-political limitations. Nevertheless, from the theoretical literature, an overall positive argument regarding the FD-efficiency and FD-effectiveness relationships can be identified, which, to some extent, provides a potential answer to the research questions of doctoral study. However, despite the relatively well-developed theoretical arguments, there is surprisingly a dearth of empirical investigation into these research questions.

Accordingly, in the next section, the limited empirical literature related to the effects of FD on public service efficiency and effectiveness will be reviewed.

3.2 The impact of FD on public service efficiency and effectiveness: evidence from empirical studies

3.2.1 Introduction

After presenting a detailed review of theories in regard to the impact of FD on public service efficiency and effectiveness, this section moves on to examine the empirical evidence relating to the above theoretical arguments. As suggested by Stegarescu (2005) and Adam et al. (2014), a key step before conducting empirical FD research is to measure the concept. Most previous studies usually measure FD from two sides: the expenditure side and the revenue side (Tanzi 1995; Jin and Zou 2002; Baskaran 2010). Therefore, to support the following discussions, an explanation of expenditure decentralisation (ED) and revenue decentralisation (RD) is provided at the beginning of this section.

ED refers to the vertical distribution of expenditure authority from the central to lower levels of government (Joanis 2014), which is the major concern of the FD theorem developed by Oates (1972). In general, the role of government includes socioeconomic stabilisation, redistributing of resources to groups and individuals, and the provision of public services and goods (Musgrave 1959). As argued by McLure (1999) and Martinez-Vazquez (1999), the vertical assignment of expenditure authority for the above responsibilities should be in accordance with the range of the region that can be benefited from the above responsibilities (McLure 1999; Martinez-Vazquez 1999). Accordingly, expenditure authority concerning socioeconomic stabilisation should be given to the central or federal government (Ahmad et al. 1997), because any issues related to social and economic stability in a region can be easily diffused to other jurisdictions and generate nationwide effects (Ahmad et al. 1997; McLure 1999; Martinez-Vazquez 1999). Second, expenditure authority related to redistribution should also be kept in the central government, in order to avoid inequity due to inter-regional gaps in fiscal capacity (Musgrave 1959; Bahl et al. 2002; Boadway and Shah 2009). Third, the provision and allocation of public services are usually viewed as a joint responsibility of both the central/federal government and subnational governments. The expenditure authority for those services with nationally homogeneous demand (e.g. national defence and foreign affairs) should be provided by the central government, while those with clear regional preferences (e.g. local culture facilities) should be undertaken by local governments (Ahmad et al. 1997; Boadway and Shah 2009). Apart from nationwide public services and "pure" local services, there are some "quasi-public" services with private nature, which cannot be fully provided in the private market, such as education, healthcare, and social services (Brueckner 1977; Reiter and Weichenrieder 1999; Iregui 2005). In practice, expenditure authority regarding such "quasi-public" services is usually shared between the central and local governments. This is because although the demands have, more or less, local preferences (Iregui 2005, Boadway and Shah 2009), and an overly high demand might significantly increase the congestion cost[7], such services

[7] Excessive demands for quasi-public services under a fixed supply level will lead to stronger competition in the user group – because those services are not "non-rivalrous", resulting in higher congestion costs (Brueckner 1977; Reiter and Weichenrieder 1999). Congestion costs could be reduced by decentralising the responsibility of service provision to local jurisdictions, as long as the service is provided within the jurisdictional border and externality can be controlled.

are usually associated with redistribution objectives (Reiter and Weichenrieder 1999; Brueckner 1977). Thus, it is required that the central government should play a certain role in service provision (e.g. setting minimum standards or providing transfer grants).

RD between different levels of government is another key issue in FD research and practice (McLure 1999). As claimed by Oates (1972), an effective system for ED should be supported by a well-established system for RD. Because tax is the major source of local government revenue in most countries, academic and practical discussions of RD usually focus on the vertical assignment of the taxing power, in which four main principles have been considered.

First, those taxes which are related to national sovereignty, suitable for economic stabilisation and redistribution purposes, should be mainly controlled by the central government to support its role in socioeconomic stabilisation and redistribution (Boadway and Shah 2009). Therefore, customs, personal income tax, and corporation income tax should be considered as tax sources for the central/federal government.

Second, because public money should be spent by the government which is closest to the beneficiaries, the decentralisation of revenue-raising power also needs to satisfy the “benefit principle”: expenditure for a public service should be paid for by the beneficiaries of such service (Oates 1972; Norregaard 1997). Therefore, those taxes (and government charges) that are linked with local public services should be controlled by the local government, so that the tax welfare can be “internalised” to local taxpayers (McLure 1999; Martinez-Vazquez 1999). For example, Norregaard

(1997) suggest that user fees for local facilities (e.g. parking charges), which are not tax payments but follow the benefit principle, should be levied and used by the local government.

Third, those taxes with highly mobile tax bases should be controlled by the central government, otherwise, the distribution of tax bases will be distorted, and total tax revenues might be reduced. That is to say, taxpayers tend to immigrate to jurisdictions with lower tax rates (Oates 1999), and local governments might not be enthusiastic to collect such taxes for the purpose of tax competition (Wilson 1986; Boadway and Tremblay 2012). In other words, those taxes with relatively fixed tax bases can be controlled by the local governments. Therefore, property tax is usually considered as an important revenue source for the local government (Norregaard 1997; Oates 2005). Although income taxes should mainly be controlled by the central/federal government for redistribution purposes, individuals/households and enterprises usually lack full mobility because of immigration costs and legal limitations (such as the "hukou" system in China). For this reason, a certain proportion of revenues from personal income taxes can be shared by subnational governments (McLure 1999; Boadway and Shah 2009).

Finally, the central government should also control taxes with a significant "tax exporting" effect. Tax exporting refers to the situation where people pay taxes for other jurisdictions but the benefits from such taxes cannot be accessed by them (Norregaard 1997). To make sure that all taxpayers are able to benefit from their payments, those taxes with strong "tax exporting" effects, such as excise taxes and sale taxes levied at the moment of manufacture, should be controlled by the central government (Morgan et al. 1996; Norregaard 1997)

The principles of ED and RD are summarised in the above paragraphs. However, in empirical studies, researchers are often more concerned with the best way to measure ED and RD. Although it is a topic always under debate (Dziobek et al. 2011), researchers have recognised that there is no perfect indicator that can be used to fully measure ED and RD in all backgrounds (Gu 2012). As suggested by Martinez-Vazquez and Mcnab (2003), different ED and RD indicators are usually considered as complementary rather than contradictory because a more comprehensive reflection of FD can be achieved by choosing multiple indicators. Nevertheless, only a few indicators based on expenditure and revenue data are widely adopted by multidisciplinary researchers in empirical studies (Gu 2012).

Regarding ED measurements, the share of a lower-level government's expenditure on one or more public services relative to the expenditure by the higher-level government is the most common indicator (OCED 2003; Afonso and Hauptmeier 2009). Another widely adopted ED measurement is sometimes regarded as an indicator of expenditure autonomy (Psycharis et al. 2016), which is the proportion of a local government's expenditure that is financed by its own taxes (Rodden 2004). Similarly, the share of local revenue relative to the revenue of the higher-level government and the proportion of local revenues that are controlled by the local government (excluding revenues that are determined by the higher-level government such as transfer grants with special purposes) are also adopted to measure RD (Oates 1972; Grossman 1989;

Jin and Zou 2002). In the following two sub-sections (3.3.2 and 3.3.3), empirical papers on the impact of ED and RD on public service efficiency and effectiveness are reviewed.

3.2.2 FD and public service efficiency

The concept of public service efficiency, as suggested by Andrews and Entwistle (2013), can be classified into four dimensions: productive efficiency, allocative efficiency, distributive efficiency, and dynamic efficiency. As described in Section 3.1, most theoretical discussions revolve around FD's contributions to productive efficiency and allocative efficiency. While for most empirical studies, due to the difficulties of measuring allocative efficiency with quantitative indices (e.g. market prices) which reflect the demand-supply relationship, it is mainly the relationship between FD and productive efficiency that is analysed, as for this PhD study. Empirical findings about the impact of ED and RD on the productive efficiency of public services are separately reviewed in the following paragraphs.

3.2.2.1 Expenditure decentralisation and public service efficiency

Although the impact of FD on public service efficiency is at the centre of theoretical FD research, there are only a few published papers which have empirically investigated the relationship between FD and productive efficiency. Regarding the impact of ED, the research of Barankay and Lockwood (2007) is "one of the first" to empirically look into this issue. Panel data for 26 Swiss Cantons (sub-national jurisdictions) from 1982 to 2000 and a popular indicator of ED (i.e., the ratio of local expenditure to central/federal government expenditure) were adopted. Empirical findings statistically confirmed that when expenditure authority was decentralised, a better education service outcome (measured as the passing rate for being admitted into a university) can be achieved under a given level of educational input. In other words, ED enhanced productive efficiency. Different from other studies, this paper did not directly measure education efficiency by techniques such as the Data Envelopment Analysis (DEA) (e.g. Arends 2017) or the Stochastic Frontier Analysis (e.g. Sow and Razafimahefa 2015). As shown in the following equation, both FD and resource inputs (*INPUT*) for education services are introduced as independent variables, while the educational outcome is the dependent variable.

$$OUTCOME = a_c + \beta_t + \gamma FD_{ct} + \delta INPUT_{ct} + \mu_{ct}$$

As the positive relationship between FD and the educational outcome was statistically confirmed (the coefficient of FD "γ" is larger than zero), the study claimed that at a given level of input (which is controlled in the model), a one-unit growth of FD will lead to an increase of outcome by γ units. In other words, FD increased the ratio of outcome to input – greater productive efficiency. However, this estimation approach failed to identify a direct correlational relationship between FD and better educational service efficiency. Thus, this finding may suffer from the lack of internal validity.

Two recent studies by Widmer and Zweifel (2012) and Alfada (2019) also identified a positive relationship between ED and the productive efficiency of public services. Based in the context of Switzerland, a country with a unique federal system, Widmer and Zweifel (2012) investigated the impact of FD on the input-output (productive) efficiency of six public services: general administration, public safety, education, healthcare, transportation, and public economy. Different from Barankay and Lockwood (2007), this study employed DEA, a non-parametric technique to measure the productive efficiency of one decision-making unit (DMU) by comparing its performance (efficiency) with the performance of the "benchmark" DMU (Cook et al. 2014). Using data from 26 Swiss Cantons over the period of 2000-2004, this paper confirmed that ED in Switzerland is positively related to the integrated efficiency score of the six public services. However, the separate relationship between ED and the productive efficiency of any selected public services was not addressed in this study. The research of Alfada (2019) was concerned with the impact of ED on the productive efficiency of four basic public services (infrastructure, education, health and social protection) in Indonesia. This study adopted a DEA-based Malmquist Productivity Index (MPI) to measure the efficiency of local public services in each year and efficiency changes over time (Raphael 2013). Moreover, this study decomposed the changes in productivity into changes in efficiency and changes in technologies, improving the accuracy of efficiency measurement (Camanho and Dyson 2006). Using data from 26 Indonesian provinces from 2004 to 2015, statistically significant positive relationships between ED and the input-output efficiency of the four selected services were identified.

Another empirical research which confirms the positive relationship between ED and public service efficiency was conducted by Alonso and Andrews (2019). This study was concerned with the overall efficiency level of six services (children and young people, adult social care, environment, housing, libraries and leisure, and benefits) provided by English local governments. As the one which reflects the up-to-date findings with regard to this academic topic, there are two distinguishing features in this paper. The first one is that by using publicly accessible data (from the Comprehensive Performance Assessments performed by the UK's Audit Commission), a value-for-money ratio which reflects both the scale and quality of output was constructed to measure productive efficiency. Secondly, a unique ED indicator, which is the proportion of a local government's expenditure that is funded by the locally collected property tax, was adopted. As suggested by the authors, this indicator is also known as 'expenditure autonomy' (Psycharis et al. 2016). Using data from 148 English local governments from 2002 to 2008, a positive relationship between ED and public service efficiency was identified. This relationship still existed when efficiency was measured by the DEA approach, showing the robustness of this finding.

Following the research of Widmer and Zweifel (2012), Arends (2017) conducted a study on the relationship between ED and the productive efficiency of healthcare services in 32 OECD countries (excluding Mexico and Turkey) for the 1995-2013 period. Also drawing on DEA to measure efficiency and the ratio of local healthcare expenditure to central/federal government healthcare expenditure to measure ED, the fixed effects models identified a negative relationship between ED and the productive efficiency of healthcare services – opposite to the findings of Widmer and Zweifel (2012). Nevertheless, these contradictory results can be interpreted from two perspectives. First, the study of Widmer and Zweifel (2012) focuses on the overall degree of efficiency of six public services, while Arends (2017) solely looks into the efficiency of the healthcare sector. Second, under two very different socioeconomic settings – the single country of Switzerland and 32 OECD countries, it is perhaps not so surprising that contradictory results arise.

A negative relationship between ED and the efficiency of public environmental services was also identified in the paper of Zhang et al. (2019). To measure productive efficiency, the rate of wastewater treatment, the rate of industrial pollution removed, and the rate of domestic waste that was collected and recycled were aggregated to calculate an output indicator, and government spending on such services was the input indicator. Using panel data for 30 provincial jurisdictions in mainland China from 2007 to 2016, this study found that ED was negatively related to the productive efficiency of environmental services. The authors claim this is because the stronger incentive to promote economic development with the support of greater fiscal authority diminishes local governments' willingness to provide environmental services in a more efficient way. However, this finding has been challenged by subsequent studies because the central and local governments, as well as the general public of China, nowadays, are paying greater attention to environmental protection and consider it as a key indicator for local cadres' performance evaluation (Tu et al. 2019; Wu et al. 2020; Liu et al. 2020a). Also, the ED indicator adopted in this study (the ratio of a province's total per capita fiscal expenditure to the central government's total per capita fiscal expenditure) is problematic, as it is not a specific one measuring environmental service decentralisation. Moreover, by no means the central government's per capita fiscal expenditure to all Chinese provinces would by no means be equal.

Apart from direct positive or negative relationships between ED and the efficiency of various public services identified in some studies, a non-linear relationship was identified by the cross-national research of Sow and Razafimahefa (2015). In their study, efficiency is measured by the SFA approach, which is similar to DEA but is a parametric frontier estimation approach (DEA is a nonparametric technique) (Balaguer-Coll et al. 2007). The efficiency analysis of this study, as the authors pointed out, is concerned with the improvement of education and healthcare service outcomes (measured by the mortality rate of infants and the school enrolment rate) under a given input. 64 developed, developing, and emerging countries were included in this analysis. Using panel data from 1990 to 2012, this research identified a U-shape relationship between ED and healthcare service efficiency, with the tipping point equal to 35.7% (100% indicates full decentralisation). It means that ED and efficiency had a negative relationship when the level of ED was smaller than 35.7%. After that, a positive relationship existed. To explain this result, further analysis suggested that most ED values lower than 35.7% were for developing countries with worse performance in democracy and anti-corruption, which turned the ED-efficiency relationship negative. After exceeding the tipping point of ED (35.7%), most observations were for developed countries with better performance in democracy and lower corruption levels, thus, the ED-efficiency relationship became positive. The above findings are in line with the theoretical argument that an unfavourable social-political environment can significantly undermine local government capacity, which jeopardises the positive relationship between FD and efficiency (Tanzi 1995). However, the ED indicator adopted in this study only measured the expenditure decentralisation between the national and sub-national governments for all services, which may not reflect the real level of ED for healthcare and education.

3.2.2.2 Revenue decentralisation and public service efficiency

There are seven empirical papers that have discussed the impact of RD on public service efficiency. Initial findings were provided by the research of Porcelli (2009) which was concerned with the effects of the 1998 tax decentralisation on healthcare service efficiency in Italy. Using DEA and regional-level data between 1991 and 2005, a positive relationship between RD and the productive efficiency of healthcare services was identified. However, only a dummy variable was adopted to measure RD in this empirical paper. Before 1998 – the year of tax decentralisation reform, the level of RD in all regions is "0", while the level of RD in all areas was changed to "1" from 1998 to 2005. Using dummy variables is a simple and useful way to measure variables that can be difficult to quantify, such as policy and strategic changes (Suits 1957; Yip and Tsang 2007). However, for this study, it is far from perfect to only rely on a dummy variable to measure FD (RD). This is because the dummy variable (0 or 1) in this study only reflects whether the policy (i.e., tax decentralisation) exists or not, but cannot evaluate the degree to which this policy is fully implemented in different jurisdictions as well as in different periods. Therefore, this study only looked into the impact of introducing an RD policy on healthcare service efficiency, but failed to further identify the relationship between different levels of RD and different degrees of efficiency, showing the necessity of conducting a more detailed study.

The above limitation was addressed in the research of Boetti et al. (2012) by adopting another RD measure, the ratio of a local government's own taxes to its total revenues, to investigate the causal relationship between RD and public service efficiency in Italian municipalities. Five basic public services, including general administration, waste collection and disposal, education, elderly care, and transportation were concerned. The efficiency scores of the above services were measured by the SFA and the DEA approach. Drawing on data from 262 municipalities in 2005, this study confirmed that RD is positively related to the efficiency of the five selected public services. However, the validity of the above finding might be limited, because only cross-sectional data were analysed.

Apart from the above two preliminary studies, a positive relationship between RD and public service efficiency was also identified by Sow and Razafimahefa (2015). Although the main decentralisation indicator of their multinational research is ED, they also controlled RD in their regression models and statistically confirmed a positive relationship between RD and the efficiency of education and healthcare in 64 selected countries. The limitation in Porcelli's (2009) research was also avoided in this paper by using the share of a local government's revenue to the per capita revenue of the central government. Another empirical study which controlled both RD and ED was conducted by Arends (2017). Using the same RD measurement, regression results from random effects models suggested that RD was positively related to the efficiency of healthcare services in 32 OECD countries (excluding Mexico and Turkey).

By contrast with most prior research, the recent study of Alfada (2019), which adopted the central-provincial revenue ratio to measure RD, revealed a negative relationship between RD and public service efficiency in the context of Indonesia. In the integrated model which included control variables for measuring political stability and ethnolinguistic fractionalisation, RD was found to be negatively related to the efficiency of the education and social care services in Indonesian provinces during 2004-2015. However, the impact of RD on the efficiency of infrastructure and healthcare services cannot be statistically confirmed.

In the empirical study of Adam et al. (2014), the share of local revenue to total or central-government revenue, which is a popular RD measurement, was criticised because this ratio cannot fully capture the revenue-raising power of local governments. Instead, this study adopted the ratio of revenues from those taxes with locally determined tax rates/bases to total tax revenues to measure RD in 21 OECD countries from 1970-2000. Following the DEA and SFA techniques to measure the efficiency of education and healthcare services, this empirical study suggested that an inverted U-shape relationship existed between RD and the efficiency of both services. The threshold value of RD for education is 34.46 (the maximum RD value is 100) while the threshold for healthcare is 47.1. This result is different from the research of Sow and Razafimahefa (2015) in which a U-shaped relationship between FD (measured by ED) and healthcare efficiency was identified. To explain the inverted U-shaped findings, Adam et al. (2014) claim that when the level of RD is growing from a relatively low level, the benefits of the increased electoral accountability are larger than the losses from diseconomies of scale. After the level of RD exceeds the threshold value, the losses from diseconomies of scale would prevail over the benefits of accountability, thus reducing the level of public service efficiency.

The above paragraphs provided a detailed review of the previous empirical studies focusing on the relationship between FD and public service efficiency in different contexts. A summative table (Table 3.1) is listed below, in which the basic information of the above papers is shown. Although the results from these studies are sometimes contradictory, overall, most of them indicate a positive relationship between FD and efficiency (4 for ED, 4 for RD), rather than a negative (2 for ED, 1 for RD) or a nonlinear relationship (1 each for ED and RD). Nevertheless, it can be seen that the number of studies directly addressing the FD-efficiency relationship is still limited, and only four of the above studies specifically focused on healthcare (Arends 2017; Sow and Razafimahefa 2015; Porcelli 2009; Adam et al. 2014). Furthermore, some of these studies suffer from limitations regarding the FD indicators (Sow and Razafimahefa 2015; Zhang et al. 2019) and the estimation approach (Barankay and Lockwood 2007). These problems demonstrate the timely importance of conducting this doctoral research. In the next section (3.2.3), empirical findings in regard to the impact of FD on public service effectiveness will be reviewed.

Table 3.1 Studies of the relationship between FD and public service efficiency

Study	Research context/period	FD indicators	Selected services	Methods for measuring efficiency	Findings
Barankay and Lockwood (2007)	26 Swiss cantons (1982-2000)	ED: share of total expenditure	Education	Input/outcome ratio	ED: positive
Widmer and Zweifel (2012)	26 Swiss cantons (1982-2000)	ED: share of total expenditure	Education, healthcare, general administration, public safety, transportation, and public economy	DEA (public service input and service output)	ED: positive
Alfada (2019)	26 of 33 provinces in Indonesia (2004-2015)	ED: share of total expenditure RD: share of revenue	Education, healthcare, infrastructure, and social protection.	DEA-based MPA	ED: positive RD: negative for education and social care
Alonso and Andrews (2019).	148 English local governments (2002-2008)	Expenditure autonomy: share of expenditure supported by local taxes	Children and young people, adult social care, environment, housing, libraries and leisure, benefits	Value-for-money ratio, DEA	ED: positive
Arends (2017)	32 OECD countries, 1995-2013	ED: share of healthcare expenditure; share of total expenditure RD: share of revenue	Healthcare	DEA	ED: negative RD: positive
Zhang (2019)	30 provinces in China (2007-2016)	ED: share of total expenditure	Environmental services	DEA	ED: negative
Sow and Razafimahefa (2015)	64 countries (1990-2012)	ED: share of total expenditure RD: share of revenue	Healthcare and education	SFA	ED: a U-shape relationship between ED and healthcare efficiency RD: positive for both two services (no non-linear relationship exists)

Porcelli (2009)	15 Italian regions (1991-2005)	A dummy variable separating the years before/after the 1997 tax decentralisation reform	Healthcare	DEA	RD: positive
Boetti et al. (2012)	262 Italian municipalities 2005	RD: share of self-collected revenue to total revenue	General administration, waste collection and disposal, education, elderly care, and transportation	DEA, SFA	RD: positive
Adam et al. (2014)	21 OECD countries (1970-2000)	RD: share of revenue from taxes strictly determined by sub-national governments to total tax revenue	Healthcare and education	DEA, SFA	RD: an inverted U-shape relationship exists between RD and the efficiency of both two services

3.2.3 FD and public service effectiveness

As mentioned above, public service effectiveness is measured by the extent to which the formal objectives of public service provision are achieved. Most empirical studies, including the papers reviewed in this thesis, consider outcome as a synonym for effectiveness (Mandl et al. 2008). Almost all empirical studies discussing the FD-effectiveness relationship are concerned with healthcare (17 of 20) and education services (5 of 20). Accordingly, in this section, empirical studies of the effectiveness of healthcare and education services are reviewed.

3.2.3.1 FD and the effectiveness of healthcare services

The empirical discussion of this topic usually starts from the selection of effectiveness indicators for healthcare services, of which, globally, the infant mortality rate (IMR) is the most popular one (e.g. Young 2005; Cavalieri and Ferrante 2016). This is because IMR reflects both children's health and pregnant women's health – improving their health status is therefore undoubtedly a formal objective of healthcare services (Kang et al. 2012). Moreover, IMR is also considered an exhaustive indicator which also reflects the overall level of healthcare service effectiveness (Young 2005; Kang et al. 2012). Additionally, Jiménez-Rubio and Smith (2005) claim that IMR is a superior indicator of effectiveness to life expectancy – another popular effectiveness indicator for healthcare services, as the statistically forecasted values of LE are easier to manipulate than IMR, and IMR tends to be more sensitive to policies such as FD

(Porcelli 2014). Regarding the FD indicators, in line with empirical studies on FD and efficiency, both ED indicators and RD indicators are adopted. However, ED indicators appear to be more popular in FD-effectiveness studies. In the following paragraphs, empirical contributions about the impact of ED and RD on the effectiveness of healthcare services are reported.

Initial findings with regard to the impact of ED on the effectiveness of healthcare services are provided by Robalino et al. (2001). ED was represented by the ratio of sub-national expenditures to central-government expenditures. Using panel data from low-income, middle-income, and high-income countries for 1970-1995, this study confirmed that ED can significantly reduce IMR, and that this relationship was particularly strong in low-income and high-income countries. Additionally, moderating effects of citizens' political rights and ethnolinguistic fractionalisation were also identified: the contribution of ED to IMR reduction was greater in countries with more political rights or with a lower level of ethnolinguistic fractionalisation. Additionally, the authors claimed that the positive effect of ED was robust enough to offset the negative impact of corruption on the IMR. Using the same ED and effectiveness indicators, Brock et al. (2015) discussed the same issue in the Chinese context. Collecting data at the national and provincial level in 1980, 1981, 1989, 1990, 2000, and 2003, this study suggested that decentralising spending authority from the central government to provinces failed to decrease IMR, which is contradictory to most empirical findings in this field. However, this study employed old data and due to data accessibility issues, only six years of the data from 1980 to 2003 were collected. Moreover, both two studies of Robalino et al. (2001) and Brock et al. (2015) might suffer from validity issues as they only employed one expenditure-side FD indicator, i.e., the percentage of fiscal expenditures spent by the subnational governments. This limitation was partially addressed in another Chinese-based study by Uchimura and Jutting (2009) by adopting two expenditure-side FD indicators: 1) total expenditure ratio i.e., the percentage of all counties' fiscal expenditures to the fiscal expenditure of the province; 2) expenditure autonomy, i.e., the ratio of all counties' expenditures to all counties' own revenue. This study found that during the period of 1995 to 2001, both indicators are related to a lower IMR in Chinese provinces. However, indicators directly measuring healthcare decentralisation are not employed, not to mention that the data employed in this study are from over 20 years ago, showing the necessity of conducting new research with updated data and FD indicators.

Similarly, in an Indian-based study conducted by Asfaw et al. (2007), the share of local expenditures covered by state transfers, per capita local expenditure, and the share of a local government's own revenue to total revenue were adopted to establish an integrated FD measure for local jurisdictions. Using panel data from 1990 to 1997, a negative relationship between FD and IMR in rural areas of India was confirmed. Moreover, the role of political decentralisation as a moderator was also identified. Specifically, the marginal growth of FD can reduce the IMR by 17.16% in jurisdictions with political decentralisation indices higher than the average value, while the reduction of IMR was only 8.64% in jurisdictions with political decentralisation indices lower than the average level. This finding corroborates the theoretical arguments that the positive effect of FD on PSP is reliant on the support of a well-established socio-political system (Prud'homme 1995). In line with the study of Brock et al. (2015), Jin and Sun (2011) also identified that ED, which is measured by the total per capita expenditures of a province divided by the total per capita expenditures of the central government, can significantly increase IMR in China. Moreover, a dummy variable was adopted to identify the impact of the revenue centralisation reform in 1994 (i.e., the 1994 Tax Sharing System reform), and an unfavourable effect of revenue centralisation on IMR was confirmed. However, as suggested by Jiménez-Rubio (2011a), a dummy variable might not be the best choice to measure decentralisation because it fails to reflect the “real” degree of decentralisation in jurisdictions where the same decentralisation policy is implemented to different extents.

To better reflect the level of ED on a specific service, several papers adopted the ratio of local healthcare expenditure to central/federal government healthcare expenditure. For example, two empirical studies conducted by Cantarero and Pascual (2008) and Jiménez-Rubio (2011b) respectively looked into the impact of decentralising healthcare expenditures on the effectiveness of healthcare services in Spain and Canada. Using panel data from 1992 to 2003, Cantarero and Pascual (2008) confirmed that decentralising healthcare expenditures can significantly reduce the IMR and increase life expectancy in Spanish regions. Such a relationship between ED and a lower IMR was also identified in 10 provinces of Canada over the period of 1979 to 1995 (Jiménez-Rubio 2011b). Another two cross-national studies using the same ED indicator were conducted by Prieto and Saez (2006) and Kang et al. (2012). Focusing on the background of 15 EU countries over the period from 1990 to 2003, Prieto and Saez (2006) found that ED was negatively related to the IMR and was positively related to life expectancy. However, based on the contexts of 22 OECD countries, Kang et al. (2012) found that ED had a nonlinear effect on the effectiveness of healthcare services. With the increase in ED, the level of IMR was initially reduced, but it subsequently turned into an upward trend when the level of ED reached a certain level. Therefore, the authors argue that FD is not a "panacea" for promoting healthcare service effectiveness. The explanation for this U-shaped relationship is quite similar to that for the inverted U-shape between ED and public service efficiency: with the increase of ED, the costs to internalise the externalities become larger (Breuss and Eller 2004). When the losses outweigh FD's benefits in matching local preferences and utilising local resources, the reduction of productive and allocative efficiency finally leads to a decrease in healthcare service effectiveness.

In addition to the expenditure ratio, the proportion of expenditures afforded by own revenues was adopted by Soto et al. (2012) to measure the degree of ED in 1080 municipalities in Colombia over a 10-year period (1998-2007). A negative relationship between ED and IMR was confirmed, with this causal relationship becoming stronger in wealthier regions. This phenomenon can be attributed to theoretical arguments regarding the impact of FD on public service effectiveness: the better socio-political institutions, the lower level of corruption, and citizen participation might significantly amplify the positive effects of FD on PSP (Enikolopov and Zhuravskaya 2003; Boadway and Shah 2009). Also in a Latin American context, insignificant results were identified when using the same ED indicator to investigate this issue in Brazilian municipalities, showing the potentially limited impact of ED on the effectiveness of healthcare services (Rocha et al. 2016). Using the same ED indicator, Cavalieri and Ferrante (2016) also found that a higher proportion of expenditures supported by own revenues was associated with a lower IMR in 20 Italian regions over the period 1996-2012. Moreover, this study adopted the ratio of self-controlled tax revenues to total tax revenues to measure RD, and a negative relationship between RD and IMR is identified. Supported by two similar RD indicators: the ratio of locally controlled fiscal resources to total resources collected and the ratio of provincial taxes to locally controlled fiscal resources, a negative relationship between RD and IMR was also confirmed by Habibi et al. (2001) for 23 provinces in Argentina during the years 1970 to 1994.

Apart from the RD indicator measuring the percentage of self-controlled revenue to total revenue, the ratio of local revenue (per capita) to central/federal revenue (per capita) is also adopted in two empirical studies: Jiménez-Rubio (2011a) measured RD with two indicators 1) the ratio of revenue from sub-nationally controlled taxes (i.e., those taxes for which the local government determines the tax rate and/or tax base) to total tax revenue of the country and 2) the ratio of total sub-national tax revenue to total tax revenue of the country. This study found that the higher level of RD (measured by the first indicator, i.e., the ratio of revenue from sub-nationally controlled taxes) was associated with lower IMR rates in 20 OECD countries between 1970 to 2001. Such a relationship was also identified by the study of Dada (2015) which employed the local-central revenue ratio and focused on both healthcare services and educations service in Nigeria between 2002 to 2010 (Dada 2015).

3.2.3.2 FD and the effectiveness of education services

Only five empirical studies have discussed the impact of FD on the effectiveness of educational services. To measure effectiveness, two types of indicators were adopted: students' examination results and the ratio of graduates enrolling in higher-level and non-compulsory education institutions that are owned by the public. Using the Programme of International Student Assessment (PISA) results of OECD countries in 1995, 2002, 2005, and 2008, Fredriksen (2013) found that both ED (measured by the decentralisation of total expenditures) and RD were associated with a better PISA result. Drawing on the local-government education expenditures to central-government education expenditures to measure ED, Barankay and Lockwood (2007) and Freinkman and Plekhanov (2009) investigated the impact of ED on the effectiveness of education services in Switzerland and Russia. While a higher ED level was associated with a higher percentage of students meeting the university entry requirements in Switzerland (Barankay and Lockwood 2007), the study of Freinkman and Plekhanov (2009) identified no relationship between ED and students' average marks in mathematics and language in Russia. However, Freinkman and Plekhanov (2009) found that the percentage of expenditure covered by self-controlled revenues was significantly and positively related to students' performance in examinations. A similar finding was also found in the Argentine-based research (Habibi et al. 2011) mentioned above, in which the effectiveness of educational services was measured by the ratio of young people enrolling into higher-level educational institutions, and RD was the ratio of self-controlled revenues to total revenues. Additionally, Dada (2015) found that under the background of a developing country (Nigeria), not only the effectiveness of healthcare services but also that of education (measured by the literacy rate) was promoted through RD from 2002 to 2010.

To conclude, the above discussions summarised the empirical papers which focused on the impact of FD on the effectiveness of healthcare and education services, in which most of them confirmed the existence of a positive relationship between FD and effectiveness. A summative table (Table 3.2) is given below, with a brief account of all empirical papers mentioned above.

Table 3.2 Studies of the relationship between FD and public service effectiveness

Study	Research context/period	FD indicators	Selected services	Effectiveness indicators	FINDINGS
Robalino, Picazo, and Voetberg (2001)	A panel of low- and high-income countries (1970-1995)	ED: share of total expenditure	Healthcare	IMR	ED: reduces IMR Moderators: citizens' political rights (+) and ethnolinguistic fractionalisation (-)
Brock et al. (2015)	31 Chinese provinces (1980-2003)	ED: share of total expenditure	Healthcare	IMR	ED: improves IMR
Uchimura and Jutting (2009)	26 Chinese provinces (1995–2001)	ED: share of total expenditure; ratio of self-controlled revenue to total expenditure	Healthcare	IMR	ED: reduces IMR
Asfaw et al. (2007)	Rural areas of India (1990 – 1997)	ED: share of total expenditure; ratio of self-controlled revenue to total expenditure	Healthcare	IMR	ED: reduces IMR Moderator: political decentralisation (+)
Jin and Sun (2011)	31 Chinese provinces (1980-2003)	ED: share of total expenditure RD: a dummy variable reflecting the 1994 TSS reform	Healthcare	IMR	ED: improves IMR RD: reduces IMR
Cantarero and Pascual (2008)	Spanish regions (1992 - 2003)	ED: share of healthcare expenditure	Healthcare	IMR; life expectancy	ED: reduces IMR, improves life expectancy
Jiménez-Rubio (2011b)	10 provinces of Canada (1979 – 1995)	ED: share of healthcare expenditure	Healthcare	IMR	ED: reduces IMR
Prieto and Saez (2006)	15 EU countries (1990-2003)	ED: share of healthcare expenditure	Healthcare	IMR and life expectancy	ED: reduces IMR, improves life expectancy
Kang et al. (2012)	22 OECD countries (1995-2005)	ED: share of healthcare expenditure	Healthcare	IMR	Threshold effects: a U-shape relationship between ED and IMR
Soto et al. (2012)	1080 municipalities in Colombia (1998-2007)	ED: ratio of self-controlled revenue to total healthcare expenditure	Healthcare	IMR	ED: reduces IMR
Rocha et al. (2016)	Brazilian municipalities (2000 to 2007)	ED: ratio of self-controlled revenue to total healthcare expenditure	Healthcare	IMR	ED: insignificant
Cavalieri and Ferrante (2016)	20 Italian regions (1996-2012)	ED: ratio of self-controlled revenue to total healthcare expenditure	Healthcare	IMR	ED: reduces IMR RD: reduces IMR

		RD: share self-controlled tax revenue to total revenue			
Habibi et al. (2011)	Argentina (1970-1994)	RD: share of self-controlled revenue to total revenue	Healthcare	IMR	RD: reduces IMR
Jiménez-Rubio (2011a)	20 OECD countries (1970-2001)	RD1: share of revenue RD2: share sub-nationally self-controlled revenue to total revenue of the country	Healthcare	IMR	RD1: reduces IMR RD2: insignificant
Dada (2015)	36 Nigerian states and federal capital territory (2002 to 2010)	RD: share of revenue	Healthcare	IMR	RD: reduces IMR
Fredriksen (2013)	OECD countries in 1995, 2002, 2005 and 2008	ED: share of total expenditure RD: share of revenue	Education	Students' performance in PISA	ED/RD: improves education service effectiveness
Barankay and Lockwood (2007)	26 Swiss cantons (1982–2000)	ED: share of education expenditure	Education	maturité rate (passing rate)	ED: improves education service effectiveness
Freinkman and Plekhanov (2009)	73 regions in Russia (2004 and 2005)	ED: share of education expenditure; share of self-controlled revenue to total educational expenditure	Education	Average mark of mathematics and language	Educational expenditure ratio: insignificant Share of self-controlled revenue: improves education effectiveness
Habibi et al. (2011)	Argentina (1970-1994)	RD: share of self-controlled revenue to total revenue	Education	Enrolment rate of secondary schools	RD: improves education service effectiveness
Dada (2015)	36 states and federal capital territory (2002 to 2010)	RD: share of revenue	Education	Literacy rate	RD: improves literacy rate

3.3 Summary

This chapter provided a comprehensive review of the theoretical and empirical arguments with regard to the impact of FD on public service efficiency and effectiveness. Theoretical explanations were given at the beginning. Both generations of fiscal federalism (FF) – the theoretical underpinnings of FD – argue that FD mobilises local governments to promote service performance. FF 1.0 suggests that local governments are altruistic and thus they would make more efforts to promote service performance once more power and responsibilities are decentralised to them. While FF 2.0 argues that politicians and bureaucrats are egoistic and aim to maximise their private benefits by winning elections and attracting population/capital inflow. With FD, local governments play a greater role in local responsibilities, thus, local officials have greater motivation to maximise their own benefits by satisfying people's needs for better PSP. The above analysis of the two generations of FF is reliant on a key assumption that local governments, compared with the central government, have greater advantages in collecting/managing local information and resources (Oates 1999; Balaguer-Coll et al. 2010).

The review of empirical studies about FD's impact on public service efficiency and effectiveness highlights that there is still a lack of empirical research on this topic – particularly the impact of FD on healthcare efficiency. Furthermore, the existing studies suffer from limitations regarding FD indicators and/or the timelessness of the data. Regarding FD indicators, when measuring decentralisation from the expenditure side, most studies adopt the ratio measuring total expenditure allocation between the higher and lower-level governments (apart from a few exceptions such as Arends 2007). However, this indicator fails to measure the level of decentralisation for a specific service, leading to “aggregation bias” (Barankay and Lockwood 2007). Also, most of these empirical studies employ secondary data before the 2010s, thus failing to investigate the impact of FD on PSP over the recent years.

Moreover, the assumption of FF about local governments’ greater capacity for action should not be taken for granted. In fact, in underdeveloped regions, the central/federal government might have a greater capacity than their local counterparts. In this case, the benefit of FD on PSP might be significantly undermined (Prud’homme 1995). However, this issue is mostly ignored in empirical studies. Also, political features which support the existing theoretical and empirical analysis of the FD-PSP relationship such as multi-party elections and strong external oversight do not exist in China. In other words, arguments such as people’s ballot-voting behaviours motivating local bureaucrats to improve performance cannot be generalised to Chinese provinces. Recognising the above gaps, this doctoral research discusses the impact of FD on healthcare service performance in Chinese provinces, using multiple FD indicators and paying special attention to the impact of local governments’ capacity on the relationship between FD and healthcare performance. To support the Chinese-based empirical analysis and discussions, the next chapter (Chapter 4) will introduce the political and administrative backgrounds, arrangements of FD, and the evolution of the healthcare sector in China. Then, hypotheses based on the Chinese context will be developed in Chapter 5. Finally, hypotheses will be tested and discussed in Chapter 6 and Chapter 7, respectively.

CHAPTER 4 RESEARCH CONTEXT

The former two chapters provided a detailed introduction to and review of the academic and empirical literature regarding PSP, FD, and the relationship between FD and PSP. This research context chapter begins by introducing the administrative hierarchy of China and the Chinese provinces where FD and healthcare policies are formulated and implemented. Then, the second part summarises the evolution and current arrangements of Chinese-style FD. Finally, the healthcare sector in China is introduced.

4.1 Understanding intra-provincial FD in China: an introduction

In the past few decades, intergovernmental fiscal and administrative relations in China have been substantially changed (Wang and Ma 2014; Donaldson 2017a). As Chien (2010) highlighted, decentralisation is one of the answers for the internationally remarkable achievements made by Chinese policymakers in socioeconomic reforms (i.e., Reform and Opening-up or "*Gaige Kaifang*"). Although China remains a unitary political system, now it has established a highly decentralised governmental system with administrative features of federalism (Zheng 2007). Thus, this section provides a brief introduction to the Chinese administrative hierarchy, and then the evolution and current arrangements for inter-governmental fiscal and administrative relations are summarised, so as to provide a solid foundation for the empirical analysis of FD and PSP.

4.1.1 Administrative hierarchy of China

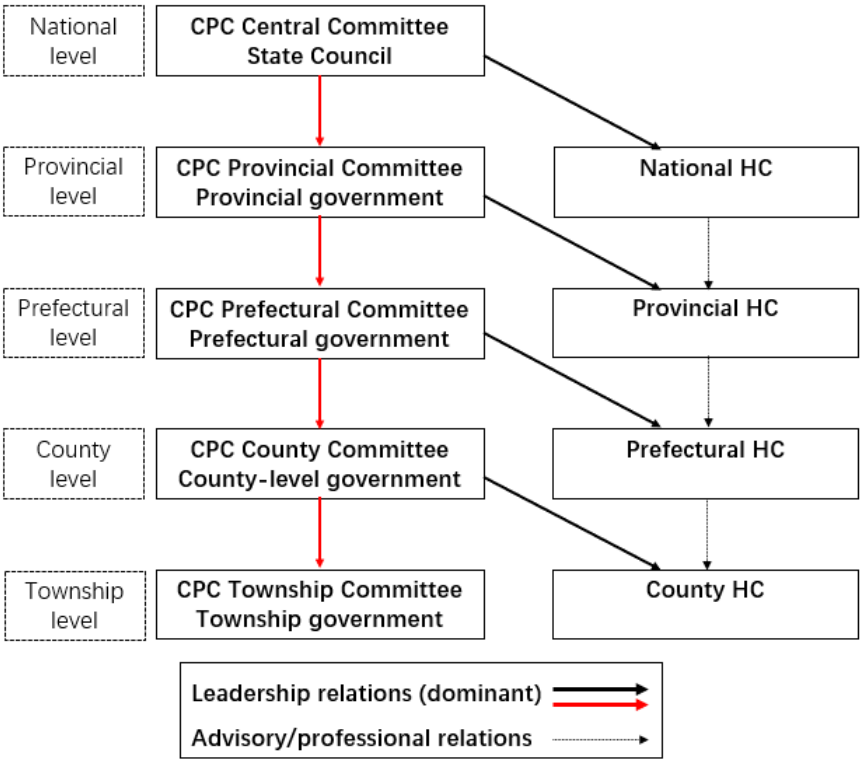
To manage a country with such a massive population and territorial size under a unitary system, nowadays governments in China are vertically divided into five tiers, as shown in Figure 4.1. In this administrative hierarchy, the CPC Central Committee and the Central government (formally named the State Council), are at the top, holding supreme power and reflecting the unitary nature of the state. For example, the Constitution of PRC (Article 89) stipulates that the establishment or abolition of jurisdictions and changes of the geographic border for all administrative jurisdictions (apart from township regions at the grassroots level) can only be approved by the central government. The central government also holds a privileged position in terms of legislation and the control of local governments: it defines the functions and duties of local governments, and it can impose nationwide laws or abolish local regulations (Zheng 2007). Within the central government, there is a set of functional agencies by whom specific policies are formulated and implemented. Usually, these functional agencies are administratively one-level lower than their corresponding party and state government (as shown in Figure 4.1). For example, the central government controls nearly 70 ministry-level agencies[8] which are located at the second level of the administrative hierarchy (Ma and Christensen 2020), including the National Health Commission (HC) which is in charge of nationwide healthcare issues[9]. Then, provincial governments are located in the second tier. Nowadays in the PRC, there are

[8] Ministries, commissions, state bureaus, state-level offices, and other ministry-level institutions

[9] Accordingly, provincial functional agencies (e.g. the provincial HC) and prefecture-level agencies (e.g. the prefecture-level HC) are located at the third and fourth tiers, respectively.

23 provinces[10], 5 provincial Autonomous Regions for ethnic minorities, and 4 provincial municipalities[11][12]. Sub-provincial governments are located below the second tier, including prefectural governments, county-level governments, and township governments respectively at the 3rd, 4th, and 5th tiers[13]. The CPC Committees which politically lead the government at the same administrative levels are located within the hierarchy as well (see Figure 4.1). Also, the provincial and sub-provincial governments control a set of functional agencies (e.g. the provincial HC and the city HC) which are administratively one level lower than them.

Figure 4.1 The accountability hierarchy of Chinese governments and CPC organisations



[10] Including Taiwan which is claimed by the PRC as its province but is not ruled by the PRC.
[11] 4 municipalities include Beijing – the capital of the PRC, Shanghai – the economic centre of China, Tianjin – the economic centre of North China, and Chongqing – the central city of West China
[12] For the convenience of analysis and descriptions, unless otherwise specified, all provincial jurisdictions discussed in this study are called "provinces".
[13] Township governments are directly controlled by the counties above them and thus could be considered as 'branches' of the county government (Donaldson 2017b).

Different from western settings with competitive elections, under the Chinese political context, the authority of local governments is derived hierarchically from the upper level (Tsui and Wang 2008; Lam 2010). In other words, an “upward accountability” relationship exists throughout the system. Within Chinese provinces, the direction of upward accountability among governments at each administrative level (e.g. the provincial government vs prefecture-level city governments in a province) appears to be clear-cut (as shown by the bold vertical red arrows in Figure 4.1). However, for most functional agencies, their activities are undertaken locally, and their service outputs are mainly accessed by local residents with strong regional features and a lower level of externality. Thus, apart from a few exceptions (e.g. the State Security Bureau), most functional departments are led by and thus upwardly accountable to their local governments, as shown by the bold upsloping black arrows in Figure 4.1. Accordingly, the superior functional agencies play an advisory but not mandatory role in policy-making (Burns and Zhou 2010; Ma 2017b). In the next two parts of the chapter, the evolution and current arrangements for inter-governmental relations in China will be explained in detail.

4.1.2 Intergovernmental administrative and fiscal relations before 1994

To clearly depict the institutional background where the effects of FD on healthcare service performance are realised, the general central-local relations as well as the relations between the provincial government and sub-provincial governments are explained in the following paragraphs. Over the past decades, intergovernmental relations in China have experienced various substantial changes. After the establishment of the PRC in 1949, there was a short period of decentralisation in the first few years. Thereafter, to suppress local rebellions and control socioeconomic activities, the central government divided China into six major districts and granted extensive power and responsibilities to district leaders. However, with growing concerns about excessive localism, from 1957, top leaders of the CPC started to centralise decision-making and administrative power within the party, especially in southern provinces such as Guangdong. With the establishment of the centrally planned socioeconomic system, centralisation was further institutionalised in the late 1950s and early 1960s (Xie and He 2010).

In the following 20 years before the early 1980s, both centralisation and decentralisation initiatives occurred, but the basic trend of such changes was still towards centralisation (Li 2010). In line with the vertical distribution of administrative power, before the early 1980s, fiscal power in China was strictly controlled by the central government as well (Dabla-Norris 2005). On the revenue side, profits from state-owned enterprises - the major source of income, and tax revenues were all controlled by the central government. Only a small proportion of government revenues came from tax payments, and personal and corporate income tax did not even exist (Shen et al. 2012). On the spending side, although local services such as healthcare, education, and social security were provided locally, under the centrally planned system, local governments had no decision-making power for the above activities but were to follow strictly the mandatory directions from the higher-level government (Shen et al. 2012; Burns and Huang 2017). The lack of autonomy strongly diminished the incentives of local governments in socioeconomic development (Wong 1991). More seriously, as Shen (2001) highlighted, the processes of decentralising and recentralising fiscal and administrative authorities were a zero-sum game, because both two processes were initiated without clear arrangements of responsibilities and benefits between stakeholders. Thus, no matter decentralisation or recentralisation, one of the two sides – the central government and local governments, experienced great losses of control. The increasing central-local tensions brought by the above processes were considered as a reason for Mao Zedong to launch the Cultural Revolution, which led to complete socioeconomic stagnation and strongly undermined the government's ability to perform its basic functions and responsibilities (Clark 2008).

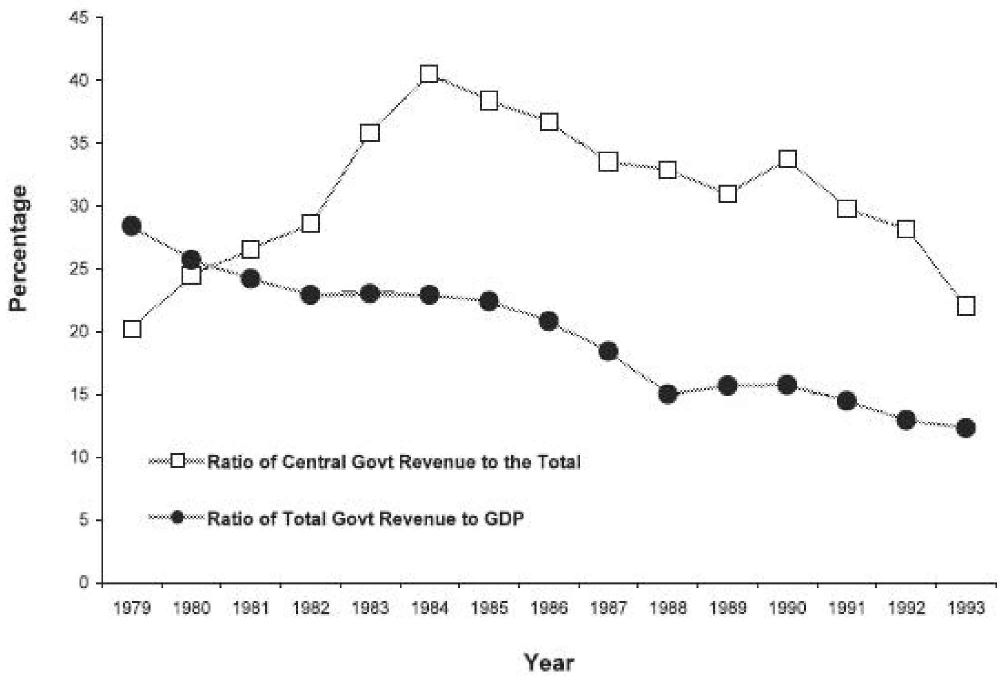
After Mao's death in 1976, his political legacies were gradually dismantled by the second generation of CPC leaders who soon initiated the Reform and Opening-up process in 1978 to eliminate the chaos of the Cultural Revolution and carry out socioeconomic reforms with a substantially changed strategy (Li 2010). To address the problems caused by the highly centralised inter-governmental relations, FD was located at the centre of this reform from the beginning (Su and Zhao 2004), and it has been considered as one of the factors contributing to China's remarkable achievements in recent decades (Shen et al. 2012). From the early 1980s, greater power was given to provincial governments to formulate local economic policies, attract investments, and manage local administrative affairs, such as appointing lower-level officials (Donaldson 2017b). Similarly, the provincial government started to share the authority with prefectural governments under their administration. From the mid-1980s, most prefectures in China were gradually reorganised into prefecture-level cities, where the local government became a separate administrative entity and was no longer a detached agency of the province[14]. As Chien (2009) suggested, decentralisation within the provinces is even more radical than central-provincial decentralisation, as sub-provincial governments are closer to the ground and have more local resources under their control. The implementation of most regional policies in a province, such as urban planning and development, public service provision, and managing natural resources and the environment, relies on sub-provincial governments (Wu et al. 2006; Chien 2009).

Along with administrative decentralisation, the power of raising revenue was largely devolved to local governments from the 1980s (Su and Zhao 2004). Only a small proportion of revenues needed to be handed over to the central government, according to the contracts set by separate central-provincial negotiations (Shen et al. 2012). The above changes effectively supported local cadres to perform their roles and

[14] The government of a prefecture-level city is an independent administrative entity, which is located at the third level of China's administrative hierarchy (see Figure 4.1). Differently, the government of a prefecture is a detached agency of the provincial government, rather than an independent administrative entity. Nowadays, only 7 of the 333 prefecture-level jurisdictions are prefectures – the rest are prefecture-level cities.

accelerated the Reform and Opening-up process in the 1980s (Wong 1991). However, the above changes gradually pushed central-local relations from one extreme to the other, which strongly reduced the ability of the central government in formulating and implementing nationwide policies, as reflected in Figure 4.2. Moreover, the central government lacked fiscal capacity and central taxes to support poorer jurisdictions through transfer payments, leading to greater regional socioeconomic disparities (Wong 1991; Lou 2008; Dabla-Norris 2005). In response to these problems, the Tax Sharing reform was launched in 1994 and re-shaped central-local relations in the following 20 years.

Figure 4.2. Ratios of central government revenue to the total fiscal revenue and GDP from 1979 to 1993



Source: China Statistical Yearbook 2004

4.1.3 The 1994 tax-sharing reform and current arrangements of fiscal decentralisation: changes and effects

The 1994 tax-sharing reform is considered as one of the most influential rearrangements for intergovernmental fiscal and administrative relations in the history of the PRC (Shen et al. 2012), in which three central issues were addressed: i) redesigning tax assignments to centralise government revenues; ii) clarifying spending responsibilities between central and subnational governments; and iii) rebuilding the system for transfer payments.

For revenue assignment, as shown in Table 4.1, taxes were categorised into central taxes, local taxes, and central-local sharing taxes, important taxes such as tariffs, value-added taxes, and income taxes, were fully or largely controlled by the central government (Su and Zhao 2004; Wang 2010; Niu 2013). Tax bases and rates were formally set and reunified by national laws and regulations, rather than by separate negotiations between the central governments and provinces. Within a province, the provincial government was allowed to set tax rates for a limited range of local taxes as well as the sharing ratio of the revenues from central-local sharing taxes (Ahmad 2008; Li 2010; Niu 2013). Similarly, such revenue-sharing arrangements dominated by the prefectural government also exist within prefecture-level cities (Man 2011; Wang and Herd 2013).

Table 4.1 Current tax sharing system in China

Taxes	Central(%)	Local(%)
Central Tax		
Tariffs	100	0
Consumption Tax	100	0
Shared Tax		
VAT	50	50
Business Tax (abolished in 2016)	3	97
Stamp Tax on Security Exchange	97	3
Personal Income Tax	60	40
Company Income Tax	60	40
Local Tax		
Resource Tax	0	100
Urban Maintenance and Development Tax	0	100
Urban Land Using Tax	0	100
Agriculture Tax (abolished in 2016)	0	100
Tax on Contracts	0	100
Tax on the Use of Arable Land	0	100
Vehicle Purchasing Tax	0	100
Other Local Taxes	0	100

Source: Ministry of Finance of PRC

On the spending side, the previous assignments of expenditure responsibilities between the central and local governments were basically unchanged in the reform and remain to the present (Shen et al. 2012), as shown in Table 4.2. In line with international practices and principles mentioned in the Literature Review (Chapter 3), "pure" public services and responsibilities with nationwide effects, such as national defence and macroeconomic planning, are carried out by the central government. Local governments are responsible for the local economy, operating local institutions, and providing services with regional effects and preferences (Shen et al. 2014). Specifically, within provinces, responsibilities related to public services such as healthcare, education, and social security are largely decentralised by the provincial government to sub-provincial governments and functional agencies (Su and Zhao 2004; Dollar and Hofman 2006; Niu 2013). Moreover, for these specific responsibilities, local cadres are not only responsible for spending money but also in charge of formulating and implementing policies related to these expenditures (Chien 2010; Donaldson 2017b).

To support local governments in performing their responsibilities, a formal transfer payment system was built for the first time, including tax rebates, general or unconditional transfers, and specific or conditional transfers. The amount of tax rebates in an area is positively determined by the size of gross tax revenues, thus most tax rebates are given to wealthy regions. In the 1990s, over 75% of transfer grants in China were tax rebates. To reduce fiscal disparity, in recent years the latter two forms of transfer payments, particularly unconditional transfers, were increasingly given to underdeveloped jurisdictions. Therefore, the ratio of tax rebates in total transfer grants decreased from 75% to 14% in 2011 (Niu 2013).

Table 4.2: Spending responsibilities of the central and local governments

Central government spending responsibilities	• National defence, armed police force, and foreign affairs • The operation of the central government and central departments • Macro-economic control and coordination of economic development • Providing funds for universities, hospitals, research institutions, and newspapers/media controlled by the central government
Local governments spending responsibilities	• Providing local public infrastructures and services, including compulsory education, healthcare service, culture and entertainment service, local transportation, and social service • Urban development and maintenance • The operation of the local government and departments; the operation of local public security bureaus (police station), procuratorial organs, and local courts • The development of the local economy; local development and construction projects

Source: Ministry of Finance.

In recent years, there are also minor changes to the existing tax-sharing system: from June 2019, the proportion of value-added tax revenues shared by the local government was increased from 25% to 50%, to compensate for the losses of local governments from the 1994 reform which replaced the business tax – the main source of local tax revenues, with a value-added tax (Li 2016). Also, over the past two decades, the above decentralisation arrangement was sometimes criticised for lacking detailed specifications and regulatory underpinnings (Dollar and Hofman 2006), until the State Council issued an official reform proposal in January 2018 where the roles of the central government and provinces in public services were clearly institutionalised (State Council 2018).

4.1.4 Underpinning FD with upward accountability

The above paragraphs summarise the evolution of inter-governmental fiscal and administrative relations over the past decades, from which it appears that a high level of FD exists in the Chinese background. However, despite a large proportion of revenue still belonging to the central government, the noticeably high level of fiscal and administrative power within local governments might still jeopardise the aforementioned upward accountability mechanism that upholds the inter-governmental hierarchy of China. In fact, during the 1980s, with the decentralisation of fiscal revenues and responsibilities, the central government lost control over those wealthy provinces, where local leaders gained decisive power in local policies and the management of local officials. With the authority of the central government and the CPC Central Committee increasingly being challenged, the upward accountability system existed in name only. While recognising the benefits of decentralisation – extreme poverty and famine were largely eradicated during the 1980s, since the early 1990s, two control mechanisms have been established and promoted by top policy-makers to underpin upward accountability without sacrificing the advantages of local governments in local affairs: the first mechanism is the target responsibility system (TRS), and the second one is the top-down PMS (which has been briefly introduced in Chapter 2). Over the past three decades, these two control mechanisms have not only been developed and employed by the central government and the CPC's top leadership but have also been increasingly applied within Chinese provinces.

The first control mechanism, TRS - the Chinese version of "Management by Objectives" (Whiting 2001; Edin 2003; Tsui and Wang 2008), as implied by Chan and Gao (2008) and Walker and Wu (2010), is nowadays used by almost every level of Chinese governments to assure their political control over the subordinated governments and functional agencies. Through the TRS, the higher-level government formulates target contracts jointly with the lower-level governments (e.g. the provincial government sets contracts with governments of the affiliated prefecture-level cities). Within a jurisdiction (e.g. a prefecture-level city), in line with the direction of upward accountability, the local government then separately formulates contracts with the subordinate functional departments (e.g. a prefecture-level city's HC) covering the detailed targets related to their specific responsibilities. In these contracts, not only the targets of the lower-level government are specified, but also, the relevant performance criteria are jointly set (Chan and Gao 2008; Zhou 2009; Shi and Ni 2017). Within a province, for the government of a sub-provincial jurisdiction, the general targets can be roughly classified into three categories: economic performance, performance in social development (including PSP), and party-building performance which focuses on the work of local CPC organisations (Edin 2003; Tsui and Wang 2004). With TRS, the target setting of Chinese local governments has been gradually institutionalised. The key policy agendas, requirements, and preferences of the upper-level government, via TRS, can be transmitted to the lower-level governments and functional agencies and reflected in the jointly formulated targets. Meanwhile, as the co-setters of targets, local cadres are sufficiently authorised to utilise their information advantages in terms of the local status, demands, and resources in target formulation (Ho 2010; Liu et al. 2021). Moreover, they have a high level of autonomy to make and implement detailed policies in the target-attaining stage (Gao 2009).

The TRS ensures that the key agendas, requirements, and preferences of superiors can be delivered to lower-level governments/agencies and reflected in the jointly formulated local targets. While the second underpinning mechanism of the upward accountability regime, i.e., the top-down PMS ensures that local cadres, especially major leaders of the local government (e.g. the mayor of a city) and functional departments (e.g. a prefecture-level city's HC), can be evaluated in accordance with their performance in fulfilling major targets. In line with the directions of upward accountability, the top-down PMS of local government leaders and leaders of the functional agencies are carried out by their direct one-level above superior government (Lam 2010; Liang and Langbein 2015). For example, within a prefecture-level city, the performance of the government leader (e.g. the city's mayor) is managed by the province, while the performance of a functional agency's (e.g. the city's HC) leader is determined by the city's government. Consistently, within a functional department, the departmental leader also holds his/her subordinates accountable through the top-down performance evaluation (Ye and Ni 2013). Despite differences in specific performance targets, those cadres – not only the major leaders of local governments and functional agencies but also grassroots cadres within a functional agency – who fail to pass the performance evaluation could be punished by demotion, dismissal, or by cutting down the funding of their departments (Huang 1995; and Tsui and Wang 2008; Liang and Langbein 2015; Qiu and Macnaughton 2017). As illustrated by Ma (2016a), with high-powered incentives such as promotion and funding, local cadres in China are effectively encouraged to devote attention to the targets and the performance goals highlighted by their superiors. Moreover, the rewards or punishments of local cadres are determined by not only their performance but also the performance of their peers at the same administrative level in similar situations – those performing "relatively" worse might be punished as well (Chen et al. 2005).

The above arguments show that by clarifying the upper-level government's policy agendas and preferences and by linking personal benefits with the performance in fulfilling responsibilities, local governments and agencies are held accountable to the upper-level government (Gore 2019). In other words, the upward accountability mechanism is strengthened. In this case, FD will not encourage local cadres to deviate from their superiors. Instead, their local policies will be more consistent with the key agendas, preferences, and requirements of the higher-level government leaders. Nevertheless, local cadres have the space and motivation to fulfil their policy targets in their own way (Chien 2010), so as to satisfy their superiors and maximise their personal benefits. The upward-accountability-underpinned FD, as argued by many researchers, is a key explanation for the socioeconomic achievements throughout the ongoing process of Reform and Opening-up since the late 1970s (Wong and Bird 2008; Shen et al. 2012).

4.1.5 Conceptualising the Chinese-style FD

Over the past decades, numerous theories have been developed to describe the unique pattern of FD with a strong underpinning of upward accountability in China. Among these theories, a consensus is reached that although a top-down unitary political leadership remains, a highly federated administrative system has been established for the vertical arrangements of government authorities and their related spending power (Wong 1991; Montinola et al. 2004; Shen et al. 2012; Wingender 2018). However, differences have emerged when researchers conceptualise such unique federalism. Researchers such as Shih et al. (2004) and Liu et al. (2018) define the distorted federal system with fiscal decentralisation and political authoritarianism in China as "predatory federalism". They suggest that fiscal responsibilities between different levels of governments deliberately remained unclear. Thus, the upper-level government can take advantage of their political power to grasp more resources, while leaving enormous responsibilities to the grassroots-level governments (Tsui and Wang

2008). To some extent, the concept of predatory federalism reflects the inter-governmental fiscal relations after the 1994 Tax Reform and before the early 2000s. However, in recent years, the allocation of duties between the central government and provinces and the allocation of duties within provinces have been increasingly becoming clearer (Bai and Liu 2020). Also, although the current institution of revenue allocation still, to some extent, favours the central government, the size of transfer grants, particularly unconditional transfers, has increased significantly, and it is clearly stipulated that local governments' fiscal gap in public service provision will be covered by the higher-level government (Niu 2013). The above facts indicate that predatory federalism may fail to fully describe the inter-governmental relationship in contemporary China and in the future.

Different from Shih et al.'s (2004) generally negative view, some researchers conceptualise the inter-governmental relationship in China as 'market-preserving federalism' where the vertical allocation of fiscal and administrative power is designed to avoid local markets being encroached upon by the central government (Qian and Weingast 1996, 1997; Montinola et al. 1995). As summarised by Montinola et al. (1995), the governmental system in China is highly hierarchical and tightly specifies that each government at a certain level is autonomous and in charge of specific duties within its sphere of authority, and the vertical arrangement of power and authorities cannot be easily changed by the upper or lower-level governments. The central government undertakes nationwide duties such as providing national public services (e.g. national security) and maintaining the common market, while sub-national governments play a dominant role in local socioeconomic affairs with the support of local resources. In this case, local cadres are accountable to not only the upper-level government but also the local residents – to avoid being voted out by their feet (Jin et al. 1999). Compared with the theory of predatory federalism, market-preserving federalism clearly reflects the "double-accountability" of local cadres in China, but it ignores the unique feature of the Chinese-style decentralisation, that is, the legitimacy of the local government is authorised by the upper-level government instead of being obtained in a bottom-up manner (Feng et al. 2013).

Combining the above two theories, another widely accepted definition named *de facto federalism* was proposed by Zheng (2006, 2007) to conceptualise the Chinese intergovernmental relation and to explain how both decentralisation and unitary features co-exist and play their roles simultaneously, with the following definition:

> A relatively institutionalised pattern which involves an explicit or implicit bargain between the centre and the provinces, one element in the bargain being that the provinces receive certain institutionalised or ad hoc benefits in return for guarantees by provincial officials that they will behave in certain ways on behalf of the centre (Zheng 2006, p. 107).

Based on the above definition, it appears that the *de facto federalism* theory - although developed to explain decentralisation between the central government and provinces - most successfully conceptualises the key features of Chinese-style FD within provinces: first, *de facto federalism* argues that there is a hierarchical system where the power of governments at each level can be arranged in a relatively institutionalised way – even though sometimes through informal institutions. This, to a large extent, reflects the institutional background where the Chinese-style FD is located and operated - as indicated by the above descriptions about the allocation of revenue and spending responsibilities (see Table 4.2). In particular, the design of FD in China follows a layer-cake pattern: the central government decentralises to the provinces. Then, in each province, the provincial government determines the allocation of responsibilities and revenues with local prefectural jurisdictions within the province, and so on (Wang and Herd 2013; Huang et al. 2017). Nowadays, the above arrangements between the central government and provinces and the arrangements within provinces (between the provincial and sub-provincial governments) have been increasingly becoming clearer (State Council 2018; Bai and Liu 2020).

Second, *de facto federalism* correctly identifies that the institutionalised allocation of power makes it unrealistic for the central government to directly impose its requirements on the provinces without bargaining with them. This second key argument of the *de facto federalism* theory is also consistent with the evolution of FD within Chinese provinces over the past decades: before the 1980s, most prefectural units in provinces were prefectures (not prefecture-level cities). In prefectures, governments were dispatch offices staffed by officials sent by the provincial government with almost no discretion in raising revenues and undertaking local responsibilities (Donaldson 2017b). From the 1980s to the early 2000s, most prefectures in China were upgraded as prefecture-level cities for which the government is no longer the dispatch office of the provincial government. Accordingly, prefecture-level cities have gained greater power in formulating and implementing local policies in various fields, such as local economy, urbanisation, public service provision, and environmental protection (Wu et al. 2006; Chien 2009; Lam 2010; Donaldson 2017b). In fact, the power granted to prefecture-level cities during the past decades is even greater than that decentralised from the centre to the provincial governments (Chien 2009; Duan and Zhan 2011).

Third but not least, *de facto federalism* highlights that decentralising power related to local affairs to provincial governments does not violate the central government's role in nationwide duties. Instead, in return for their autonomy, provincial cadres have to obey the leadership of the central government and the CPC in terms of the key political and policy agendas. The above arrangements, within provinces, are also reflected in the establishment of the aforementioned TRS and the top-down PMS within Chinese provinces. For TRS, researchers have widely discussed its application in Chinese grass-root jurisdictions, in which the roles of the higher-level government (usually the provincial government) as an “interpreter” of China’s top leadership as well as a “dominator” of the local key policy agendas and performances are demonstrated (Chan and Gao 2009; Walker and Wu 2010). While for the top-down PMS, within provinces, the institution that the “higher-level government manages the officials one level down” (“Xiaguan Yiji” in Chinese) has also been established (Chan and Gao 2008; Lam 2010; Liang and Langbein 2015; Donaldson 2017a).

In summary, the above arguments demonstrate that a high level of decentralisation can be seen in most provinces of China, where sub-provincial governments undertake major roles in local socioeconomic policies and the provision of public services. Nevertheless, a provincial government, similar to the central government, is still able to control local governments and officials and avoid them violating the key policy agendas and requirements of the central and provincial governments. This arrangement of inter-governmental relations within provinces, to a large extent, can be viewed as a duplication of the central-provincial relations in China. Accordingly, the term *de facto federalism* developed by Zheng (2006) to describe the central-local relations still appears to be applicable to conceptualise the intra-provincial FD in the Chinese context. In the next section, the evolution of China’s healthcare sector will be reviewed in detail.

4.2 Healthcare governance and reforms in China

In the history of the PRC, people have witnessed a series of evolutions in healthcare governance and administration. Thus, after giving a detailed introduction to FD in China, this section provides a detailed introduction to the historical and in particular, current systems of healthcare governance in China, so as to better understand the potential relationship between FD and healthcare performance.

4.2.1 Before the 1980s: the centrally planned healthcare system and achievements

In a short period after 1949, a Soviet-style central planning healthcare system was established in China (Jin and Sun 2011). Policies regarding healthcare services were mostly formulated by the central government, and the role of local governments and healthcare departments was to carry out these policies and only partially fund local healthcare institutions (Burns and Huang 2017). In line with the hukou system which strictly limited urban-rural migration, healthcare institutions were also separately divided into three tiers in urban and rural areas. In rural areas, there were village healthcare organisations (village clinics and “barefoot” doctors operating out of village clinics), township health centres, and county hospitals. In urban areas, this system started from street clinics, district hospitals and specialist clinics, to provincial and city hospitals. At that stage, all healthcare institutions were state-owned and funded by government expenditures and insurance schemes, following the free-to-access principle (Hu et al. 2011; Burns and Huang 2017).

During the 1950s to early 1980s, there were three basic insurance schemes covering more than 90% of the population: the Cooperative Medical Scheme was provided for rural residents, for which only a quite small portion of insurance payments were self-funded (0.5% - 2% of annual income). In urban areas, the mandatory Government Insurance Scheme and Labour Insurance Scheme were respectively provided for government employees and staff of public-owned enterprises. These two insurance schemes were fully financed by the government and enterprises (Lennart et al. 1996). Despite the ideology of egalitarianism, the design and operation of healthcare institutions and insurance schemes still showed a preference towards urban residents. This could be seen as a reflection of the Soviet-style socioeconomic system which allows for sacrificing agriculture to achieve industrialisation (Harrison 1996; Burns and Huang 2017).

Fiscally, in an underdeveloped context, it was a great burden for central and local governments to play a dominant role in healthcare service provision and insurance schemes. To cope with such burdens, the design and operation of healthcare institutions sought to follow the most cost-efficient pattern. As highlighted by Sidel and Sidel (1975), to save costs, healthcare services in China between the 1950s to the 1970s paid more attention to traditional Chinese medicine, disease prevention, and primary healthcare services. In rural areas, healthcare service provision mainly relied on “barefoot doctors” rather than grassroots healthcare institutions (Jin and Sun 2011). Moreover, under the ideology of collectivism, mass campaigns were frequently initiated by the CPC to deal with various infectious diseases, by which significant outcomes were achieved without paying much for healthcare services (Hsiao 1995).

Despite lacking resources and the presence of urban-rural inequality, internationally noticeable outcomes were still achieved in China: from 1950 to 1982, the yearly IMR reduced from 19.5% to 3.4%, and life expectancy rose from 40 to 68 (Blumenthal and Hsiao 2005). Such outcomes were achieved at a relatively low cost: on average, healthcare expenditure only accounted for 3% of the total GDP (Ramesh et al. 2014). More remarkably, the Chinese pattern of healthcare service delivery was recommended by WHO for developing countries (Hu et al. 2011).

4.2.2 Healthcare governance between the 1980s and 2009: the introduction of market incentives and its problems

After 1978 and before the mid-1980s, healthcare was not a key concern in the process of Reform and Opening-up (Hu et al. 2011). However, from the mid-1980s, with the gradual disintegration of the centrally planned socioeconomic system, healthcare reform was started, following similar strategies adopted in the reforms of central-local relations and state-owned enterprises (Hillier and Shen 1996; Ramesh et al. 2014). Accordingly, the responsibilities of central and local governments in healthcare financing and provision were decentralised and partially replaced by market elements (Jin and Sun 2011; Akin et al. 2005). In 1993, marketisation in healthcare service provision was further recognised as a national policy by the CPC (Central Committee of the CPC 1993). In this context, several substantial changes in healthcare governance had been witnessed: first, public healthcare institutions gained more fiscal independence. Especially after the 1994 Tax Reform, healthcare spending and the power of formulating local healthcare policies were mostly decentralised to the sub-national governments (Yip et al. 2012). In this period, the predominant focus of local governments was economic development and generating revenues, while providing healthcare services was not economically attractive, thus decentralisation in the 1990s led to a further reduction of fiscal expenditure on healthcare services (Zhou 2010). To cover the fiscal gap, all healthcare institutions, even the Centres for Disease Control and Prevention, were viewed as profit-making units and were authorised to create revenues from service provision and drug mark-ups (Wang and Wang 2007).

Second, apart from authorising public healthcare institutions with stronger profit-making power, private healthcare institutions were allowed from the 1980s, in order to create a competitive market and increase the efficiency of public healthcare institutions (Li and He 2019). Even in some areas such as Xinzheng, Henan Province and Heze, Shandong province, the privatisation of public healthcare institutions was conducted as a policy pilot (Lin 2018; Li and He 2019). However, privatisation was not rolled out due to legal and ideological concerns related to the Socialist system of the PRC (Gu and Zhang 2006). Additionally, apart from the above changes, starting from the 1990s, two of the three basic insurance schemes, the Cooperative Medical Scheme and Labour Insurance Scheme, were struggling because of the disintegration of collective farms and the bankruptcy of many state-owned enterprises. From the late 1970s to 1993, the proportion of citizens covered by health insurance reduced from more than 90% to 20% (World Bank 2003).

Various problems emerged in the 1990s because of the above changes. At first, the marketisation reform distorted the incentives of healthcare service providers from public interest to profit maximisation (Gu and Zhang 2006; Wang and Wang 2007; Ramesh et al. 2014). As Liu and Hsiao (1995) and Xu et al. (2010) discussed, because of this reform, many healthcare institutions tended to provide only profitable services and prescribe expensive drugs for customers. Meanwhile, the demise of insurance schemes significantly increased the fiscal burden of ordinary people: from 1978 to 2006, the proportion of healthcare costs attributable to out-of-pocket expenses increased from less than 20% to nearly 50% (Jin and Sun 2011). More seriously, the quantity and quality of healthcare services in rural areas were drastically undermined, because the limited healthcare expenses drifted to urban areas, and many rural residents cannot afford healthcare products in the private market (Hillier and Shen 1996). The above changes made it more difficult for people to access healthcare services (Ramesh et al. 2014). In an evaluation report published in 2005, the government admitted that the earlier reform in healthcare was basically a failure (Development Research Centre of the State Council 2005; Huang 2009).

4.2.3 The current system of healthcare governance: results of the second-phase healthcare reform

Recognising the problems with the early-stage reforms, top leaders of China restarted to highlight the importance of promoting healthcare service performance in their political agendas such as “harmonious society” and “building a service-oriented government” (Central Committee of the CPC 2006). In 2009, the second round of healthcare system reform was initiated by the central government, which is still ongoing and shaping the current healthcare system in China. In an official guidebook for this reform, the central government re-acknowledged that healthcare is a public service where the public sector should play a dominant role in policy-making, funding, operation, and healthcare supervision (Ramesh et al. 2014). Moreover, in recent years, the dominant role of local governments particularly sub-provincial HCs in formulating detailed healthcare policies and providing local healthcare services was further clarified (Hipgrave et al. 2012) – until now, 31 of 33 provincial units in Mainland China have institutionalised the inter-governmental relations in terms of healthcare responsibilities within the provinces. Accordingly, various policies have been formulated and implemented by local HCs under the decentralised settings of healthcare spending and governance (Bloom 2011; Hipgrave et al. 2012; Yip et al. 2012), in which primary healthcare and the professional management of the healthcare sector are two key issues.

Regarding primary healthcare in urban communities and rural areas, ongoing reforms are mainly concerned with service delivery in healthcare institutions such as community healthcare centres and village clinics. Over the past two decades, there has been a serious resource-responsibility mismatch between healthcare institutions at different levels (Sussmuth-Dyckerhoff and Then 2017). Based on size and functions, healthcare institutions in China are designated as Class III hospitals, Class II hospitals, Class I hospitals, and other primary healthcare institutions such as community healthcare centres, township hospitals, and village clinics. Class III hospitals are located at the top of the hierarchy of healthcare institutions, engaging in specialist health services, medical education, and scientific research. Class II/Class I hospitals and primary healthcare institutions provide general health services, preventive care, and treatments for chronic and general diseases. However, many patients, no matter what illnesses they are suffering from, prefer to visit Class III hospitals rather than lower-level healthcare institutions (Sussmuth-Dyckerhoff and Then 2017; Li et al. 2017), because they believe that Class III hospitals have higher quality facilities and personnel (Eggleston et al. 2008; Yip et al. 2012). Also, there is no patient referral system or appointment system to limit unnecessary access to Class III hospitals (Liu et al. 2017). The above problems led to a severe distortion of workloads and duties between Class III hospitals and other institutions: Class III hospitals spent massive resources on minor and general diseases rather than specialist health services, research, and education. Meanwhile, the overloaded operation of Class III hospitals also caused other problems such as long waiting times and patient-doctor tensions (Liu et al. 2017). On the other hand, medical resources in primary institutions – particularly those located in urban areas – were underutilised due to fewer visits of patients, leading to lower productive efficiency. While in rural areas, especially in underdeveloped provinces, a more urgent problem for primary healthcare institutions was the lack of healthcare resources and professionals (Sussmuth-Dyckerhoff and Then 2017). As mentioned above, the urban-rural gap in healthcare resources became greater during the first phase of healthcare system reform. In 2008, the per capita healthcare expense (including government and private expenses) in urban and rural areas was 1862.3 RMB and 454.8 respectively[15].

To deal with the above problems, since 2011, local spending on primary healthcare services in rural areas has significantly increased (Hipgrave et al. 2012, Ramesh et al. 2014). These expenses were not only allocated to update medical infrastructure but also to staff-training programmes (Sussmuth-Dyckerhoff and Then 2017). Moreover, preferential policies such as higher reimbursement ratios were also given to encourage people to visit primary healthcare institutions (Hipgrave et al. 2012). However, until now, the referral system has only been established in a few places (Ramesh et al. 2014), and it still needs a long period to reverse people's preference for Class III hospitals. Additionally, the urban-rural gap in healthcare resources and quality is reduced, but still exists (Li et al. 2020).

[15] Data sources: National Bureau of Statistics

Regarding the professional management of the healthcare sector, key issues are about strengthening the oversight of healthcare institutions while respecting and promoting professional management. For decades, professional healthcare officials, particularly those of public hospitals, are blamed as simultaneously being operators and supervisors, which creates opportunities for regulatory capture (Meng et al. 2012). Researchers also identified that some administrative bodies, especially the HC, might collude with hospital managers and physicians to maximise their private interests (Hsiao 2007). To deal with the above problems, the central government provided a series of policy suggestions from 2015: first, the healthcare administrative, managerial, and regulatory bodies should be separated, so as to give greater autonomy to the managerial team while keeping them under a more intensified oversight (Hipgrave et al. 2012). As explained by Liu et al. (2017), this initiative is effective in holding public hospitals more accountable to their agencies in local governments, thus ensuring public benefits. Meanwhile, the regulatory control of the higher-level governments over subordinate governments and healthcare agencies was consolidated as well, which strengthens vertical accountability and motivates local cadres to pay closer attention to healthcare improvement (Ramesh et al. 2014). Second, within healthcare institutions, greater autonomy of personnel management should be given to the professional management team, as long as they are under close and independent oversight (Liu et al. 2017). With greater flexibility, personnel management within healthcare institutions is intended to be more consistent with the characteristics of the medical profession. Nowadays, an obsolete tenure system with tiered income structures still exists in public healthcare institutions. By carrying out the above initiatives, the performance of public healthcare institutions and professionals is expected to be promoted (Sussmuth-Dyckerhoff and Then 2017).

4.2.4 Summary

In recent years, international experiences have fully demonstrated the importance of government intervention in healthcare service provision (Ramesh 2008): providers usually pay more attention to public interest if the government has a strong determination to shape healthcare policies and perform their regulatory functions. This argument is also tenable in the Chinese background. When the central and local governments in China attempted to initiate marketisation reforms, various problems emerged – not only within the healthcare sector but also throughout society. With the role of government in healthcare re-emphasised in second-phase reforms, these problems are a major focus for public service improvement efforts within and across Chinese provinces, for which various treatments are expected to be initiated by the decentralised local healthcare sector. Thus, starting from the next chapter, whether the decentralised sub-provincial governments improved healthcare service performance will be investigated empirically.

CHAPTER 5 METHODOLOGY

The previous chapter provided a detailed review of the context and institutional background of this doctoral research. This chapter presents the research methodology – the systematic way to solve the research problem (Kothari 2004). It starts by introducing the research paradigm, research strategies, and research design in the first section, in which the positivist stance, quantitative strategy, and the correlational research design are justified. Then, the following sections provide detailed specifications about the quantitative methods, starting from the units of analysis and the choice of data sources (5.2). After that, Section 5.3 and 5.4 focus on the evaluation of healthcare efficiency and effectiveness – two key performance aspects discussed in this study as well as the measurements of FD, in which summary statistics are also reported. The fourth and fifth parts (5.2.4 and 5.2.5) provide detailed explanations of the variables, quantitative models, and estimation methods. Finally, a summary of this chapter is given in Section 5.3.

5.1 Research Paradigm, strategies, and design

This section starts by justifying the paradigm for this research. Despite numerous, sometimes conflicting definitions (Mackenzie and Knipe 2006), scholars broadly conceptualise research paradigms as a set of beliefs and thoughts shared by researchers to direct the whole process of investigation (Guba and Lincoln 1994; Creswell 2009). A paradigm is considered as the worldview of researchers which not only illustrates how they see the world but also how it influences their way of framing a research topic, looking into this topic, and interpreting the research data (Lather 1986; Willis 2007). Thus, as summarised by Grix (2002), a research paradigm refers to the philosophical foundation of thinking which is concerned with two major concepts: ontology and epistemology.

5.1.1 Ontology

Ontology denotes the philosophical position researchers hold about the nature of social reality or existence (Krauss 2005), which, as argued by Saunders et al. (2016), is about how the world works. There are two main perspectives on ontology: the first one is objectivism, arguing that social entities exist external to social actors (Bryman 2012). The second one, subjectivism, holds a constructionist view that all social phenomena are subjectively created and continuously renewed by social actors via their interactions (Saunders et al. 2007). The researcher of this study follows the objectivist stance, that is, believing that social entities objectively exist and "have a reality external to social actors" (Bryman 2012, p. 75).

5.1.2 Epistemology

Different from the focus of ontology on the nature of reality, epistemology is about the nature of knowledge. It is concerned with what should be perceived as acceptable knowledge and how such knowledge can be acquired (Blaikie 2003; Saunders et al. 2007). A central question for epistemology is whether principles and approaches of natural science can and should be adopted in the field of social science (Bryman 2012). Major epistemological stances, including positivism, interpretivism, critical realism, and pragmatism, are explained in the following paragraphs.

- **Positivism**

Following the objectivist ontological stance that considers social entities exist independently, positivism argues that acceptable knowledge can be obtained from "observable facts and data" (Saunders et al. 2016, p. 150). Focusing on causal relationships, positivism aims to find law-like and highly generalisable knowledge from observable social facts. Thus, in their research processes – from data collection to analysis, researchers should be impartial and avoid influencing or being influenced by their research subjects (Remenyi et al. 1998). Positivists usually follow a deductive pattern of reasoning, that is, proposing hypotheses based on existing theories and knowledge. Then, hypotheses are scientifically assessed to justify or challenge the existing theories through which new knowledge and underlying logic can be recognised (Ghauri and Gronhaug 2002; Saunders et al. 2007). Therefore, positivists prefer to employ the highly structured and systematic methods used in natural sciences for social science inquiries (Mertens 2005). In other words, positivists aim to identify highly generalisable knowledge by collecting numerical data from a large sample group and analysing the data with quantitative methods (Gill and Johnson 2002; Creswell 2009).

- **Constructionism**

Ontologically, constructionism stands on a subjectivist position, suggesting social realities are constructed by the members of social groupings (Mertens 2005). Constructionists believe that acceptable knowledge is created by people from their personal experiences and interactions with others (Burr 2015; Saunders et al. 2016). Thus, constructionist researchers locate themselves in the human world to understand "the world of human experience" (Cohen and Manion 1994 p.36), in this case, they are far from impartial or value-free (Sandberg 2005). Recognising all knowledge is developed under certain social contexts where the research subjects are located, constructionists believe that it is impossible to generate law-like and generalisable knowledge within various contexts (Mackenzie and Knipe 2006), and it is superficial to just focus on causal relations when probing into human actions (May 2011). Accordingly, constructionists hold an inductive stance, that is, research does not start from a mature theory and the tests of theories, instead, it is about generating theories via an inductive pattern of data collection and interpretation (Creswell 2009). Thus, regarding methods, constructionists prefer to conduct in-depth research with qualitative data from a relatively small sample group (Bryman 2012).

- **Critical realism**

Critical realism is viewed as a "middle ground" between positivism and constructionism (Krauss 2005). Similar to positivism, ontologically, critical realism argues that there are realities in the external world that are independent of people's descriptions and minds (Saunders et al. 2016). Epistemologically, critical realism believes what we see is only the image of reality rather than reality itself (Saunders et al. 2016). Therefore, to understand and impact the social world, researchers need to look deeply into the social structures which produce social phenomena and discourses via their practical and theoretical enquiries (Bhaskar 1989). This process of understanding social structures cannot be accomplished without involving social actors in the knowledge derivation process (Bryman 2012), which, to some extent, can be seen as critical realists' critique of the positivist stance. Given the aforementioned descriptions, critical realism supports the use of any numerical-based or non-numerical approaches to perform "retroductive" reasoning, which refers to an inference about the likely explanations for an observation that can be tested later (Blaikie 2003; Morin et al. 2021).

- **Pragmatism**

Different from the above three epistemological stances, pragmatism is not committed to a fixed philosophical system (Creswell 2009). It does not perceive the social world as an absolute unity. Moreover, the choice of a philosophical stance should be based on research questions – that is, choosing the most appropriate one for solving research questions (Saunders et al. 2016). Put another way, pragmatism might be the suitable choice if the research question does not evidently match any single philosophical stance such as positivism or constructionism. Accordingly, any research methods available to solve the research question can be adopted (Rossman and Wilson 1985;

Patton 1990; Creswell 2009). Thus, the pragmatic stance, as researchers have suggested, prefers pluralistic methods, especially the simultaneous adoption of quantitative and qualitative methods (Tashakkori and Teddlie 1998; Mertens 2005; Morgan 2007).

5.1.3 Epistemology of this study

This study recognises that all epistemological stances have their strengths and weaknesses, nevertheless, based on the research topic, a positivist stance is adopted due to the following reasons: first, this study aims to investigate the impact of one variable, i.e., FD on another key concept, i.e., healthcare service performance (proxied by efficiency and effectiveness), which, as explained by Kim (2003), can be better addressed by a positivist stance. Second, the positivist stance works well for developing highly generalisable knowledge, while this study also focuses on a broad context, i.e., Chinese provinces. Third, recognising the objective existence of social reality, the researcher of this study aims to conduct this research as impartially as possible, that is, to "neither affect nor be affected by the subject of the research" (Remenyi et al. 1998, p. 33), which, as explained above, is also one of the key arguments of positivism. Keeping impartial from the research subject under the positivist stance also facilitates replication, ensuring the reliability of this research (Kim 2003). Fourth, given the aforementioned research aims and questions, this study intends to propose hypotheses based on existing theories and research. Then, data will be collected from a large sample and analysed using highly structured methods.

Finally, new knowledge will be identified to illustrate, refine, or challenge the existing theories. Such a deductive–inductive process, as argued by Bryman (2012) and Saunders et al. (2007), is also a key principle of the positivist stance. Finally, positivism is proved by empirical studies to be suitable for business and management research (Thomas 2004) as well as being the dominant philosophical stance for FD studies (e.g. Barankay and Lockwood 2007; Widmer and Zweifel 2012; Ghuman and Singh 2013). Therefore, based on the reasons explained above, positivism is employed to direct the empirical part of this research.

5.1.4 Research strategies and design

In line with the research questions and philosophical stance, a suitable research strategy[16] of inquiry should be chosen to direct research procedures (Creswell 2009; Marczyk et al. 2005). There are two major strategies: the quantitative strategy and the qualitative strategy. Additionally, a study that combines both quantitative and qualitative features is referred to as a mixed-method one. (Creswell 2009; Bryman 2012). The quantitative strategy stands on the objectivist and positivist stance (Bryman and Bell 2011) and aims to identify social phenomena with highly structured approaches (Wilson 2010). Consistently, quantitative studies entail a deductive-inductive process to test and challenge the existing theories with massive numerical data and econometric approaches (Bryman 2012). In contrast, the qualitative strategy usually follows a subjectivist and constructionist stance, aiming to perform in-depth

[16] Some researchers call it research methodology (e.g. Mertens 1998; Creswell 2009, p. 11).

research and generate context-specific knowledge in a relatively limited setting, with popular designs such as ethnography, grounded theory, and case studies (Creswell 2009). The process of a qualitative research usually follows an inductive process, that is, to generate rather than test theories. The above two research strategies have been widely adopted in previous empirical studies, with the following advantages and disadvantages summarised in Table 5.1:

Table 5.1: Advantages and disadvantages of quantitative and qualitative strategies

	Quantitative strategy	**Qualitative strategy**
Advantages	● Data objectivity: researchers keep impartial and adopt objective data analysis methods ● High external validity and reliability: results are replicable ● High generalisability: massive data are analysed to identify highly generalisable results ● High efficiency: data can be collected and analysed in a short time frame	● Allowing researchers to merge into and understand the context ● Good at obtaining rich information about the subjective values, perceptions, beliefs, and behaviours of participants ● Good at performing in-depth studies in a relatively limited context ● Data interpretation: researchers have a high level of flexibility
Disadvantages	● Lacking detailed interpretation from real-world participants ● Risks in internal validity: variables are mistakenly measured, data are estimated with improper methods, etc ● Lacking understanding and analysis based on the research context	● Low generalisability: data are collected and analysed based on a limited context ● Results are highly subjective and cannot be verified, unknown reliability and validity ● Time-consuming

Recognising the above strengths and weaknesses, a quantitative research strategy should be adopted. This doctoral research probes into the relationship between FD and healthcare service performance in Chinese provinces, as well as the role of relative wealth in this relationship. This is achieved by formulating hypotheses based on existing theories and arguments at first, followed by the collection of extensive numerical data from a large sample group for structured data estimations (Sale et al. 2002). This ensures that impartial and highly generalisable findings can be developed to validate or challenge the theories about how phenomena occur (Johnson and Onwuegbuzie 2004). Obviously, a quantitative strategy performs well in directing the above research processes. This aligns with most empirical studies focusing on the relationship between FD and PSP (e.g., Adam et al. 2014; Arends 2017; Alonso and Andrews 2019).

- **Research design**

A suitable research design is vital for researchers to refine the research methods and successfully structure their studies (Churchill and Iacobucci 2002). When selecting a research design, it is crucial to choose the one that enables the most effective exploration of the research problems while considering practical constraints such as time and budget limitations (Ghauri and Grønhaug 2002). The major quantitative research designs encompass three types: descriptive, experimental, and correlational (Stangor 2011). A descriptive design is to offer a detailed 'profile of persons, events, or situations' (Robson 2002, p.59), which aims to determine the frequency of occurrences or understand the current state of the concerned issues (Churchill and Iacobucci 2002). With descriptive research, research problems could be organised and comprehended in a systematic manner (Ghauri and Grønhaug 2002). In other words, a descriptive study may serve as a foundation or precursor for experimental or correlational studies. An explanatory design focuses on identifying cause-and-effect relationships through experiments or quasi-experiments. Accordingly, such a study requires an experimental group and a control group to gauge the impact of the panned intervention or manipulation (Stangor 2011). For an experiments, the control and experimental group should be randomly selected, and there is no difference between these groups apart from the intervention/manipulation. These above restrictions could be relaxed in a quasi-experiment. Similarly, a correlational design also aims to identify relationships among two or more variables. Unlike experimental studies, correlational analysis does not require human manipulation and can rely solely on secondary data. Thus, with accessible dataset, correlational studies can flexibly examine any relationships in everyday life events with single or multiple regression analysis (Aiken and West 1991). Based on the above descriptions, a correlational design appears to be the most suitable for this research. Guided by this design, the research units, data, variables, and estimation methods will be proposed and explained in detail in the next section.

5.2 Units of analysis, the period of research, and data sources

To empirically investigate the impact of intra-provincial FD on healthcare performance, and in line with the research aim and questions, this study chooses all 22 provinces and four of the five Provincial Autonomous Regions for ethnic minorities across mainland China as the units of analysis. In line with the province-based research of Uchimura and Jutting (2009), the four provincial municipalities are not considered in this study. This is because their socioeconomic status (e.g. demographic characteristics, level of economic development, urbanisation) and administrative arrangements are significantly different from the other provinces. Moreover, different from ordinary provinces, top leaders of the four cities are all members of the CPC Politburo – the core executive committee of the party-state. This distinction set the four cities apart in terms of the inter-governmental interactions within these cities as well as the interactions between these cities and the central government (Li, 2007). Regarding the five Provincial Autonomous Regions for ethnic minorities, it is legally stipulated that they have the right to “formulate self-government regulations in accordance with local conditions of the ethnic group”. However, in practice, such autonomous power is strictly limited and all five autonomous regions are still under the leadership of the CPC and have the same inter-governmental political, administrative, and fiscal arrangements as other provinces (Zhang 2012). In other words, they are not substantially different from other provincial jurisdictions. Thus, apart from Tibet (due to the lack of data), four of the five provincial Autonomous Regions are also covered in this study. For simplicity, unless otherwise specified, these four provincial Autonomous Regions are described as provinces in the following chapters of the thesis.

The data for empirical analysis were collected from 2006 to 2017. First, this is due to data unavailability: before 2006, fiscal transparency in Chinese provinces was low. While there is a lack of data after 2017 because the newest government statistical reports are still awaiting publication. Second, as explained above, since 2006, improving services has become a top policy aim of the CPC and the central government (Central Committee of the CPC 2006). The New Healthcare Reform aiming to promote healthcare service performance was initiated in 2009 (State Council of China 2009), leading to the implementation of numerous local policies and initiatives, such as increasing healthcare expenditures, providing financial aid for rural and community healthcare, and carrying out new human resource management practices (Liu et al., 2017). Whereas FD prior to 2006 may have motivated local governments to focus on increasing productive expenditures (e.g. infrastructure) at the expense of people's welfare, the launch of the 2006 initiative of "building a harmonious socialist society" (Hu 2007) has required local governments to pay more attention to the performance of public services. In other words, after 2006, it is likely that with FD, local cadres would have devoted more attention to healthcare, aligning with the updated priorities set by higher-level government. Therefore, it is of timely importance to empirical investigate FD's impact on healthcare service efficiency and effectiveness from 2006 to 2017. By doing so, this study, in contrast to others based on outdated data, would provide different insights and up-to-date evidence regarding the FD's impact on the evolving commitments of local cadres towards different socio-economic targets.

In line with previous Chinese-based studies (e.g. Uchimura and Jutting 2009; Jin and Sun 2011; Hao et al. 2021), all quantitative data are collected from statistical yearbooks and budgetary reports that are openly published by the national and provincial Bureaus of Statistics as well as the central and provincial governments, ensuring consistency of data quality and reliability. Specifically, expenditure and revenue data for calculating FD indicators as well as data for other fiscal control variables were collected from the Finance Yearbooks of China, China Statistical Yearbooks for Regional Economy, and the provincial governments' yearly budgetary reports. Data for estimating healthcare efficiency and effectiveness and other healthcare-related control variables were collected from the Chinese Health Statistical Yearbooks. Data for other socioeconomic control variables (e.g. GDP, population, and urbanisation) were gleaned from the China Population and Employment Statistics Yearbooks. Summary statistics for all variables are shown in the following sections of this chapter and the Findings chapter. All the above data for empirical analysis are secondary and numerical, which to a large extent, avoids the occurrence of ethical problems. However, this study still strictly follows all requirements of the Research Ethics Committee of the Business School and the University, and the ethical approval was obtained before data collection.

5.3 Evaluating fiscal decentralisation within Chinese provincial jurisdictions

In the Literature Review chapter, the theories and principles of FD have been discussed in detail. Thus, this section mainly focuses on the selection of suitable indicators for gauging FD. In the past decades, numerous FD indicators have been proposed and adopted in empirical studies, most of which, as mentioned above, can be classified into three categories: expenditure-side indicators, revenue-side indicators, and combined indicators considering both revenues and expenditures (Cheol 2012). For the expenditure side and the revenue side, the proportion of expenditure spent by a lower-level government to that spent by its higher-level government, and the share of a local government's revenue relative to the revenue of the higher-level government, are two major indicators (Oates 1972; Grossman 1989; Jin and Zou 2002; OCED 2003; Afonso and Hauptmeier 2009). Additionally, the ratio of a local government's own revenue to its total revenue is also a popular indicator to measure revenue autonomy (see Boetti et al. 2002; Otsuka et al. 2014). Similarly, revenue and expenditure autonomy can also be measured by combined indicators such as the ratio of a local government's own revenue to its expenditure (see Martinez-Vazquez and Timofeev 2009; Alonso and Andrews 2019), and a local government's transfer payments to its total revenue and/or total expenditure (Rao and Singh 2007; Baskaran 2010; Bahl and Wallace 2007).

From the above classifications, three conventional but popular measures are chosen for this study: (1) healthcare expenditure decentralisation (HED) measuring the ratio of the healthcare expenditure spent by sub-provincial governments in a province to the total healthcare expenditure of this province (i.e., sub-provincial governments' expenditure plus provincial governments' expenditure, see Cantarero and Pascual 2008; Arends 2017); (2) total expenditure decentralisation (TED) measuring the total fiscal expenditure spent by sub-provincial governments in a province as a percentage of the total fiscal expenditure of the province (Uchimura and Jutting 2009); and (3) revenue decentralisation (RD) measuring the own revenue of sub-provincial governments in comparison to the total revenue of the province (i.e., sub-provincial governments' own revenue plus provincial governments' own revenue).

Despite wide application in FD research, the above three FD measures sometimes are criticised for their limitations in estimating the actual decision-making power in terms of fiscal spending (Stegarescu 2005; Gu 2012). Also, the real degree of RD might fail to be reflected by the ratio of revenue allocation between higher and lower-level governments if transfer payments are not excluded from the revenue of local governments. However, the first issue is not a serious concern in this research, because in China, the decentralisation of decision-making power in formulating and implementing local policies, as mentioned in the background section, has followed the trend of FD (Hao et al. 2020). In other words, HED and TED are two indicators that largely reflect both the decentralisation of spending responsibilities and the decision-making power in regard to spending policies. Then, to avoid the second limitation regarding RD, following Jiménez-Rubio (2011a), this study only considers revenues from local taxes, sharing taxes, and local non-tax revenues when measuring RD, while transfer payment revenues are excluded.

The measurement of revenue autonomy (i.e., a local government's own revenue / total revenue) and the combined indicators mentioned above, as indicated by (Gu 2002), are able to avoid the limitations of conventional measures. However, they are not suitable for this study. This is because these indicators fail to directly reflect the fiscal relationship between the provincial government and sub-provincial governments within a province, instead, they represent to which extent sub-provincial governments rely on revenues from all external sources which include but are not limited to the provincial government. Moreover, the data of transfer payments given by the provincial governments to sub-provincial governments in China are not accessible for the study period. Thus, considering the above issues, HED, TED, and RD are chosen to measure intra-provincial FD in China. As this study specifically focuses on the decentralisation and efficiency of a specific public service (healthcare), HED is considered the preferred measurement of FD. Summary statistics of HED, TED, and RD are reported together with other variables at the beginning of the Findings chapter (summary statistics of FD indicators by province are given in the Appendix). Figure 5.1 shows the yearly variation of HED, TED, and RD values from 2006 to 2017. The figure shows that for healthcare services within most Chinese provincial jurisdictions, expenditure authority, and government revenues, are highly decentralised, and an overall upward trend of decentralisation is demonstrated. Also, the level of HED is around 10% higher than that of TED and RD. Figures 5.2-5.4 map the average level of FD across provinces, with those that are more decentralised represented by darker colours. These figures show that jurisdictions in eastern and central China, generally, have slightly higher levels of decentralisation than their western counterparts.

Figure 5.1: Yearly variation of HED, TED, and RD

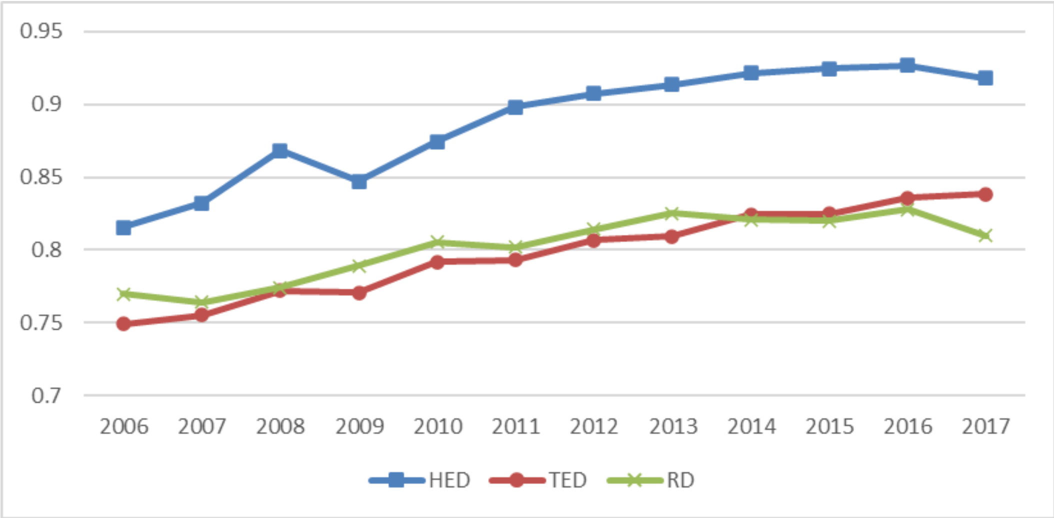


Figure 5.2: Map of average HED from 2006 to 2017

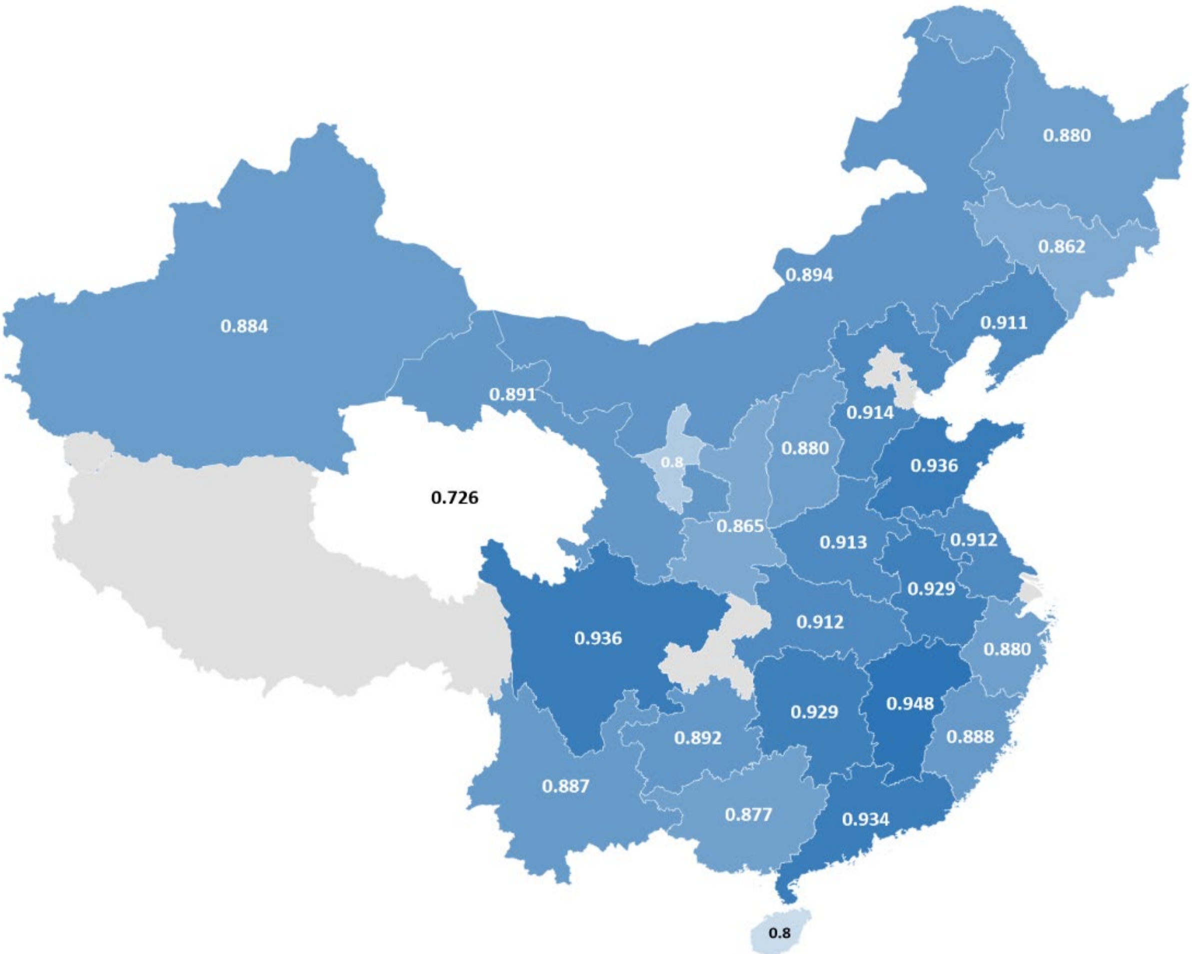


Figure 5.3: Map of average TED from 2006 to 2017

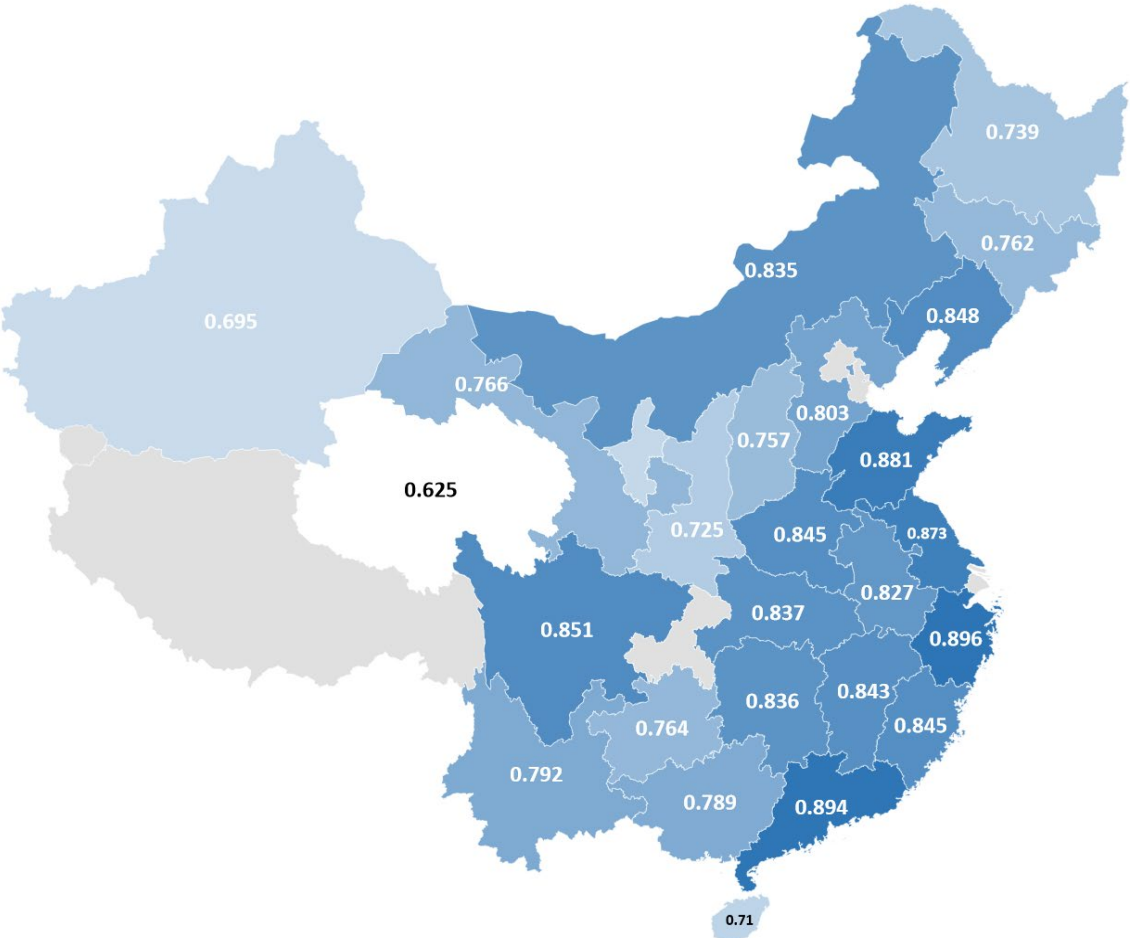
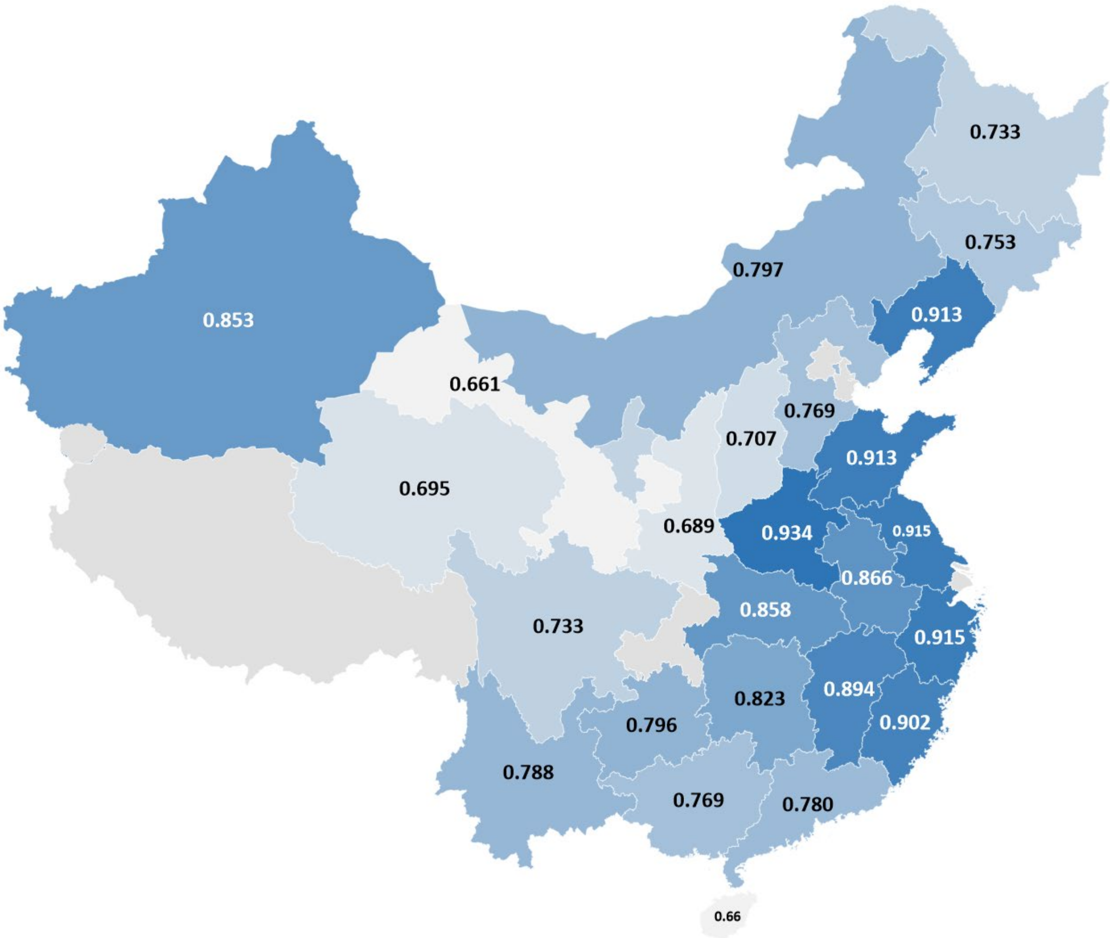


Figure 5.4: Map of average RD from 2006 to 2017



5.4 Measuring healthcare efficiency and effectiveness in Chinese provinces

Broadly following the 3E model (Boyne 2002), this PhD study is concerned with two key dimensions of healthcare performance: efficiency and effectiveness. In this section, measures of these two performance dimensions are explained in detail. Efficiency measures are calculated by the researcher using two popular approaches named DEA and super-efficiency DEA. Considering data accessibility and in accordance with previous studies (e.g. Habibi et al. 2003; Jiménez-Rubio and Smith 2005; Asfaw et al. 2007), healthcare effectiveness is measured by the survival rate of stillbirths and new births (i.e., during the perinatal period). The first part of this section (5.2.3.1) explains the efficiency calculation methods and reports summary statistics for efficiency calculations, while all details regarding the effectiveness measure are reported in the second part (5.2.3.2).

5.4.1 Measuring healthcare efficiency

For decades, both academia and society have considered efficiency as the centrepiece of PSP, and a broad consensus on the meaning of efficiency has been established (Walker 1937; Simon 1976). As explained in the second chapter, there are four types of public service efficiency: productive efficiency, allocative efficiency, distributive efficiency, and dynamic efficiency. This study focuses on the productive efficiency of healthcare services in Chinese provinces. As demonstrated in the second Literature Review chapter (Chapter 3), local governments have advantages in terms of collecting local information and utilising local resources. Therefore, FD ensures that local government could take advantage of their advantages to better satisfy the demands of residents in more cost-efficient manners. In other words, both theoretical and empirical studies have fully identified that the influence of FD is mainly targeted at the productive and allocative efficiency of public services. However, there is a lack of suitable measurements for allocative efficiency over a long time period. Therefore, this study focuses on the productive efficiency of healthcare services in Chinese provinces. In the following paragraphs, the methods for calculating productive efficiency are introduced.

- **Data Envelopment Analysis: theoretical basis**

Parametric and non-parametric approaches to measuring productive efficiency have been developed by researchers. Parametric approaches assume that input and output relationships are established following a specific functional form, such as the Cobb-Douglas or Translog forms (Cook et al. 2014). Then, various output indicators are weighted into a single output indicator. Finally, productive efficiency is calculated following the chosen function and econometric methods (Lin and Chen 2018). By contrast, non-parametric approaches do not hold assumptions in terms of input-output functions, but use linear programming techniques to measure the distance from each input-output combination to the most efficient scenario (Li and Lee 2010). Moreover, multiple inputs and outputs can be chosen when calculating productive efficiency with non-parametric techniques (Afonso and Fernandes 2008; Fonchamnyo and Sama 2016). Due to their high flexibility, non-parametric approaches are now more prevalent than their parametric counterparts in public sector research (Geys and

Moesen 2009). This study thus adopts two non-parametric approaches, DEA and the super-efficiency DEA, to estimate the level of healthcare service efficiency in Chinese provinces.

DEA was first introduced in the 1970s by Charnes et al. (1978) to measure the efficiency of public entities where revenue and profit are not considered as performance criteria. The basic notion of DEA is to build an efficiency frontier that "envelops" all input/output vectors of selected decision-making units (DMUs), with the support of linear programming techniques (Bowlin 1998). Notably, the formulation of such an efficiency frontier is based not only on the input/output performance of selected DMUs but also on all possible input/output combinations which could exist (Geys and Moesen 2009). Thus, for a DMU, an efficiency score is given by comparing its distance to those real and virtual DMUs located at the efficiency frontier. The highest score is 1, indicating that the selected DMU is efficient and located at the efficiency frontier. For inefficient DMUs, scores from 0 to 1 (excluding 1) are given in terms of their distance to the efficiency frontier (Afonso and Fernandes 2008; Hauner and Kyobe 2010).

In recent years, DEA has been increasingly adopted by studies in various disciplines and topics, particularly those focusing on the public sector. For example, De Borger and Kerstens (1996), Worthington (2000), Afonso et al. (2006), Balaguer-Coll et al.

(2007), and Cruz and Marques (2014) adopted DEA to measure the efficiency of basic public services (e.g. water, electricity, transportation, and waste recycling services) in Belgium, Australia, EU nations, Spain and Portugal, and Spain, respectively. While Hauner and Kyobe (2010) and Fonchamnyo and Sama (2016) investigate the efficiency of education and health services in their research based on 114 countries and 4 African countries. Moreover, 7 of the 10 papers focusing on the impact of FD on public service efficiency, as shown in the Literature Review chapter, adopted DEA. As mentioned above, the popularity of DEA in empirical studies can be attributed to the flexibility of the method. Moreover, as Coelli et. al (2005) suggested, DEA is particularly powerful when measuring efficiency with input and output indicators that are in the form of indices and where price data are unavailable. In other words, DEA is a suitable choice for this study where the efficiency of public services without market prices needs to be measured.

Three steps should be followed to calculate DEA efficiency scores: First, make a selection between input-oriented and output-oriented DEA. The former measures the extent to which inputs can be minimised at a given output, while the latter is concerned with the level of output at a given input (Afonso and Fernandes 2008). In a DEA review paper, Cook et al. (2014) suggest that the choice of orientation should be guided by the side (input or output side) over which the DMUs have greater control. In most cases, local governments have more control over public service inputs than outputs, and this study focuses on the arrangement of FD within the government and its impact on healthcare service efficiency, thus the input-oriented DEA is more suitable here (see also Balaguer-Coll et al. 2007; Alonso and Andrews 2019).

Second, this study chooses the Charnes et al. (1978) model (i.e., CCR model) or the Banker et al. (1984) model (i.e., BCC model) for performing DEA analysis. The CCR model presumes that all DMUs are operated at the optimal scale with constant returns to scale (CRS), while the BCC model criticises this assumption and suggests returns to scale are variable (i.e., VRS), thus efficiency scores can be decomposed into values of technical efficiency and values of scale efficiency. As it appears unreasonable to hold the assumptions of constant returns and optimal scale, this study adopted the input-oriented BCC model to calculate DEA efficiency scores, which is consistent with previous studies (e.g. Worthington 2000; Afonso and Fernandes 2008; Alonso and Andrews 2019). The BCC-DEA approach can be sketched as follows[17]:

$$maximise: h_k = \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i \, x_{ik}} \quad (1)$$

$$subjective\ to:$$

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1; j = 1, 2, \dots, n$$

$$u_r, v_i \geq \varepsilon > 0, r = 1, \dots, s, i = 1, \dots, m$$

In the above model, h_k is the level of efficiency of DMU_k (i.e., the distance between DMU_k's efficiency to the efficiency frontier). While y_{rj} and x_{ij} respectively denote the amount of the r^{th} outputs and i^{th} input of the j^{th} DMU.

[17] For detailed descriptions of the BCC-DEA approach, see Bowling (1998). Calculations of BCC-DEA efficiency values are conducted by the computer programme DEAP 2.1 (Coelli 1996).

Thus, $1 - h_k$ equation (1) measures the distance between DMU_k's efficiency to the optimal level of efficiency. It is clear that h_k cannot be larger than 1 by the virtue of the above constraints.

Supported by the research of Charnes and Cooper (1962), equation (1) can be solved by transforming it into the following linear programming form:

$$maximise: h_k = \sum_{r=1}^{s} u_r y_{rk} \quad (2)$$

$$subjective\ to:$$

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} \le 0, i = 1, \dots, n$$

$$\sum_{i=1}^{m} v_i x_{ik} = 1$$

$$u_r, v_i \ge 0, r = 1, \dots, s, i = 1, \dots, m$$

The dual form of (2) is:

$$minimise: h_k = \theta - \varepsilon \left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+ \right) \quad (3)$$

$$subjective\ to:$$

$$\theta x_{ik} - \sum_{j=1}^{n} \lambda_j x_{ij} - s_i^- = 0, i = 1, \dots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{ij} - s_i^+ = y_{rk}, y = 1, \dots, s$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, s_i^- s_i^+ \ge 0, \qquad for\ i = 1, \dots, m; r = 1, \dots, s; j = 1, \dots, n$$

In equation (3), s_i^-, s_i^+ are two slack variables. Here, for a DMU, s_i^- measures the level of inefficiencies from the perspective of input (the distance between a DMU's input to the optimal level of input), while s_i^+ represents the distance between a DMU's output to the optimal output. Thus, if θ =1 and meanwhile $s_i^- = s_i^+ = 0$, the DMU_k is efficient (i.e., h_k^* =1[18]). Based on the above two constraints $\sum_{j=1}^{n} \lambda_j x_{ij} = \theta x_{ik} - s_i^-$ and $\sum_{j=1}^{n} \lambda_j y_{ij} = y_{rk} + s_i^+$, for an inefficient DMU, the following equations explain the two directions in which efficiency can be achieved:

$$\Delta x_{ik} = x_{ik} - \left(\theta^* x_k^i - s_i^{-*}\right), i = 1, \dots, m \quad (4)$$

$$\Delta y_{rk} = (y_{rk} + s_r^{+*}) - y_{rk}, r = 1, \dots, s$$

According to equation (4), for an inefficient DMU with an efficiency value lower than 1, there are two directions to improve its efficiency: decrease the level of input by Δx_{ik} or increase the level of output by Δy_{rk}.

The third step is to select input and output indicators for efficiency calculation. In line with previous studies (e.g. Arends 2017; Ding et al. 2018), the research questions of this doctoral research, and fully considering data accessibility, the following outputs and inputs indicators of healthcare services are selected. Output indicators include i) the number of inpatients treated per year; ii) the number of outpatients treated per year; iii) the number of operations performed a year; and iv) the number of health checks performed a year. These four indicators can be seen as major products of healthcare services and represent the direct outputs of the healthcare sector, which is consistent with the explanations of productive efficiency in the 3Es and IOO models – the key analytical frameworks of PSP explained in Chapter 2 (Boyne 2002). For input indicators, three indicators representing the overall level of resource inputs are

[18] The star mark indicates the optimal value in equivalent (1) and (3) (Bowlin 1998).

chosen: i) the number of public medical institutions[19]; ii) the number of professional staff in these medical institutions; and, iii) the number of patient beds in medical institutions. Additionally, as suggested by DEA researchers, the sum of input and output indicators for efficiency calculation should not be larger than the number of DMUs divided by 2 (Golany and Roll 1989) or 3 (Banker et al. 1989). There are 26 DMUs, 4 output indicators, and 3 input indicators in this research. Thus, this prerequisite is satisfied.

- **Data Envelopment Analysis: data collection and results**

In line with the above explanations, data capturing the input and output indicators over the period of 2006 to 2017 were collected to measure efficiency using the input-oriented BCC-DEA approach. To ensure reliability, data for the 3 resource input indicators and 4 output indicators come from the Health Statistical Yearbooks of China, in which all data were originally collected from the national and local healthcare agencies (the National and local HCs). Although one of the advantages of the DEA approaches is that indicators with different units can be used together to measure efficiency, researchers still advise normalising the data if the magnitude of some indicators is significantly larger than others (Sarkis 2007; Sueyoshi and Goto 2013). Thus, each input or output observation of a DMU should be divided by its average. Another important issue is about the timing of the data for efficiency calculation. Many previous studies calculate periodical DEA efficiencies separately

[19] Government-funded medical institutions, including general hospitals, public health institutions (e.g. specialised hospitals), community healthcare centres, and township hospitals/clinics.

and then combine DEA scores in all selected time periods for panel data analysis. However, this approach appears to be problematic, because in this case, a DMU's efficiency score in one time period is calculated by measuring the distance between its performance to the optimal performance in the same time period. While the level of optimal performance varies across the selected periods, thus an efficient DMU in one period may be inefficient in other periods. In other words, for all DMUs, if their efficiencies for each period are estimated separately, those inter-period efficiencies are not comparable and thus cannot be empirically analysed together. Therefore, to ensure comparability, this research follows Hauner and Kyobe (2010) and Alonso and Andrews (2019) to calculate the DEA efficiency scores of all DMUs between 2006 to 2017 at a time. However, when adopting the BCC model, several DMUs are fully efficient with a score of 1. Among these efficient DMUs, despite a few DMUs having the best performance in all possible input-output combinations, many of them are considered as efficient because some of their input-output combinations are located at the efficient frontier while others are not. In other words, it is problematic to claim that all "efficient" DMUs have the same performance in practice (Adler et al. 2003).

To re-rank those DMUs with an efficiency score of 1, Banker and Gifford (1988) and Banker et al. (1989) developed a super-efficiency DEA model which breaks the constraint that the efficiency value of a DMU cannot be larger than one[20] (Cook et al. 2009). In this case, the DMUs that are efficient in DEA models might have efficiency

[20] It is an approach developed based on the conventional DEA technique, but can measure the extent to which an efficient DMU can further improve its input and/or reduce its output without becoming inefficient (Banker et al. 1989).

values larger or smaller than 1 (Lin and Chen 2018). However, the early-stage super-efficiency DEA models can only be applied when the DMUs are assumed to have constant returns to scale (CRS), and the inputs (outputs) are changed simultaneously and following a proportional (radial) function (Chen and Sherman 2004; Chen 2005). If the above two conditions are not satisfied, the linear programme process of some DMUs might be infeasible. To avoid this problem and in line with the VRS assumption, a non-radial VRS super-efficiency model developed by Tone (2001) is chosen for this study[21]. For this model, a higher super-efficiency DEA score represents a greater level of efficiency, which is in line with the BCC-DEA scores. Variations of the nationwide average BCC-DEA and super-efficiency DEA scores from 2006 to 2017 are shown in Figure 5.5, in which an upward trajectory is clearly illustrated, and the yearly average super-efficiency DEA scores are significantly lower than those calculated by the BCC-DEA approach. The provincial cross-year average BCC-DEA and super-efficiency DEA scores are mapped in Figures 5.6 and 5.7. Summary statistics of efficiency values by province are given in the Appendix.

[21] For detailed specifications of this model see Cooper et al. (2007 pp. 317-319). Super-efficiency calculations are conducted by DEAFrontier developed by Professor Joe Zhu of the Worcester Polytechnic Institute (Zhu 2014, pp. 399-407).

Figure 5.5: Average BCC-DEA and super-efficiency DEA scores over 2006 to 2017

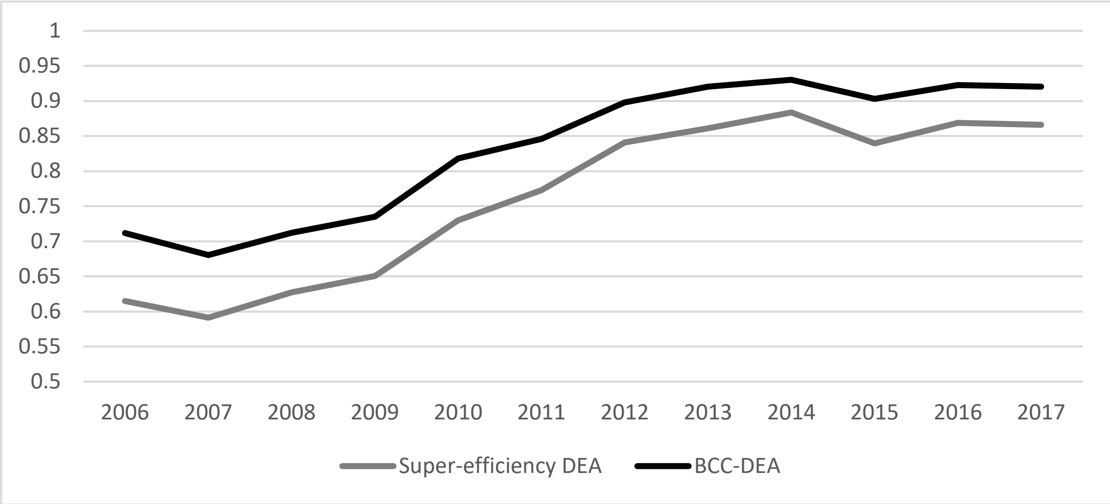
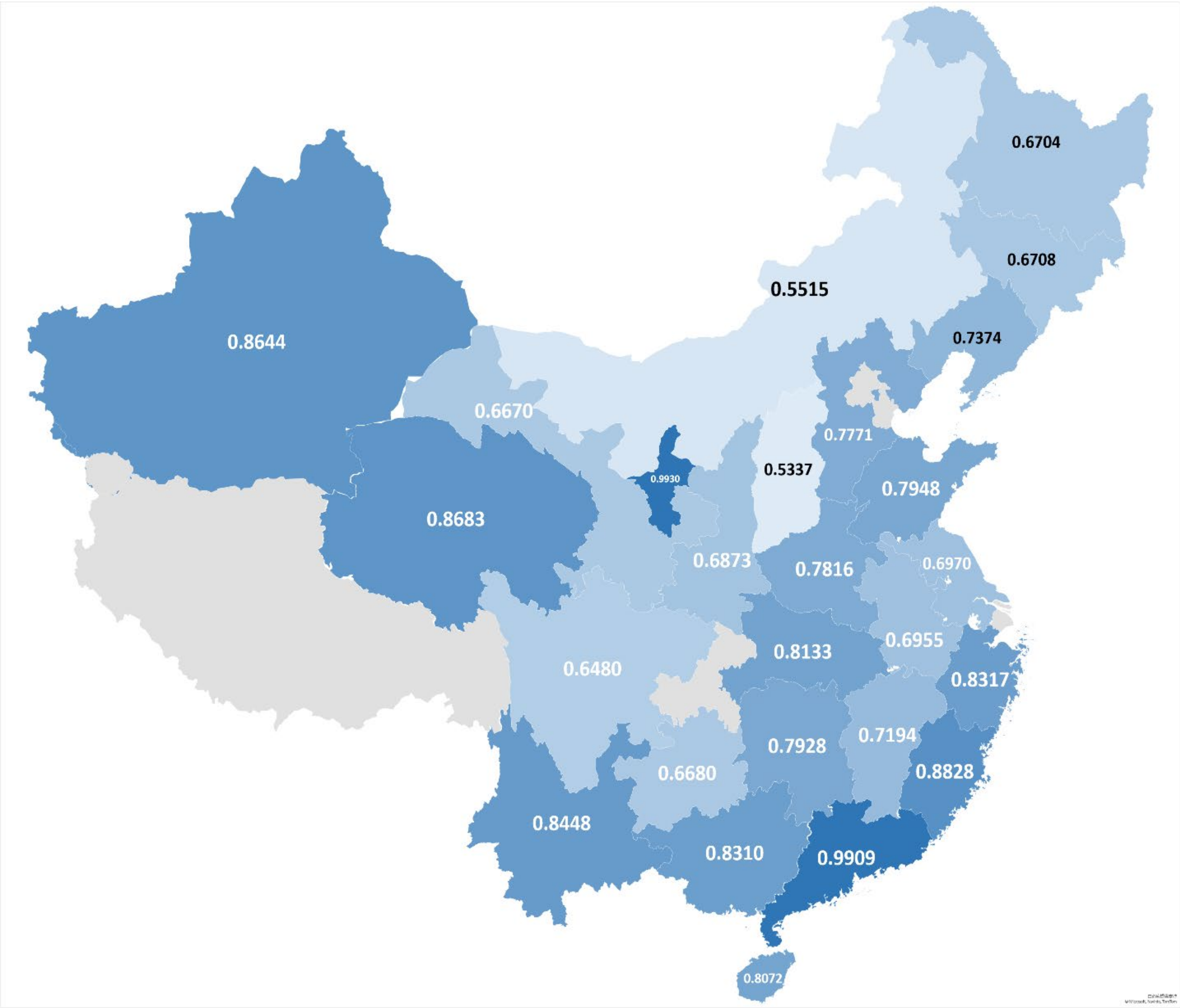


Figure 5.6: Map of average BCC-DEA scores from 2006 to 2017

Figure 5.7: Map of average super-efficiency DEA scores from 2006 to 2017



5.4.2. Measuring healthcare effectiveness

As mentioned in the Literature Review chapter, effectiveness refers to the extent to which the formal objectives of a public service have been achieved. Internationally, researchers usually rely on two quantitative measures to proxy for healthcare effectiveness, namely mortality rate (particularly IMR) and life expectancy (Kang et al. 2012). In empirical studies, IMR is more popular. On the one hand, IMR directly measures maternal and infant health status, and reducing IMR is undoubtedly a formal objective of healthcare service provision. Moreover, IMR is also a good proxy for the quantity and quality of other healthcare services in an area. Thus, researchers widely consider IMR as the most exhaustive indicator that represents not only the effectiveness of maternal and infant care but also the overall effectiveness of healthcare services in society (Young 2005). From another perspective, compared with LE, IMR is more sensitive to policy changes in healthcare services (Habibi et al. 2003; Asfaw et al. 2007). Also, compared with IMR, life expectancy is usually estimated according to the assumed lifespan in an area and thus is more easily manipulated (Jiménez-Rubio and Smith 2005). However, in China, there is a lack of yearly data on LE and IMR for each province over the selected research period. Thus, this research employs the perinatal mortality rate (PMR), an alternative to IMR and also a key indicator of maternal and infant care (Cartlidge and Stewart 1995; Richardus et al. 1998; Vogel et al. 2014), to gauge healthcare service effectiveness. Obviously, a higher PMR represents a lower level of healthcare effectiveness. This is different from the efficiency indicators where higher DEA and SDEA values directly proxy for a higher level of efficiency. Therefore, to ensure consistency and clarity when presenting empirical results, this research develops another indicator named perinatal survival rate (PSR) by subtracting the PMR value from 100[22] for empirical analysis.

Following international standards, PMR in China measures the number of stillbirths that occur after the 28th week of pregnancy plus the number of infants that die in the

[22] This transformation does not affect regression results, as PSR values are still directly calculated based on PMR.

first seven days per 100 total births. All PMR data for the period of 2006 to 2017 were collected from China's Health Statistical Yearbooks to ensure reliability. These data supported this research to calculate PSR values. As shown in Figure 5.8, from 2006 to 2017 the nationwide average PSR significantly increased from 98.9 to 99.5. The yearly averages of PSR for the selected provinces are mapped in Figure 5.9, showing that the west and north part of China, in general, has relatively lower PSRs. Also, summary statistics of the PSR data by province are given in Appendix.

Figure 5.8: Average PSR from 2006 to 2017

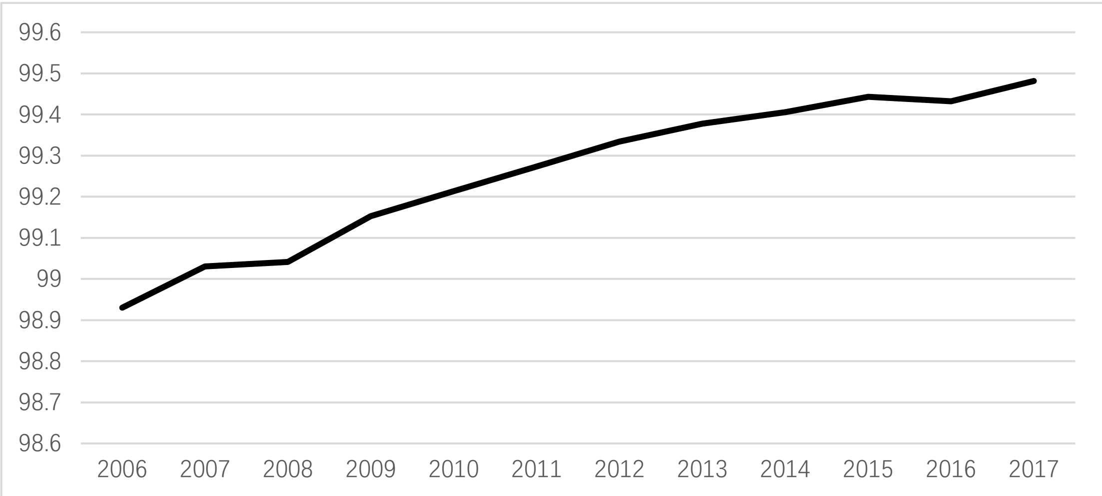


Figure 5.9: Map of average PSR from 2006 to 2017

In the above two sub-sections (5.2.2 and 5.2.3), the measures of healthcare service efficiency and effectiveness (the two dependent variables) and FD (the key independent variable) are presented and explained in detail. To identify the relationship between the levels of decentralisation and healthcare service efficiency/effectiveness, this study moves to the next part where relevant hypotheses and variables are proposed for empirical analysis.

5.5 Quantitative models and variables

As a study supported by panel data, the initial model applied for both the case of healthcare efficiency and effectiveness, is as follows:

$$PSP_{it} = \beta FD_{it} + \gamma Z_{it} + \varphi_i + \mu_t + \varepsilon_{it}$$

where PSP_{it}, respectively, refers to the healthcare service efficiency and effectiveness of province i in year t. As introduced in the above section, healthcare efficiency is calculated by the BCC-DEA approach and the super-efficiency DEA approach, while effectiveness is measured by PSR. FD_{it} refers to the degree of fiscal decentralisation within a province – measured by three indicators *HED*, *TED*, and *RD*. Z_{it} denotes a set of control variables. φ_i *and* μ_t refer to the time-specific effects and province-specific effects, respectively. ε_{it} is the disturbance term. In the following paragraphs, FD – the core explanatory variable and its hypothesised association with healthcare efficiency and effectiveness, are discussed in detail. Then, this sub-section separately presents the control variables for the efficiency and effectiveness models, which leads to the discussion about estimation strategies in the next sub-section (5.2.5).

5.5.1 FD – the core independent variable

- **Effects of FD on healthcare performance in Chinese provinces**

Most theoretical and empirical studies suggest that FD can improve public service efficiency and effectiveness. Regarding theoretical arguments, theories of FF 1.0 assume that local bureaucrats always aim to maximise the benefits to residents in their jurisdiction (Samuelson 1954; Musgrave 1959; Arrow 1971). Thus, FD provides local governments with more opportunities to employ their information and technical advantages in public service provision, which increases service efficiency (Oates 1972; 1999; 2005). FF 2.0 theories suggest that bureaucrats are in fact self-interested and identify two major channels between FD and better public service efficiency and effectiveness (Brennan and Buchanan 1980; Qian and Weingast 1997; Baskaran 2010). The first channel is ballot voting which, in a democratic background, supports local residents to vote out incompetent politicians (Tiebout 1956). With FD, local politicians play a greater role in local service provision. Thus, they have to fully utilise their information and technical advantages to improve public service efficiency and effectiveness, so as to win elections and maximise their own benefits (Persson and Tabellini 2000; Besley and Coate 2003). The second channel is the mechanism of resident sorting (Tiebout 1956; Cho 2009; De Mello 2011), which refers to people migrating to jurisdictions where public goods satisfy their tastes. In this case, a decentralised local government tends to improve PSP because it is motivated to attract population inflow, thereby bringing greater benefits to local bureaucrats.

However, with no multi-party and competitive elections, the first mechanism mentioned above, i.e., voting with ballots, appears to be inapplicable in Chinese provinces (Liu et al. 2006; Cho 2009). Nevertheless, supported by the mechanism of upward accountability, FD is still expected to benefit healthcare service performance. As explained in the Context chapter, each level of the government in China are accountable to the government that is administratively located one level above (Tsui and Wang 2004; Lam 2010). The relationship of upward accountability is ensured by two policy tools. The first tool is the TRS. It delivers general agendas and performance targets step-by-step to lower-level governments by whom detailed local policies are formulated and implemented with full consideration of local situations (Edin 2003; Gao 2009). The second tool is the top-down PMS, which imposes high-powered incentives on leaders of the lower-level government to ensure their fulfilment of pre-set policies and performance targets (Liang and Langbein 2015; Ma 2016). Moreover, the relative performance of local leaders in terms of inter-jurisdictionally shared policies and missions (e.g. economic development and the provision of basic public services) is also a key consideration in performance evaluation. That is, the performance of cadres in other regions with similar socioeconomic backgrounds is considered as a yardstick, leading to top-down yardstick competition (Caldeira 2012). Thus, local cadres have to not only fulfil pre-set policies and targets but also perform better than their counterparts in other (usually neighbouring) jurisdictions (Tian et al. 2020), so as to be fully accountable to the expectations of their superiors.

The above arrangements of upward accountability also exist in Chinese provinces. As illustrated by Figure 4.1 (see the Context chapter), a functional agency which provides local services, such as a prefecture-level city's HC, is primarily accountable to the prefecture-level city's government. This is ensured by the city's government's control of city HC's funding as well as the performance evaluation, rewards, and punishments of city HC's major leaders (Ma 2017a). Then, a prefecture-level city's government is accountable to the provincial government, which in turn, is accountable to the central government. With upward accountability, FD will not lead lower-level governments and functional agencies to deviate from the higher-level government, instead, they have more space and motivation to satisfy the key expectations of their superiors in their own way (Chien 2010).

From the 1980s to the early 2000s, the economic miracle in China fulfilled people's enduring demand for basic needs and getting rid of absolute poverty, which effectively strengthened the performance legitimacy of the CPC and the CPC-led government (Duckett and Wang 2017; Duckett and Munro 2022). However, from the late 2000s, the gradual weakening of economic growth, as well as the problems accompanying economic development such as corruption, regional inequality, and the degradation of public services, started to erode economy-based performance legitimacy. Thus, recognising people's growing needs for better public services including healthcare and the difficulty of maintaining high-speed economic growth, the Chinese top party and state leadership has been placing greater weight on healthcare in their key policy agendas to achieve "good governance" and regain performance legitimacy (Zhu 2011; Li 2011; Zeng 2014; Dickson et al. 2016; Duckett and Wang 2017; Zhang 2020).

Accordingly, the central government and the central committee of the CPC started to pay greater attention to healthcare. In 2009, "new healthcare reform" was initiated, which proposed clearer and higher requirements for performance-related issues such as improving the accessibility of healthcare services (especially basic and primary healthcare services) and promoting the oversight and professional management of healthcare institutions (Yip and Hsiao 2015; Then and Sussmuth-Dyckerhoff 2017). To a large extent, the new healthcare reform focusing on the above areas reflected the core values of "harmonious society[23]" and "Scientific Outlook on Development" – two top policy agendas of the CPC and PRC leader Hu Jintao at that time (Ngok and Zhu 2010; Wong 2010). That is, building a harmonious society relies on a scientific pattern of development, which should not only revolve around the economy but also focus more on people's well-being in various aspects such as health, education, housing, and public safety (Hu 2007; Geis and Holt 2009).

In 2012, Xi Jinping succeeded Hu Jintao and became the leader of the PRC and CPC. Under the new leadership, healthcare has been consistently perceived as one of the PRC and CPC's priorities. As stressed by Xi (2017), the "Chinese Dream" – Xi's political manifesto cannot be realised without satisfying Chinese people's pursuit of a better life. Healthcare services, undoubtedly, are associated with people's life and well-being. Also, Xi's other top political agenda named "Modernising Governance System and Capacity" clearly justifies the necessity of improving healthcare service performance for all people (Jing et al. 2015). Following Xi's emphasis on healthcare,

---

[23] Harmony is an important concept in Chinese culture. Traditionally, Confucianism argues that harmony should be a principle for a person, family, state, and the world. At the state level, Confucian philosophers explain that harmony refers to "governance with virtue" (Confucius 1999, Analects 2.1). It requires rulers to balance people's demands in all categories, improve people's welfare, and care for the weak (Li 2008). The traditional idea of harmony was used by the CPC to develop its key agenda, the "harmonious society" (Geis and Holt 2009). To a large extent, this agenda is a response to the challenges caused by the excessive emphasis on economic development from the 1970s to the early 2000s (Central Committee of the CPC 2006). Accordingly, satisfying people's needs in areas other than economic development, such as healthcare, education, and social security, is located at the centre of the harmonious society agenda (Yu 2008; Geis and Holt 2009).

the Central Committee of the CPC and the central government introduced the "Healthy China 2030" plan, which prioritises health as one of the core socioeconomic issues and provides a broad guide for local governments and society to achieve "health for all, and all for health" by 2030 (Central Committee of the CPC and State Council 2016; Tan et al. 2017; Tan et al. 2018).

The above description suggests that over the past decade, China's top leadership has placed an increasing emphasis on healthcare and healthcare service performance. This emphasis, supported by TRS and top-down PMS, has been transmitted to Chinese provinces and eventually to the sub-provincial governments and HCs – the providers of local healthcare services (Ratigan 2022). Previous studies have widely identified that within Chinese provinces, in line with the requirement of superiors, the provision and performance of public services including healthcare has been given greater weight in target setting and performance evaluation (Chien 2010; Zhou 2010; Zhang 2020). In this case, a greater level of FD – whether on the expenditure side (HED and TED) or the revenue side (RD) – will supply sub-provincial governments and HCs with more responsibilities and fiscal power in managing their jurisdictions. Accordingly, to be fully accountable to the requirements of their superiors, they have greater motivation and higher capacity to promote healthcare service efficiency and effectiveness through various ways such as initiating localised policies, innovations, and policy learning (Yu et al. 2016; Zhu 2017).

In addition to upward accountability, residential sorting – the second channel argued by theoretical studies that makes decentralisation work – might be generalisable to Chinese provinces as well. This is because, since the mid-1980s, the relaxation of the household registration system (the Hukou system) has created greater opportunities for inter-jurisdictional migration (Chan and Buckingham 2008). Li and Li (2015) and Zhang et al. (2017) suggest that the performance of local public services is a significant factor that attracts inter-regional migrants. However, only those families with better economic backgrounds are able to move among different jurisdictions due to the high migration costs. Thus, it is possible that under a fiscally decentralised context, sub-provincial governments in China are motivated to attract wealthy households by promoting the efficiency and effectiveness of local public services, so as to facilitate social-economic growth with the support of those newcomers (Cai and Wang 2008; Chan 2012; Shen and Li 2020). Moreover, as argued by Wang et al. (2018), residents of Chinese counties tend to compare the healthcare performance of their hometown with that of the adjacent regions. These comparisons affect their satisfaction with local healthcare services, which in turn, might impact their residential sorting actions. In this case, decentralised sub-provincial governments may be motivated to perform better than their counterparts in healthcare performance, so as to win the yardstick competition for population (Salmon 1987; Baicker 2001; Boadway and Tremblay 2012).

In summary, the above explanations suggest that within Chinese provinces, a great level of FD, supported by the mechanisms of upward accountability and residential sorting, is likely to promote the efficiency and effectiveness of healthcare services in Chinese provinces. Thus, the following hypotheses are given:

H1a: Intra-provincial FD will be positively related to healthcare efficiency.

H1b: Intra-provincial FD will be positively related to healthcare effectiveness.

- **The moderating effects of relative wealth on the FD-PSP relationship**

To provide a deeper insight into FD theories, this study investigates whether relative wealth influences the relationship between FD and healthcare performance. Following previous studies (e.g. Adam et al. 2014; Brock et al. 2015; Arends 2017), relative wealth is measured by the natural logarithm form of per capita local GDP.

Researchers usually suggest that the growth of wealth can promote healthcare service performance. Because in wealthier regions, the healthcare sector is likely to have stronger human, technical, managerial, and organisational advantages, which ensure services can be more efficiently and effectively produced (Hillestad et al. 2005; Burns et al. 2012). Moreover, researchers suggest that the governments and citizens in a wealthier territory will have greater financial resources to afford healthcare costs, which may increase service demands (Qian et al. 2009; Dou et al. 2018), leading to the achievement of scale economies (Evans et al. 2001; Allin et al. 2015; Zheng et al. 2015). In wealthy communities, individuals and families may not only have better healthcare status but also have more time and money to contribute to the development of local services than their poorer counterparts (Saich 2000; Jiménez-Rubio 2011a).

Moreover, relative wealth might also have a major influence on the relationship between FD and healthcare efficiency and effectiveness. In other words, relative wealth may perform as a "moderator" to strengthen the hypothesised positive FD-efficiency/effectiveness relationships. As explained above, FD provides self-interested cadres in China with greater motivations and more opportunities to improve healthcare service performance, so as to grasp personal benefits such as promotion opportunities (Caldeira 2010; Zhang et al. 2017). Nevertheless, the aforementioned impact of FD on performance is also likely to be determined by the capacity of local governments in service provision. As explained by Ingraham and Donahue (2000, p. 294), government capacity refers to its ability to "marshal, develop, direct, and control its human, physical and information capital to support the discharge of its policy directions". Relative wealth, measured by per capita GDP, represents the resources that can be employed for state-building and is therefore also an indicator of government capacity (Batley and Larbi 2006; Ding et al. 2018). As Prud'homme (1995) and Arends (2017) suggested, compared with the central government, local governments in less-wealthy areas may have disadvantages in terms of organisational, managerial, and technical capacity. Such gaps in capacity could weaken the positive impact of FD on efficiency, even turning the relationship into a negative direction. In other words, it means that with the growth of relative wealth, local governments are more likely to have strong organisational, managerial, and technical skills. This, in turn, provides them with a higher capacity to coordinate public services efficiently and effectively (Batley and Larbi 2006). Additionally, in prosperous areas, residents may be better at articulating their demands and acting to support public services in meeting them (Frenkiel 2021), which strengthens the co-production capacity of local governments and thus ensures a greater level of performance improvement under a certain level of FD. In summary, the above two arguments suggest that relative wealth – measured by per capita GDP – may strengthen the positive impact of FD on healthcare efficiency and effectiveness. Thus, the following hypotheses are given:

H2a: Relative wealth will strengthen the positive effect of FD on healthcare efficiency.

H2b: Relative wealth will strengthen the positive effect of FD on healthcare effectiveness.

5.5.2 Control variables

- **Control variables for the efficiency models**

Following previous studies, five demographic variables are controlled in the efficiency models. The first one is the population size, for which the natural logarithm form is usually adopted to avoid the problems of high skewness and show the proportional effect (Feng 2014; Curran-Everett 2018). As Boyne (1995) and Otsuka et al. (2014) suggested, a larger population size proxies for a greater demand for public services and a larger public sector with stronger purchasing power, which may lead to better public service efficiency. However, empirical findings by Stastná and Gregor (2015) indicate that a larger public sector does not necessarily have a greater capacity. Thus, a given level of resource input, no matter how strong the purchasing power, might fail to bring a greater output. Such a negative relationship between population and efficiency was also found in the Chinese-based study of Ding et al. (2018). Thus, a negative relationship between population and healthcare efficiency might still exist in this study.

Then, to account for scale economies and urban advantages in accessing healthcare services (Rayp and De Sijpe 2007), population density (population per square km) and the ratio of urban population to a province's total population are controlled. In a jurisdiction with higher population density and a greater proportion of urban residents, the provision of public services can be concentrated, which may reduce healthcare inputs (e.g. fixed costs and transportation costs) while maintaining or increasing service outputs at given costs (Stevens 2005; Rayp and De Sijpe 2007). However, a higher population density may also lead to diseconomies of scope and reduce healthcare service efficiency, due to over-consumption of medical resources, which may increase operating difficulties and make it more difficult for citizens to monitor healthcare institutions (Grossman et al. 1999; Adam et al. 2014). Such a negative impact was identified by Kalb et al. (2011), Cruz and Marques (2004), and Stastná and Gregor (2015) in various contexts and also in the Chinese-based study by Ding et al. (2018). Thus, a negative relationship between population density/the ratio of urban population and efficiency might exist in this research.

The fourth demographic variable included in the models is the education level of the local population. Theoretical and empirical papers suggest that well-educated citizens, as voters, have stronger capabilities to monitor the usage of public resources, which contributes to a higher level of public service efficiency (De Borger et al. 1994;

Afonso and Aubyn 2006; Geys and Moesen 2008; Hauner and Kyobe 2010). In the Chinese context without voting and multi-party competition systems, the studies of Ding et al. (2018) and Liu et al. (2020) also argue that education is positively related to healthcare service efficiency. This is because well-educated citizens are more likely to have a healthy lifestyle and a greater consciousness to keep healthy, which reduces unnecessary medical demands and ensures the effective utilisation of healthcare resources. Following prior studies (e.g. De Borger et al. 1994; Widstrom et al. 2004; Afonso and Aubyn 2006; Afonso and Fernandes 2008; Geys and Moesen 2008; Štastná and Gregor 2015), this research also expects a positive relationship between education and healthcare service efficiency and measures this variable as the percentage of inhabitants holding higher education certificates (bachelor, master, or PhD degrees, see Afonso and Fernandes 2008).

The proportion of older people in the population is the fifth demographic variable included in the models. This is measured as the percentage of inhabitants over 65 years old. Guo et al. (2017) suggest that the total demand for healthcare services is positively related to this variable. With the increase in this ratio, more healthcare investments and better levels of productive efficiency are needed to cope with the growing demand (Ding et al. 2018; Liu et al. 2020c). Such a positive relationship was empirically identified by Hauner and Kyobe (2010) and Arcelus et al. (2015). However, if the local healthcare department fails to increase service capacity in a timely manner, the greater medical demand will create more challenges for healthcare staff and the managerial team, leading to a lower level of healthcare efficiency. In empirical studies, a negative relationship between the elder ratio and efficiency was found by Widmer and Zweifel (2012) and Cruz and Marques (2014). Given the difficulties for older people to articulate their demands to healthcare providers with social media, it is expected that in the Chinese background, a higher proportion of older residents might also fail to contribute to better healthcare service efficiency.

Next, socioeconomic circumstances are controlled using the unemployment rate for provinces, which not only captures socioeconomic disadvantages but, similar to education, can also be regarded as a proxy for citizens' health status (Mosca 2007). A higher unemployment rate signifies economic depression and thereby people's greater demand for healthcare services, which may reduce healthcare organisations' capacity to provide services in cost-efficient ways (Geys and Moesen 2008) – leading to lower healthcare efficiency. Such a negative relationship was empirically identified by Loikkanen and Susiluoto (2005), Kalb et al. (2011), and Arends (2017). However, on the other hand, with the rise in the unemployment rate and the deterioration of people's health status, more expenditures might be allocated to the healthcare sector to deal with the growing demands for efficient healthcare services (Arends 2017). Furthermore, unemployment, as a symbol of economic stagnation, might result in fiscal tightening. As a response, the local governments have to improve cost-efficiency at a given level or lower level of inputs. Thus, a higher unemployment rate might also lead to better healthcare service efficiency in the Chinese background.

After that, two fiscal variables are controlled: the per capita healthcare expenditure in a provincial jurisdiction (in a natural logarithm form) and the relative *fiscal solvency*. For the first one, researchers claim that the per capita healthcare expenditure proxies for the overall capability of the local healthcare sector and the quality of local healthcare infrastructure, which are both expected to be positively related to efficiency (Evans et al. 2001; Varabyova and Schreyögg 2013). The relative *fiscal solvency* of sub-provincial governments in a province is measured as the ratio of all sub-provincial governments' own revenues to their expenditures. Greater fiscal solvency means that sub-provincial governments rely less on external funding. Some studies (e.g. Widmer and Zweifel 2012; Alonso and Andrews 2019) consider this indicator as a proxy for FD, but here, it is treated as a control variable. Because this research specifically focuses on the FD between the provincial-level government and other lower-level governments within a province, while the fiscal solvency ratio measures the reliance on money from all external resources (including but not limited to the provincial-level government, such as the central government and local government debts). Theoretical papers (Davis and Hayes 1993; Hines and Thaler 1995) and most empirical papers (e.g. Grossman et al. 1999) suggest that local taxpayers have greater motivation to monitor the usage of public funds if more money is collected locally. Thus, it is expected that fiscal solvency is positively related to healthcare efficiency (Balaguer-Coll et al. 2010).

Finally, government fragmentation is controlled. In particular, empirical researchers focus more on the impact of horizontal fragmentation on government performance within a region[24] (e.g. Dowding and Mergoupis 2003; Hendrick et al. 2011; Goodman 2015). These studies are based on Tiebout's (1956) theory that a higher level of horizontal fragmentation brings a greater level of mobility and more choices for moving in/out (i.e., voting with feet), which motivates local bureaucrats to compete for population inflow by improving service performance with lower costs. With the relaxation of the household registration system ("hukou") over the past decades, the above impact of horizontal fragmentation on PSP might also exist in the Chinese background. Thus, to grasp the impact of horizontal fragmentation on healthcare efficiency within Chinese provinces, the number of prefectural governments in a province per 100,000 population is controlled (see Hendrick et al. 2011), and following Tiebout's (1956) argument, a positive relationship between this variable and healthcare efficiency/effectiveness is expected.

- **Control variables for the effectiveness models**

Following previous studies on healthcare effectiveness, variables reflecting socioeconomic circumstances are controlled in the effectiveness models, many of which are included in the efficiency models as well. The first one is population size. Some papers argue that larger population size is related to better healthcare service effectiveness (e.g. Soto et al. 2012; Rocha et al. 2016). As mentioned above, population size proxies for the size of the public sector (Boyne 1995; Otsuka et al.

[24] Government fragmentation can be classified into two types: horizontal fragmentation which refers to the number of jurisdictions at the same tier of a local government and vertical fragmentation which refers to the number of administrative tiers in a jurisdiction (Goodman 2019).

2014). However, the positive association between public sector size and capacity cannot be taken for granted. Thus, it is still possible that a larger population leads to a lower level of healthcare effectiveness.

Then, following the empirical studies of Andrews and Martin (2010), Kang et al. (2012), Arends (2017), and the Chinese-based study of Uchimura and Jutting (2009), this study also controls population density and urbanisation in effectiveness models. On the one hand, a higher level of population density and urbanisation might contribute to better healthcare service effectiveness. The achievement of scale economies and better infrastructure in urban areas will not only lead to better service efficiency, but also other aspects of healthcare service performance such as effectiveness might be improved (Dollery and Fleming 2006). On the other hand, with a greater level of population density and urbanisation, high usage of healthcare resources in an area might also bring excessive burdens to the healthcare sector, leading to diseconomies of scale and worse healthcare effectiveness.

Fourth, citizens' education level, measured by the percentage of inhabitants holding higher education certificates, was also controlled in the effectiveness models. Empirical studies by Arends (2017), Jiménez-Rubio and Smith (2005), and Jiménez-Rubio (2011a, 2011b) identified a positive relationship between education and healthcare service effectiveness. As explained by Jiménez-Rubio (2011b), education is a proxy for social capital. Well-educated people usually have healthier lifestyles and better living conditions, ensuring a better health status. Thus, citizens' overall level of education is expected to be positively associated with healthcare service effectiveness.

Fifth, birth rate is controlled. The empirical works of Uchimura and Jutting (2009), Jin and Sun (2011), and Kang et al. (2012) suggest that birth rate is positively related to the IMR because a higher birth rate brings excessive burdens to families and healthcare institutions (Kaplan et al. 2015). In the Chinese background, an area with a high birth rate is usually socioeconomically underdeveloped with a high infant (perinatal) mortality rate (Narayan and Peng 2006; Li and Zhang 2007). Thus, the birth rate, which is measured by the ratio of new-borns to a provincial jurisdiction's population (times 1000%), is controlled, and a negative relationship between birth rate and healthcare effectiveness (i.e., PSR, perinatal survival rate) is expected.

Then, to control for the impact of government inputs on healthcare service effectiveness, three indicators measuring resource inputs for public healthcare institutions and the total fiscal input for healthcare services (per capita) are included in the effectiveness models. Following previous studies of Jiménez-Rubio and Smith (2005), Cantarero and Pascual (2008), Jin and Sun (2006), Brock et al. (2015), Jiménez-Rubio (2011a, 2011b), and Rocha et al. (2016), the three indicators are i) the number of public medical institutions per 1,000 people; ii) the number of hospital beds in medical institutions per 1,000 people; iii) the total healthcare expenditures (per capita, measured in a natural logarithm form) in a provincial jurisdiction, and a positive relationship between these input variables and healthcare service effectiveness is expected. Finally, fiscal solvency and fragmentation are also controlled to account for the impact of fiscal self-reliance and population inflows on the effectiveness of healthcare services (Cavalieri and Ferrante 2016). In line with their relationships with healthcare efficiency, they are expected to be positively related to healthcare effectiveness.

5.6 Estimation strategies

As the units of research data sources, and variables have been clarified, this section moves on to explain the estimation strategies and techniques for regression analysis. For the efficiency models where the dependent variable is measured by the BCC-DEA approach, the fixed effects estimation is employed at first. For an individual-specific effects model using panel data:

$$y_{it} = x_{it}^{'}\beta + z_{i}^{'}\delta + \mu_i + \varepsilon_{it} \; (i = 1, \dots, n; t = 1, \dots, T) \qquad \text{Equation 1}$$

where x_{it} can include variables that change across time and individuals, and z_i refers to variables that are time-invariant. $\mu_i + \varepsilon_{it}$ are disturbance terms, where the unobservable μ_i is called the individual effect or individual heterogeneity (which is time-invariant as well). ε_{it} refers to the disturbance term that changes across time and individuals but is assumed to be uncorrelated with μ_i. If μ_i is uncorrelated with all explanatory variables (x_{it}, z_i), the model is called a random effects model. If not, the model is a fixed effects model. In general, social science usually favours the fixed effects analysis because μ_i is less likely to be uncorrelated with all explanatory variables (Baltagi 2008). To obtain consistent estimations of β using fixed-effect analysis, it is required to conduct a fixed effects transformation. To do so, the first step is to average equation over the time period 1, …, T to get the following cross-sectional equation:

$\overline{y}_i = \overline{x}_i^{'}\beta + z_i^{'}\delta + \mu_i + \overline{\varepsilon}_i$ Equation 2

Then, subtracting Equation 2 from Equation 1 for each time period yields the FE transformed equation

$y_{it} - \overline{y}_i = (x_{it} - \overline{x}_i)^{'}\beta + (\varepsilon_{it} - \overline{\varepsilon}_i)$

or

$\tilde{y}_{it} = \tilde{x}_{it}^{'}\beta + \tilde{\varepsilon}_{it}$ Equation 3

where $\tilde{y}_{it} \equiv y_{it} - \overline{y}_i, \tilde{x}_{it} \equiv x_{it} - \overline{x}_i, \tilde{\varepsilon}_{it} \equiv \varepsilon_{it} - \overline{\varepsilon}_i$.

In Equation 3, the time-invariant individual heterogeneity μ_i has been removed. As long as $\tilde{\varepsilon}_{it}$ is uncorrelated with $\tilde{x}_{it}$ – in line with one of the key pooled OLS assumptions, the OLS estimation of β will be consistent. In this case, the estimated coefficient $\hat{\beta}_{FE}$ is called the fixed effects estimator[25]. The above transformations illustrate that a fixed effects model can control for those omitted variables which are time-invariant but changing across individuals. Even though these missing variables might be correlated with explanatory variables, the regression results would not be undermined by those time-invariant omitted values (Khaleghian 2003; Verbeek 2004; Rocha et al. 2016). Given this advantage, this research employs fixed effects analysis for the efficiency models as well as the effectiveness models, in line with the empirical studies of Jiménez-Rubio and Smith (2005), Soto (2012), Cavalieri and

[25] For detailed explanations of the fixed effects transformation, see Wooldridge (2001, pp. 265-272).

Ferrante (2016), Rocha et al. (2016), and Arend (2017). Meanwhile, standard deviations are clustered at the provincial level, so as to avoid problems with serial correlations and heteroskedasticity (Bertrand et al. 2004). Furthermore, the year trend dummy *Time* which proxies for time fixed effects that are invariant across observations (but change over the years) is included. This ensures that the upward trend in the dependent variable (i.e., healthcare service efficiency), which may be attributable to unobservable effects such as national policies and technological progress (see Baltagi and Moscone 2010), could be controlled. In this way, not only the "individual-specific" fixed effects (which are invariant across years but change over individuals) but also the "time-specific" fixed effects are taken into account. Thus, the following two baseline efficiency models are established:

$$DEA_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Elderly_{it} + \gamma_8 Unemp_{it} + \gamma_9 Lnhexppc_{it} + \gamma_{10} Solvency_{it} + \gamma_{11} Fragmentation_{it} + \gamma_{12} time + \mu_i + \varepsilon_{it} \qquad \text{(Model 1)}$$

where DEA_{it} refers to efficiency for province *i* in year *t* calculated by the BCC-DEA approach. FD_{it} is the level of fiscal decentralisation within a province. Separately, it proxies for 3 FD measures: healthcare expenditure decentralisation (HED), total expenditure decentralisation (TED) and revenue decentralisation (RD). $Lngdppc_{it}$ is the per capita GDP (in a natural logarithm form). $Lnpop_{it}$ is the number of residents in a province (in a natural logarithm form). $Popden_{it}$ is the population density of a province. $Urban_{it}$ refers to the proportion of urban residents in the total population of a province. $Education_{it}$ measures the percentage of local residents holding bachelor degrees or above. $Elderly_{it}$ is the percentage of residents older than 65 in a province. $Unemp_{it}$ is the unemployment rate of a province. $Lnhexppc_{it}$ is the per capita healthcare expenditure of a province (in a natural logarithm form). $Solvency_{it}$ represents fiscal solvency, which is measured by all sub-provincial governments' fiscal revenue divided by their expenditure in a province. $Fragmentation_{it}$ is measured by the number of prefectural jurisdictions in a province per 100,000 population. $Time$ is included to capture the unobservable year trend that are fixed across individual observations in healthcare efficiency. μ_i refers to the time-invariant province-specific effects. Finally, ε_{it} is the disturbance term.

To test for the moderating effects of relative wealth on the relationship between FD and healthcare efficiency, interaction terms (*lngdppc*FD*, FD respectively refers to HED, TED, and RD) are separately included in Model 2. To avoid the serious multicollinearity problems caused by adding interaction terms (Smith and Sasaki 1979), when testing moderating effects, the FD indicators and lngdppc will be mean-centred (Balli and Sørensen 2013).

$$DEA_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Elderly_{it} + \gamma_8 Unemp_{it} + \gamma_9 Lnhexppc_{it} + \gamma_{10} Solvency_{it} + \gamma_{11} Fragmentation_{it} + \gamma_{12} Time + \gamma_{13} FD_{it} * Lngdppc_{it} + \mu_i + \varepsilon_{it} \qquad \text{(Model 2)}$$

Some researchers argue that when the data of the dependent variable is censored, the Tobit regression (Tobin 1958) may perform better in terms of obtaining consistent estimates (Chen et al. 2014; Samut and Cafri 2016). For this study, all BCC-DEA efficiency scores are located in an interval of 0 to 1 and thus can be considered as censored data. The specification of a panel-data censored Tobit regression model is as follows:

$$y_{it}^{*} = x_{it}^{'} + \mu_i + \varepsilon_{it}$$

and

$$y_{it} = \begin{cases} a, & y_{it}^{*} \leq a \\ y_{it}^{*}, & a < y_{it}^{*} < b \\ b, & y_{it}^{*} \geq b \end{cases}$$

where i represents the individuals and equals to 1, 2, …, N; t = 1, 2, …, representing the time period. μ_i refers to the time-invariant individual effect, and ε_{it} is the disturbance term. If μ_i is correlated with the regressors, the above equation is a fixed effects Tobit model. If not, it is a random-effects Tobit model (Smith and Brame 2003). However, for a fixed-effect Tobit model, the maximum likelihood estimator (MLE) is considered to be biased and inconsistent when the length of the panel data is small and fixed (i.e., the incidental parameter problem, Lancaster 2000; Greene 2004; Fernandez-Val and Weidner 2016). Although Monte Carlo methods can be employed to obtain consistent estimators in fixed-effect Tobit analysis, this approach cannot be easily applied using statistical software (Greene 2004). Thus, only the random-effects Tobit analysis is adopted in this research. For a random-effect Tobit model, the individual effect term μ is assumed to be independent of x_{it}, and both μ and ε follow normal distributions with mean = 0 and variance = σ_{μ}^{2}. The likelihood contribution of a single individual i is:

$$L_i = \int_{-\infty}^{\infty} \left\{ \prod_{t=1}^{T_i} \left[\Phi \left(\frac{a - x'_{it}\beta - \mu_i}{\sigma_\nu} \right) \right]^{I_{it}^a} \left[\Phi \left(\frac{x'_{it}\beta + \mu_i - b}{\sigma_\nu} \right) \right]^{I_{it}^b} \left[\frac{1}{\sigma_\nu} \phi \left(\frac{y_{it} - x'_{it}\beta - \mu_i}{\sigma_\nu} \right) \right]^{(1 - I_{it}^a - I_{it}^b)} \right\} \phi \left(\frac{\mu_i}{\sigma_\mu} \right) d\mu_i$$

The log-likelihood function is as follows (Bruno 2004):

$$\log L = \sum_{i=1}^{N} \log L_i$$

As explained by Greene (2012), the integrals of the log-likelihood function can be calculated using the Gauss-Hermite quadrature. Then, the log-likelihood function can be maximised by standard non-linear optimisation algorithms (Butler and Moffit 1982). Using the given variables, the following efficiency model for random effects Tobit regression is formulated:

$$DEA_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Elderly_{it} + \gamma_8 Unemp_{it} + \gamma_9 Lnhexppc_{it} + \gamma_{10} Solvency_{it} + \gamma_{11} Fragmentation_{it} + \gamma_{12} Region_i + \mu_i + \varepsilon_{it} \qquad \text{(Model 3)}$$

Accordingly, the model for moderating effects analysis is established as well:

$$DEA_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Elderly_{it} + \gamma_8 Unemp_{it} + \gamma_9 Lnhexppc_{it} + \gamma_{10} Solvency_{it} + \gamma_{11} Fragmentation_{it} + \gamma_{12} Region_i + \gamma_{13} FD_{it} * Lngdppc_{it} + \mu_i + \varepsilon_{it} \qquad \text{(Model 4)}$$

where the specifications of all dependent and independent various have been given above. Given the random features the Tobit models (Wooldridge 2001), $Time$, the variable representing the fixed effects of year trends (i.e., fixed across observations but changing over times), is not included in this random effects Tobit model. However, as a random effects model is able to estimate variables with time-invariant values, the dummy variable $Region$ is controlled, for which 1 is given to the four provincial autonomous regions for ethnic minorities (i.e., Xinjiang, Guangxi, Ningxia, and Inner Mongolia), and 0 is given to other provinces. This variable accounts for the potential effects of autonomous policies in these four regions.

Then, when healthcare efficiency is measured by the second approach, i.e., the super-efficiency DEA, fixed effects and random effects models are again employed, using the aforementioned control variables. Two fixed effects models for estimating the effects of FD on healthcare efficiency and the moderating effects of relative wealth on the FD-efficiency relationship are as follows:

$$SDEA_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Elderly_{it} + \gamma_8 Unemp_{it} + \gamma_9 Lnhexppc_{it} + \gamma_{10} Solvency_{it} + \gamma_{11} Fragmentation_{it} + \gamma_{12} Time + \mu_i + \varepsilon_{it} \quad \text{(Model 5)}$$

$$SDEA_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Elderly_{it} + \gamma_8 Unemp_{it} + \gamma_9 Lnhexppc_{it} + \gamma_{10} Solvency_{it} + \gamma_{11} Fragmentation_{it} + \gamma_{12} Time + \gamma_{13} FD_{it} * Lngdppc_{it} + \mu_i + \varepsilon_{it} \quad \text{(Model 6)}$$

Model 7 and Model 8 are then established to perform random effects estimations:

$$SDEA_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} +$$

$$\gamma_6 Education_{it} + \gamma_7 Elderly_{it} + \gamma_8 Unemp_{it} + \gamma_9 Lnhexppc_{it} + \gamma_{10} Solvency_{it} + \gamma_{11} Fragmentation_{it} + \gamma_{12} Region_i + \mu_i + \varepsilon_{it} \qquad \text{(Model 7)}$$

$$SDEA_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Elderly_{it} + \gamma_8 Unemp_{it} + \gamma_9 Lnhexppc_{it} + \gamma_{10} Solvency_{it} + \gamma_{11} Fragmentation_{it} + \gamma_{12} Region_i + \gamma_{13} FD_{it} * Lngdppc_{it} + \mu_i + \varepsilon_{it} \qquad \text{(Model 8)}$$

For the analysis of healthcare effectiveness, fixed effects and random effects models are again employed, using the control variables explained in the last section. Fixed effects models for FD-effectiveness analysis are as follows:

$$PSR_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Birthrate_{it} + \gamma_8 Hinstitutions_{it} + \gamma_9 Beds_{it} + \gamma_{10} Lnhexppc_{it} + \gamma_{11} Solvency_{it} + \gamma_{12} Fragmentation_{it} + \gamma_{13} Time + \mu_i + \varepsilon_{it} \qquad \text{(Model 9)}$$

$$PSR_{it} = \gamma_1 FD_{it} + \gamma_2 lngdppc_{it} + \gamma_3 lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Birthrate_{it} + \gamma_8 Hinstitutions_{it} + \gamma_9 Beds_{it} + \gamma_{10} lnhexppc_{it} + \gamma_{11} Solvency_{it} + \gamma_{12} Fragmentation_{it} + \gamma_{13} Time + \gamma_{14} FD_{it} * Lngdppc_{it} + \mu_i + \varepsilon_{it} \qquad \text{(Model 10)}$$

Finally, two models for random effects estimations are established:

$$PSR_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Birthrate_{it} + \gamma_8 Hinstitutions_{it} + \gamma_9 Beds_{it} + \gamma_{10} Lnhexppc_{it} + \gamma_{11} Solvency_{it} + \gamma_{12} Fragmentation_{it} + \gamma_{13} Region + \mu_i + \varepsilon_{it}$$

(Model 11)

$$PSR_{it} = \gamma_1 FD_{it} + \gamma_2 Lngdppc_{it} + \gamma_3 Lnpop_{it} + \gamma_4 Popden_{it} + \gamma_5 Urban_{it} + \gamma_6 Education_{it} + \gamma_7 Birthrate_{it} + \gamma_8 Hinstitutions_{it} + \gamma_9 Beds_{it} + \gamma_{10} Lnhexppc_{it} + \gamma_{11} Solvency_{it} + \gamma_{12} Fragmentation_{it} + \gamma_{13} Region + \gamma_{14} FD_{it} * Lngdppc_{it} + \mu_i + \varepsilon_{it}$$ (Model 12)

In the above four models, apart from those variables that have been explained above, PSR refers to perinatal survival rate which equals to 100 minus PMR (perinatal mortality rate); $Birthrate$ refers to the birth rate of a province; $Hinstitutions$ controls the number of public healthcare institutions per 1,000 person in a province; $Beds$ is the number of beds in public healthcare institutions per 1,000 person in a province.

5.7 Summary

This chapter provided a comprehensive review and explanation of the methodology employed in this research. The first section justified the positivist philosophical stance and the quantitative-dominant sequential mixed methods strategy. Then, the second section introduced the units of analysis and data sources, which is followed by the specifications of FD evaluations in the third section. After that, Section 5.4 explained how healthcare efficiency and effectiveness were measured. Subsequently, quantitative models, variables, and hypotheses were proposed in Section 5.5. Finally, Section 5.6 demonstrated the strategies for data estimation. In the next chapter, empirical findings of quantitative models will be reported.

CHAPTER 6 FINDINGS

In the Methodology chapter, hypotheses related to the impact of FD on healthcare service efficiency and effectiveness as well as the econometric models for empirical analysis have been explained in detail. In this chapter, findings from the efficiency and effectiveness models are separately reported in Section 6.1 and Section 6.2. In Section 6.1, summary statistics and the correlation matrix are presented first. As mentioned in the Methodology chapter, the dependent variable (efficiency) is measured using two approaches: 1) BCC-DEA and 2) super-efficiency DEA for which the names of the variable in the models are *DEA* and *SDEA*, respectively. Then, for each efficiency scenario, three baseline models and three models discussing the moderating effects of relative wealth on the FD-efficiency/effectiveness relationships are analysed, in which the key independent variables are HED, TED, and RD, respectively. After reporting summary statistics and the correlation matrix, empirical findings from the above regressions are reported step by step. Then, in Section 6.2, results in regard to FD and healthcare effectiveness are presented, following the same structure as the first section[26]. More detailed introductions to the structures are given at the beginning of the following sections.

[26] This thesis also tests for the non-linear relationships between FD and healthcare efficiency and between FD and healthcare effectiveness (see Appendix 8). There seems to be a U-shape FD-efficiency and FD-effectiveness relationship in some of the estimations. However, this result only applies to a few numbers (less than 10%) of the FD observations in the early years (before 2010), thus having no substantial impact on the main findings. For that reason, it is not included in the main analysis.

6.1 FD and healthcare efficiency

This section focuses on the impact of FD on healthcare efficiency. A report and brief discussions of summary statistics and the correlation matrix are given at the beginning. Then, this section reports the separate effects of FD (measured by HED, TED, and RD) on healthcare efficiency (measured by the BCC-DEA and super-efficiency DEA approach) in fixed effects, random effects, and Tobit regressions. After that, the moderating effects of relative wealth on the FD-efficiency relationship are investigated in the final part of this section.

6.1.1 Summary statistics and correlation analysis

Table 6.1 presents descriptive statistics for all variables used in the FD-efficiency analysis. As details about the FD (HED, TED, and RD) and efficiency indicators (DEA and SDEA) have been presented in the Methodology chapter, the following paragraphs mainly discuss the summary statistics of other variables.

First, for per capita GDP, the min value is 5750 (Guizhou province in Western China), and the max value is 107150 (Jiangsu province in Eastern China), showing a significant east-west regional gap in terms of socioeconomic development. Nevertheless, summary statistics of *Unemp* suggest that unemployment is not a serious issue in most Chinese provincial units from 2006 to 2017, with a maximum value of 5.1%. In fact, some underdeveloped provinces such as Guangxi, Hainan, and Gansu have the lowest unemployment rates because they are making greater efforts to narrow the economic gap with eastern provinces by attracting investments and creating job opportunities. Moreover, unemployment rates in China are all lower than those of the OECD countries, G7 countries, and the EU (27 nations) countries (OECD 2020). Full employment, along with the growing demand and investment, is widely considered to be the three key elements that have created the miracle of Chinese development over the past decades (Ghose 2008; Tsen 2010; Sharma and Sharma 2019).

There is a high degree of variation for the population data, as shown by the min value (5,480,000), max value (111,690,000), and standard deviation (26,907,798). This is also the case for *Popden* (population density) – min value (7.87), max value (753.19), and standard deviation (195.27). The data for *Urban* (%) highlight a relatively low level of urbanisation in most Chinese provinces: the min value, mean, and median value are 27.45, 49.84, and 49.70, respectively. Even the max value (69.85) is still lower than 70. While the small gap between the mean and median value (0.14) suggests that although the values of the data are dispersed (standard deviation = 9.28), the distribution of the data is roughly normal.

For the fourth demographic variable *Edu* (the percentage of residents holding bachelor degrees), the min and max values are equal to 2.72 and 19.83, respectively, with a relatively small standard deviation (3.46). However, the data show that only a limited proportion of citizens have higher-education experience: even the highest ratio is below 20%, with two-thirds of the observations being less than 10%. These data are in line with the low levels of socioeconomic development in some provinces, especially those located in the western part of China, such as Yunnan and Guizhou. Although the annual mean values of the *Elder* variable have an upward trend, the summary statistics show that only a small proportion of the population is older than 65 years old in most Chinese provincial jurisdictions (min = 5.47, max = 14.08, mean = 9.28, median = 9.11).

Regarding the fiscal variables, The min and max values of per capita healthcare expenditure (*lnhexppc*) show a huge gap in healthcare expenditure between different provinces over the period of 2006 to 2017, even though the mean and median values are close to each other as well. Then, *fiscal solvency* (%) (the ratio of sub-provincial governments' revenue to expenditure) has a wide range of values from 15.77 to 115.93. The mean value and the median value are close to each other, suggesting a relatively symmetric trend of data distribution. Moreover, summary statistics of *fiscal solvency* also indicate that most local governments in Chinese provinces cannot fully cover their expenditures using their own revenues. The mean and median values are smaller than 50%, showing a strong reliance on external financial support in most provincial jurisdictions.

In terms of the number of prefecture-level jurisdictions per 100,000 population in a province (*Fragmentation*), the min (0.02), max (0.15), mean (0.04), and median values (0.03) show that the population size of prefecture-level jurisdictions in China is relatively large – even the max value (0.15) means that there are around 660,000 residents in a prefecture-level jurisdiction. Moreover, the raw data on the number of prefecture-level jurisdictions show that most Chinese provinces are highly fragmented: provinces in China, on average, have 12.5 prefecture-level jurisdictions. While the median value is 13, showing a symmetric distribution of the observations. Finally, summary statistics for the dummy variable *Minority* are reported, which is in line with the fact that there are only 5 of the 31 provincial jurisdictions in mainland China are ethnic minority regions, and 4 of them are included in this study's sample group.

Table 6.1: Summary statistics of variables for FD-efficiency models

Variables	Obs.	Min	Max	Mean	Median	*Standard Deviation*
BCC-DEA	312	0.321	1.000	0.833	0.873	0.147
Super-efficiency DEA	312	0.263	1.068	0.762	0.785	0.172
HED	312	0.603	0.982	0.887	0.907	0.066
TED	312	0.518	0.935	0.798	0.812	0.080
RD	312	0.552	0.974	0.802	0.790	0.094
GDP per capita	312	5750	107150	35559.865	33554	18141.480
Unemp (%)	312	1.7	5.1	3.536	3.600	0.558
Population	312	5480000	111690000	48245224	43895000	26907798
Popden	312	7.866	753.189	273.988	236.953	195.269
Urbanisation (%)	312	27.453	69.854	49.838	49.700	9.283
Edu (%)	312	2.718	19.825	9.445	9.041	3.458
Elder (%)	312	5.473	14.076	9.278	9.111	1.726
Hexppc	312	47.669	2093.813	547.807	506.014	349.495
Fiscal solvency (%)	312	15.765	115.934	48.162	43.931	18.009
Fragmentation	312	0.015	0.146	0.035	0.026	0.026
No. of prefecture-level jurisdictions	312	2	21	12.544	13	4.174
Minority	312	0	1	0.154	0	0.361

Abbreviations: Unemp, unemployment rate (%); Popden, population density; Edu, education – percentage of residents with higher education certificates; Elder, percentage of residents over 65; Hexppc, healthcare spending per capita; Fragmentation, number of prefecture-level jurisdictions per 100,000 population; Minority, the dummy variable (1 for Provincial Autonomous Regions for ethnic minorities and 0 for others).

To show the potential relationship between the dependent and independent variables, the Pearson correlation test is conducted, as shown in Table 6.2. Most of the correlation coefficients between FD indicators and the efficiency indicators are positive and statistically significant, showing the potential relationship to be found from regression analysis. Also, the correlation between the two efficiency measures is high (coefficient = 0.85), highlighting that they capture a common construct. Additionally, variance inflation factor (VIF) values shown at the end of Table 6.2 models are smaller than 10, suggesting that multicollinearity is not a serious issue (Lind et al. 2010).

Table 6.2: Correlation Matrix for FD-efficiency models

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. DEA	1															
2. SDEA	0.9622*	1														
3. HED	0.2194*	0.2303*	1													
4. TED	0.2172*	0.1986*	0.7470*	1												
5. RD	0.2520*	0.2656*	0.5049*	0.6324*	1											
6. GDP per capita (ln)	0.5037*	0.5197*	0.4896*	0.5514*	0.4350*	1										
7. Population (ln)	-0.1046*	-0.0919	0.6712*	0.7164*	0.5078*	0.1605*	1									
8. Population density	0.1118*	0.1278*	0.3636*	0.6235*	0.5718*	0.3479*	0.6675*	1								
9. Urbanisation	0.3299*	0.3663*	0.2931*	0.4717*	0.3365*	0.8585*	0.1132*	0.3666*	1							
10. Education	0.4108*	0.4308*	0.3240*	0.2743*	0.2755*	0.7905*	-0.0845	0.0311	0.7113*	1						
11. Elderly	0.00970	-0.0517	0.4859*	0.5448*	0.3692*	0.3330*	0.5397*	0.3990*	0.2792*	0.2817*	1					
12. Unemployment	-0.3454*	-0.3394*	-0.2290*	-0.2326*	-0.1330*	-0.4014*	-0.0551	-0.3304*	-0.2995*	-0.2781*	0.0743	1				
13. Healthcare spending per capita (ln)	0.6133*	0.5984*	0.3661*	0.2130*	0.1108*	0.7267*	-0.1920*	-0.1192*	0.4396*	0.6885*	0.1240*	-0.3974*	1			
14. Fiscal solvency	0.0625	0.0773	0.2515*	0.5135*	0.6150*	0.4303*	0.5173*	0.7675*	0.5376*	0.1780*	0.2841*	-0.2511*	-0.1731*	1		
15. Fragmentation	0.1458*	0.1360*	-0.5720*	-0.6302*	-0.3779*	-0.1364*	-0.7947*	-0.5972*	-0.1627*	0.0746	-0.4939*	0.0827	-0.5198*	0.2056*	1	
16. Minority	0.0664	0.1179*	-0.1419*	-0.2276*	-0.0693	0.0034	-0.3663*	-0.4020*	-0.0800	0.1004*	-0.3414*	0.1062*	0.0858	-0.2142*	0.3250*	1
VIF (HED controlled)			3.98			8.97	6.03	4.33	4.65	4.10	2.12	1.75	5.33	4.54	3.23	1.48

* represents the 10% significance level

VIF values are from on the models where DEA is the dependent variable and HED is the key independent variable. For VIF values from the models controlling for TED and RD, see Appendix 7.

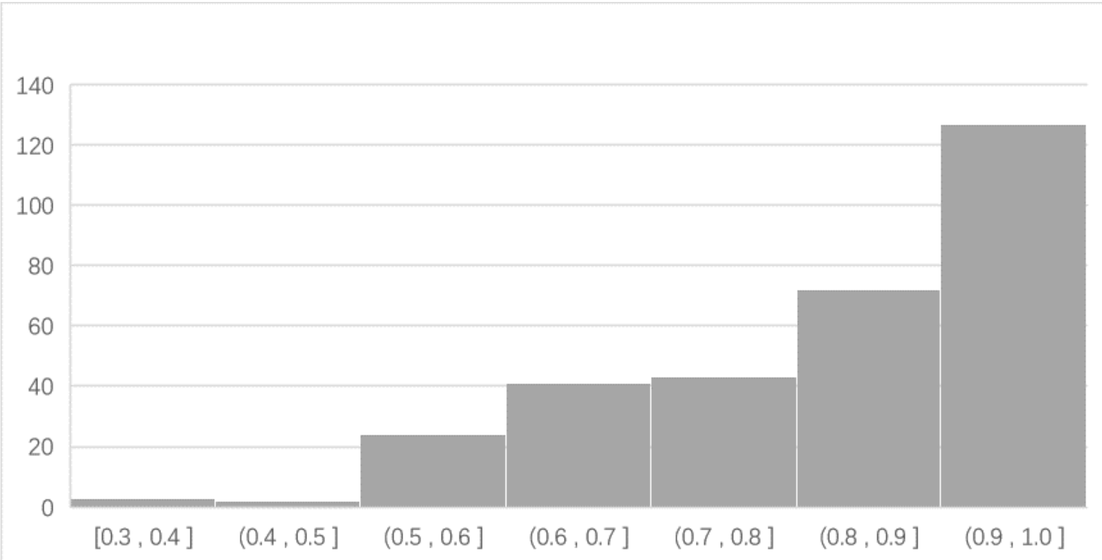
6.1.2 Regression results: FD and healthcare efficiency

This part reports the regression results from the models focusing on the impact of FD on healthcare service efficiency. As introduced above, this study calculates efficiency with two approaches, namely the BCC-DEA and super-efficiency DEA. For the key independent variable FD, there are three indicators: 1) HED, 2) TED, and 3) RD. Both the effect of FD on healthcare efficiency and the moderating effect of relative wealth on the FD-efficiency relationship are investigated using fixed effects regressions, random regressions, and Tobit regression. The following paragraph starts from the models looking into FD's effects on healthcare efficiency. Then, the moderating effects of relative wealth on the relationship between FD and efficiency are discussed.

6.1.2.1 Effects of FD on healthcare efficiency

This part starts from the effects of FD on BCC-DEA calculated efficiencies. Fixed effects regression is employed at first, for which results are shown in columns 1-3 of Table 6.3 (see page 191). It can be seen that all three FD indicators have a statistically significant and positive relationship with the BCC-DEA efficiencies, which is in line with Hypothesis 1 as well as previous studies focusing on FD and public service efficiency more generally (e.g. Barankay and Lockwood 2007; Sow and Razafimahefa 2015; Alonso and Andrews 2019). For the first FD indicator HED, the coefficient equals 0.504 and is statistically significant at the 1% significance level. Then, the coefficient of TED is 0.368 and is significant at 10%. Finally, RD has a coefficient equal to 0.564 and is significant at the 1% significance level. As values of the FD indicators are scaled between 0 and 1, the above coefficients mean that a 0.01 unit increase in these FD indicators can lead to efficiency gains of 0.00504, 0.00368, and 0.00564 units, respectively. To better illustrate the contributions of FD to efficiency, Figure 6.1 shows the distribution of the BCC-DEA values grouped in intervals of 0.1. It can be seen that a 0.01-unit increase in BCC-DEA will potentially result in the performance of a province being better than several of its counterparts. Detailed discussions about FD's impact on healthcare efficiency (and effectiveness) will be given in the next chapter.

Figure 6.1: Distribution of BCC-DEA scores



For other variables, first, the impact of *lngdppc* cannot be statistically confirmed. Then, the negative coefficients of *lnpop* (population) are statistically significant at the 10% significance level in one of the three models with FD measured by TED, showing that the size of the local population might bring little effect on healthcare efficiency in Chinese provinces. All other social and demographic variables and one of the fiscal variables, including population density, urbanisation, education, the percentage of elderly residents (*elderly*), unemployment rate, and the logged per capita healthcare expenditures (*lnhexppc*), do not have statistically significant impact. *Fiscal solvency*, which is the revenue-to-expenditure ratio of all sub-provincial jurisdictions in a province, is positively related to healthcare efficiency (at the 1% significance level) in two of the three models focusing on the impact of HED and TED. This finding suggests that a government relying more on its own funding tends to pay greater attention to the efficient usage of its revenues, which is in line with the expectation. Administrative fragmentation has no significant impact on healthcare efficiency, which possibly because the scale of intra-provincial immigration in China is smaller than that between different provinces (Zhao et al. 2018), and local governments in different jurisdictions have diversified natures (Dowding and Mergoupis 2003). Finally, the year trend variable (*time*) is statistically significant (at 5%) and positive in all three models. It suggests that the unobservable time fixed effects which are invariant over the observations (i.e., provinces) but changing over the years, such as national policies, have positive effects on healthcare efficiency.

Table 6.3: FD's effects on healthcare efficiency (BCC-DEA)

	BCC-DEA Fixed effects			BCC-DEA Tobit regression		
	(1)	(2)	(3)	(4)	(5)	(6)
HED	0.504***			0.551***		
	(0.160)			(0.161)		
TED		0.368*			0.402***	
		(0.192)			(0.151)	
RD			0.564***			0.570***
			(0.160)			(0.126)
Lngdppc	-0.117	-0.085	-0.029	-0.023	-0.003	0.047
	(0.092)	(0.088)	(0.086)	(0.060)	(0.060)	(0.058)
Lnpop	-1.398	-1.540*	-0.957	-0.091*	-0.081	-0.082
	(0.879)	(0.889)	(0.819)	(0.050)	(0.050)	(0.051)
Popden	-0.001	-0.000	-0.001	0.000	0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)
Urban	-0.000	-0.001	-0.001	-0.001	-0.002	-0.001
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
Education	0.005	0.006	0.006	0.009***	0.009***	0.009***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Elderly	-0.013	-0.012	-0.010	0.007	0.007	0.005
	(0.009)	(0.009)	(0.009)	(0.006)	(0.006)	(0.006)
Unemp	0.011	0.002	0.012	-0.011	-0.019	-0.016
	(0.021)	(0.021)	(0.020)	(0.017)	(0.017)	(0.017)
Lnhexppc	0.061	0.061	0.035	0.082***	0.086***	0.054*
	(0.043)	(0.041)	(0.037)	(0.027)	(0.028)	(0.028)
Solvency	0.003***	0.003***	0.001	0.002***	0.003***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Fragmentation	-3.208	-3.557	-2.887	0.014	-0.108	-0.653
	(5.025)	(5.369)	(4.982)	(1.202)	(1.207)	(1.219)
Minority regions				0.017	0.015	-0.012
				(0.066)	(0.066)	(0.069)
Time	0.029**	0.027**	0.024**			
	(0.011)	(0.012)	(0.011)			
Constant	25.652	27.900*	17.457	1.455	1.325	0.988
	(15.153)	(15.296)	(13.987)	(0.920)	(0.924)	(0.932)
R-sq within	0.70	0.69	0.71			
Pseudo R-sq				-0.32	-0.31	-0.34
F-test	29.19	28.94	41.68			
Prob > F	0.000	0.000	0.000			
Wald chi2				520.07	506.53	551.47
Prob > chi2				0.000***	0.000***	0.000***
LR test				200.92	192.61	206.81
Prob≥chi2				0.000***	0.000***	0.000***

N=312. Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. For Tobit regressions: uncensored observations: 276, left-censored: 0, right-censored: 36

Findings from the random-effect Tobit models are reported in columns 4 to 6 of Table 6.3. The Wald Chi-Square statistics suggest that all of the Tobit models are statistically significant at a 1% level. Also, according to the p-values of the Likelihood-Ratio (LR) test results, a random-effect Tobit analysis is statistically better than a pooled Tobit regression in all of the above cases. Regarding FD, all three indicators are found to be positively related to healthcare efficiency at the 1% significance level. The coefficients of HED, TED, and RD are 0.551, 0.402, and 0.570, respectively, which are largely consistent with the above key findings from fixed effects models (see columns 1 to 3 of Table 6.3), showing that a 0.01 units increase in the three FD indicators can respectively lead to efficiency gains of 0.0551, 0.0402, and 0.0570 units.

For the control variables in the Tobit models, first, the impact of *lngdppc* is statistically insignificant. Second, in line with findings from the fixed effects estimations, population size (*lnpop*) is negatively related to healthcare efficiency at the 10% significance level in the model with HED. Still, the impact of population density and urbanisation on efficiency cannot be statistically confirmed. Different from the fixed effects models which perform within-province estimations, in random-effect Tobit models, citizens' level of education (*edu*) shows a statistically significant and positive relationship with healthcare efficiency in all three models (significant at 1%). Provinces with more educated populations may perform better in healthcare service performance because well-educated citizens, in general, have a healthier lifestyle and a greater ability to use healthcare services efficiently (Hsu 2013; Ding et al. 2018). Then, there is no significant impact of *elderly* – the percentage of residents over 65 and the unemployment rate on healthcare efficiency. The relationship between per capita healthcare expenditure (in a natural logarithm form) and healthcare efficiency is positive and statistically significant in the Tobit models. This result could be explained by the greater infrastructure brought by a higher level of fiscal input (Evans et al. 2001; Varabyova and Schreyögg 2013). For fiscal solvency (*solvency*), consistent with the findings from the fixed effects models, the positive relationship is statistically confirmed at the 1% significance level in two of the three models focusing on HED and TED, showing the contribution of a higher level of fiscal self-reliance on the efficiency of local healthcare services. Then, the impact of administrative fragmentation still cannot be confirmed. Finally, the impact of dummy variable *minority regions* is also insignificant. These results reflect that the four provincial autonomous regions, in fact, hardly have greater autonomous power than ordinary provinces (Zhang 2012).

Turning to the impact of FD on healthcare efficiencies calculated by the super-efficiency DEA approach, highly consistent findings can be seen in Table 6.4 where fixed effects regression results and random effects GLS regression results are reported in columns 1 to 3 and columns 4 to 6, respectively. The Hausman test results favour fixed effects over random effects in all specifications. Nevertheless, in both fixed effects and random effects regressions, two of the three FD indicators, HED and RD, are statistically confirmed to be positively related to healthcare efficiencies at the 1% significance level. Regarding TED, its impact on healthcare efficiency cannot be statistically confirmed. This is probably because decentralising all responsibilities to the sub-provincial bureaucracy (measured by a higher TED) is relatively less relevant to the healthcare sector. Thus, the benefits of TED on healthcare efficiency might be less significant. Again, Figure 6.2 shows the distribution of the super-efficiency DEA scores.

Figure 6.2: Distribution of super-efficiency DEA scores

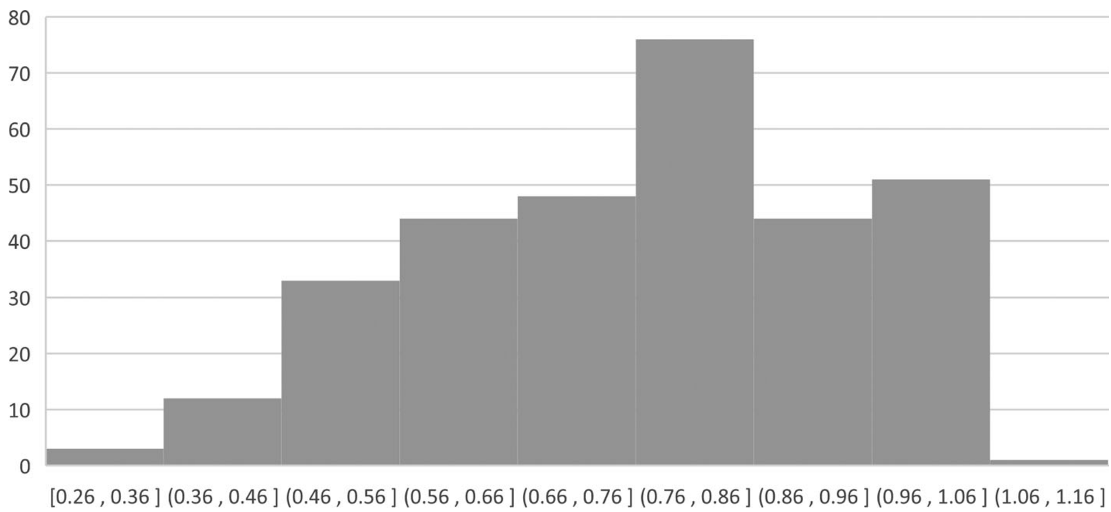


Table 6.4: FD's effects on healthcare efficiency (super-efficiency DEA)

	Super-efficiency DEA Fixed effects			Super-efficiency DEA Random effects		
	(1)	(2)	(3)	(4)	(5)	(6)
HED	0.516***			0.533***		
	(0.147)			(0.186)		
TED		0.317			0.278	
		(0.215)			(0.172)	
RD			0.601***			0.653***
			(0.144)			(0.141)
Lngdppc	-0.111	-0.074	-0.021	-0.011	0.019	0.059
	(0.106)	(0.102)	(0.099)	(0.069)	(0.069)	(0.065)
Lnpop	-1.833	-1.965	-1.362	-0.068	-0.054	-0.063
	(1.190)	(1.199)	(1.134)	(0.054)	(0.055)	(0.054)
Popden	-0.000	0.000	-0.001	0.000	0.000	-0.000
	(0.003)	(0.003)	(0.002)	(0.000)	(0.000)	(0.000)
Urban	0.001	0.001	0.000	0.001	-0.001	0.000
	(0.007)	(0.007)	(0.008)	(0.003)	(0.003)	(0.003)
Education	0.005	0.005	0.006	0.009***	0.010***	0.009***
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)
Elderly	-0.027*	-0.026*	-0.024*	0.002	0.002	-0.000
	(0.014)	(0.015)	(0.014)	(0.007)	(0.007)	(0.006)
Unemp	0.015	0.005	0.015	-0.006	-0.014	-0.011
	(0.026)	(0.026)	(0.024)	(0.020)	(0.019)	(0.019)
Lnhexppc	0.040	0.040	0.013	0.085***	0.087***	0.055*
	(0.044)	(0.043)	(0.037)	(0.031)	(0.031)	(0.031)
Solvency	0.003***	0.003***	0.001	0.003***	0.003***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Fragmentation	-2.447	-2.787	-2.096	0.716	0.528	0.094
	(7.323)	(7.679)	(7.280)	(1.292)	(1.313)	(1.289)
Minority regions				0.030	0.027	-0.001
				(0.072)	(0.073)	(0.073)
Time	0.040**	0.038**	0.035**			
	(0.015)	(0.015)	(0.015)			
Constant	33.041	35.138	24.283	0.781	0.561	0.316
	(20.507)	(20.615)	(19.405)	(1.010)	(1.026)	(1.001)
R-sq within	0.68	0.67	0.69			
R-sq overall				0.38	0.36	0.36
F-test	33.64	31.97	42.31			
Prob > F	0.000	0.000	0.000			
Wald chi2				489.83	476.39	526.55
Prob≥chi2				0.000	0.000	0.000

N=312. Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Consistent findings regarding control variables are also identified. In fixed effects regressions (see columns 1 to 3 of Table 6.4), *Elderly* and *Solvency* are found to be statistically significant. The first one measures the proportion of residents over 65 and is negatively related to healthcare efficiency at the 10% significance level in all three models. This result indicates that a high percentage of the elderly population in a region might bring excessive medical demands and burdens to the local healthcare sector (Guo et al., 2017), leading to diseconomies of scale and reducing healthcare efficiency. Second, the positive impact of fiscal solvency on efficiency is statistically confirmed at the 1% level in the two models controlling HED and TED, which, as explained above, could be attributed to the sub-provincial governments' greater motivation in improving efficiency once they have greater control over local resources. Additionally, in line with the fixed effects model controlling for efficiencies measured by the BCC-DEA approach, the year trend variable (*time*) is still positively related to healthcare efficiency at the 5% significance level. In random effects GLS regressions (see columns 4 to 6 of Table 6.4), results are still consistent with the Tobit models controlling for BCC-DEA scores. First, *education* is positively significant at the 1% level in all three models controlling for HED, TED, and RD, showing that citizens' better educational backgrounds may contribute to a greater level of healthcare efficiency. Then, per capita healthcare (in a natural logarithm form) is still positive and significant in all three models, and the positive effects of fiscal solvency can be confirmed in two models controlling for HED and TED. These results indicate the benefits of a higher level of healthcare funding and fiscal independence on healthcare efficiency.

6.1.2.2 The moderating effects of relative wealth on the relationship between FD and healthcare efficiency

This part looks into the moderating effects of relative wealth (lngdppc) on the relationship between FD and healthcare efficiency, which is represented by the interaction terms *FD*GDP* (i.e., HED*GDP, TED*GDP, and RD*GDP). In fixed effects models with BCC-DEA efficiencies, as shown in columns 1-3 of Table 6.5, the interaction term *HED*GDP* equals 0.530 and is significantly related to efficiency at the 10% level. However, coefficients of the second and third interaction terms *TED*GDP* and *RD*GDP* are statistically insignificant.[27]

[27] It is noticeable that most of the coefficients of HED, TED, and RD in the moderating models are statistically significant, as shown in Table 6.5. However, for a model with a dependent variable (Y), an independent variable (X), a moderator (Z), and an interaction term (X*Z), i.e., a multiplicative interaction model, the coefficient of X is no longer the average marginal effect of this variable on Y (Brambor et al. 2005; Hainmueller et al. 2019). Instead, this coefficient only represents the effect of X on Y when Z equals zero. Thus, for this study, the coefficients of HED, TED, and RD in the three multiplicative interaction models should not be directly interpreted as the impact of FD on efficiency. In fact, “the only clear way to gauge the average effect of X on Y is to run an unconditional model in which X is not included in a multiplicative interaction term” (Brambor et al. 2005, p.11). In other words, the findings in regard to the coefficients for HED, TED, and RD should be interpreted by inspecting the original models without the interaction variable shown in Table 6.3 and Table 6.4.

Table 6.5: Moderating effects of relative wealth on FD-efficiency relationship

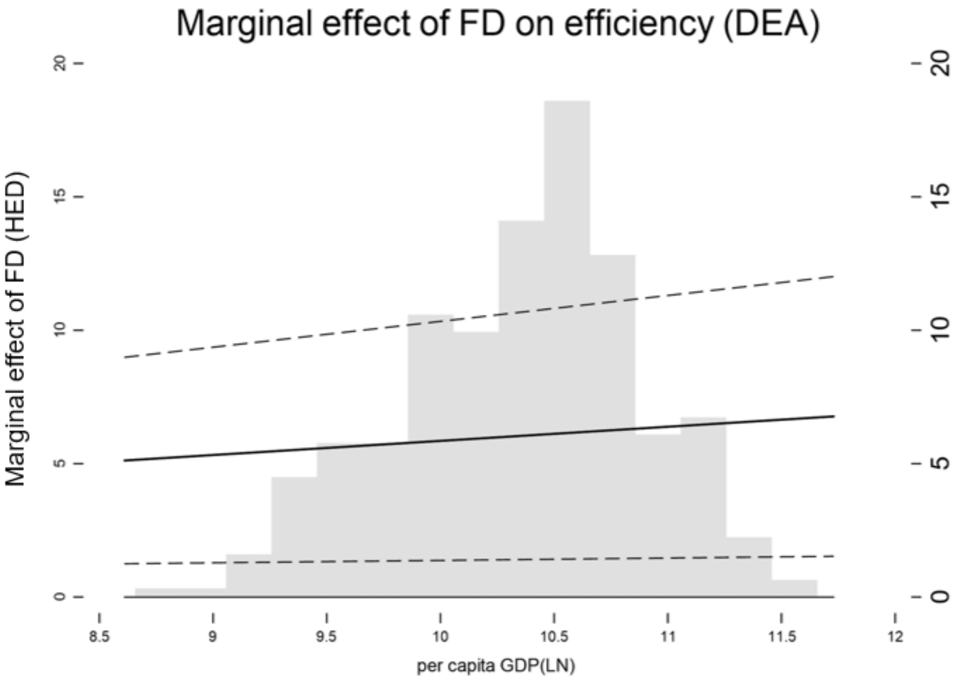
	BCC-DEA Fixed effects			BCC-DEA Tobit regression			Super-efficiency DEA Fixed effects			Super-efficiency DEA Random effects		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
FD*GDP	0.530*	-0.214	0.087	0.698**	0.267	0.196	0.594*	-0.113	0.334	0.678**	0.373	0.393
	(0.267)	(0.270)	(0.255)	(0.290)	(0.311)	(0.359)	(0.299)	(0.390)	(0.330)	(0.331)	(0.350)	(0.411)
HED	0.558***			0.645***			0.577***			0.622***		
	(0.154)			(0.165)			(0.152)			(0.189)		
TED		0.336*			0.451***			0.300			0.339*	
		(0.181)			(0.161)			(0.202)			(0.184)	
RD			0.570***			0.583***			0.623***			0.681***
			(0.159)			(0.128)			(0.147)			(0.144)
Lngdppc	-0.108	-0.091	-0.030	-0.017	-0.002	0.046	-0.102	-0.077	-0.023	-0.002	0.017	0.058
	(0.092)	(0.088)	(0.087)	(0.060)	(0.060)	(0.058)	(0.105)	(0.101)	(0.099)	(0.069)	(0.069)	(0.065)
Lnpop	-1.394	-1.540*	-0.959	-0.102**	-0.085*	-0.084*	-1.829	-1.966	-1.368	-0.082	-0.054	-0.068
	(0.863)	(0.896)	(0.818)	(0.050)	(0.050)	(0.051)	(1.172)	(1.204)	(1.131)	(0.056)	(0.051)	(0.055)
Popden	-0.001	-0.000	-0.001	0.000	0.000	-0.000	-0.000	0.000	-0.001	0.000	0.000	-0.000
	(0.002)	(0.002)	(0.002)	(0.000)	(0.000)	(0.000)	(0.002)	(0.003)	(0.002)	(0.000)	(0.000)	(0.000)
Urban	-0.001	-0.000	-0.001	-0.002	-0.003	-0.001	0.000	0.001	-0.000	-0.001	-0.001	-0.000
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.008)	(0.003)	(0.003)	(0.003)
Edu	0.005	0.006	0.006	0.008**	0.009***	0.009***	0.005	0.005	0.006	0.008**	0.009**	0.009***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.003)	(0.004)	(0.003)
Elderly	-0.013	-0.012	-0.010	0.003	0.005	0.005	-0.028*	-0.026*	-0.025*	-0.001	-0.000	-0.001
	(0.009)	(0.009)	(0.009)	(0.006)	(0.006)	(0.006)	(0.014)	(0.015)	(0.014)	(0.007)	(0.007)	(0.007)
Unemp	0.014	0.002	0.012	-0.004	-0.018	-0.015	0.018	0.005	0.016	0.001	-0.013	-0.008
	(0.021)	(0.021)	(0.020)	(0.017)	(0.017)	(0.017)	(0.027)	(0.027)	(0.025)	(0.020)	(0.019)	(0.019)
Lnhexppc	0.085*	0.052	0.038	0.098***	0.092***	0.057**	0.067	0.036	0.022	0.098***	0.097***	0.060*
	(0.046)	(0.043)	(0.039)	(0.028)	(0.028)	(0.028)	(0.049)	(0.047)	(0.039)	(0.031)	(0.031)	(0.031)
Solvency	0.003***	0.003***	0.001	0.003***	0.003***	0.001	0.003***	0.003***	0.001	0.003***	0.003***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Fragmentation	-3.466	-3.431	-2.901	-0.215	-0.173	-0.694	-2.737	-2.721	-2.151	0.422	0.496	-0.002
	(4.815)	(5.351)	(4.971)	(1.202)	(1.200)	(1.222)	(7.098)	(7.604)	(7.242)	(1.330)	(1.243)	(1.309)
Minority regions				0.012	0.014	-0.011				0.021	0.028	-0.001
				(0.064)	(0.065)	(0.069)				(0.075)	(0.068)	(0.074)
Time	0.025**	0.030**	0.024**				0.036**	0.039**	0.034**			
	(0.012)	(0.013)	(0.011)				(0.015)	(0.016)	(0.015)			
Constant	25.416	28.002*	17.484	1.478	1.338	1.021	32.775	35.192	24.387	0.826	0.542	0.379
	(14.907)	(15.401)	(13.969)	(0.911)	(0.913)	(0.934)	(20.207)	(20.661)	(19.349)	(1.040)	(0.961)	(1.019)
R-sq within	0.70	0.70	0.71				0.68	0.67	0.69			
R-sq overall										0.39	0.39	0.36
Pseudo R-sq				-0.33	-0.31	-0.34						
F-test	29.23	33.92	41.38				31.89	32.43	42.32			
Prob > F	0.000	0.000	0.000				0.000	0.000	0.000			
Wald chi2				533.45	506.28	551.87				503.16	469.98	529.49
Prob > chi2				0.000	0.000	0.000				0.000	0.000	0.000
LR test				201.38	180.83	206.51						
Prob≥chi2				0.000	0.000	0.000						

N=312. Standard errors in parentheses: *** p < 0.01, ** p < 0.05, * p < 0.1. For Tobit regressions: uncensored observations: 276, left-censored: 0, right-censored: 36

The positive coefficient for *HED*GDP* suggests that an increase in per capita GDP can potentially strengthen the positive impact of HED on healthcare efficiency. In other words, the benefit of decentralising responsibilities for healthcare provision might be greater in wealthier regions. This result is consistent with the hypothesis that the higher capacity brought by the growth of relative wealth could support the decentralised local healthcare sector to make greater a contribution to service efficiency. Detailed explanations are given in the Discussion chapter.

Following the recommendations of Golder (2006), Figure 6.3 graphs the marginal effects of GDP (*lngdppc*) on the relationship between each HED and healthcare efficiency. In this figure, per capita GDP's positive marginal effect is illustrated by the solid upward line, while values of the marginal effect are shown on the left side of the graphs. The histogram represents the percentage of *lngdppc*'s observations within a certain interval. Two dotted curves above and below the solid line represent the 90% confidence intervals for the marginal effects. If both two lines are located above or below the zero line, the effect of FD on efficiency can be confirmed at the 10% significance level (Brambor et al. 2005). Thus, as shown in Figure 6.3, the benefits of HED for healthcare efficiency are consistent over the given range of *lngdppc*'s observations.

Figure 6.3: Marginal effect of HED on BCC-DEA efficiency contingent on relative wealth (fixed effects model)



Findings from the Tobit models are shown in columns 4 to 6 of Table 6.5. In line with the fixed effects models, one of the three interaction terms *HED*GDP* is found to be positively related to healthcare efficiency at the 5% significance level. These findings are illustrated in Figure 6.4. Again, the positive effect of *HED*GDP* on healthcare efficiency is statistically significant for all given *lngdppc* observations.

For fixed effects and random effects models with efficiencies measured by the super-efficiency DEA approach, results are shown in columns 7 – 12 of Table 6.5. Still, in line with the findings from models with BCC-DEA efficiencies, only HED*GDP is positively related to efficiency (at the 10% significance level in the fixed effects model and at the 5% significance level in the random effects model). As demonstrated by Figure 6.5 and Figure 6.6, the statistically significant effects of HED*GDP on efficiency are present for all *lngdppc* values. The above findings will be discussed in more detail in the Discussion chapter.

Figure 6.4: Marginal effect of HED on BCC-DEA efficiency contingent on relative wealth (Tobit effects model)

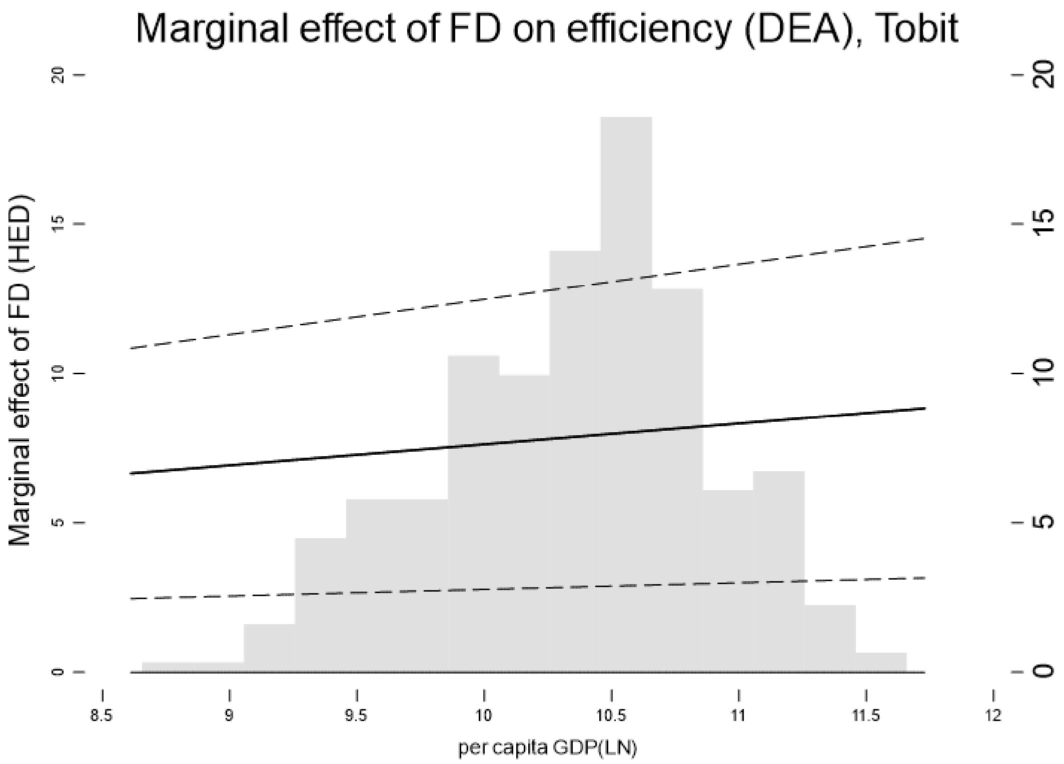


Figure 6.5: Marginal effect of HED on efficiency contingent on relative wealth (fixed effects model, Super-efficiency DEA)

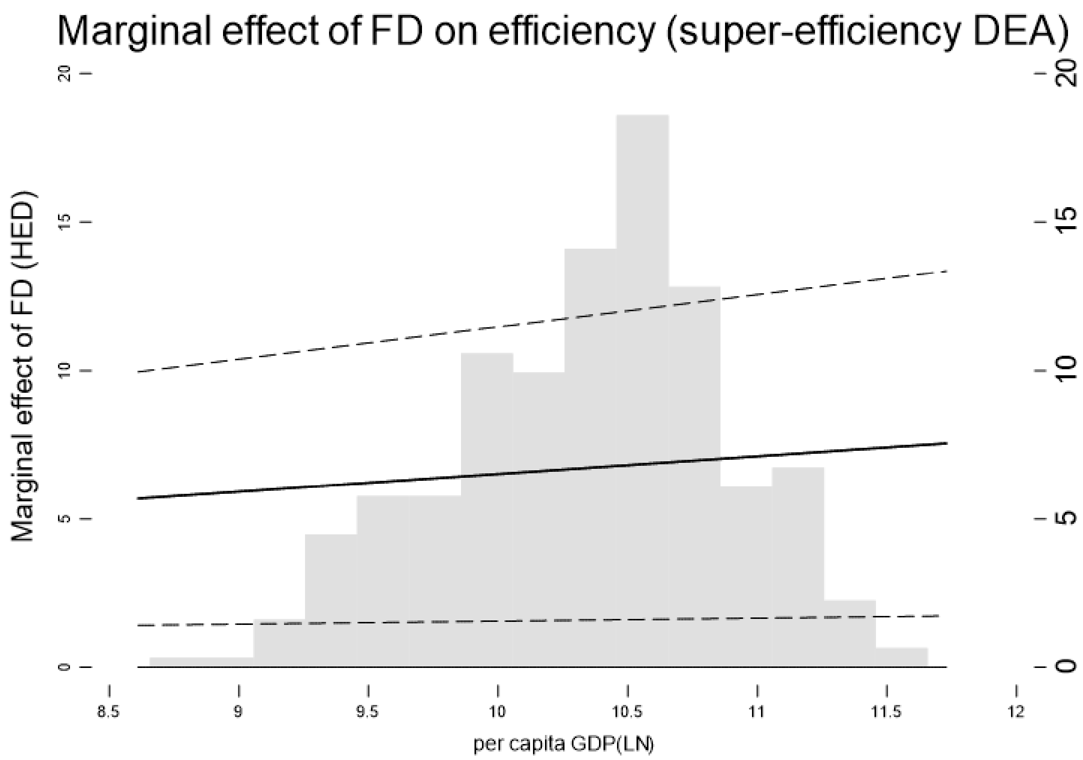
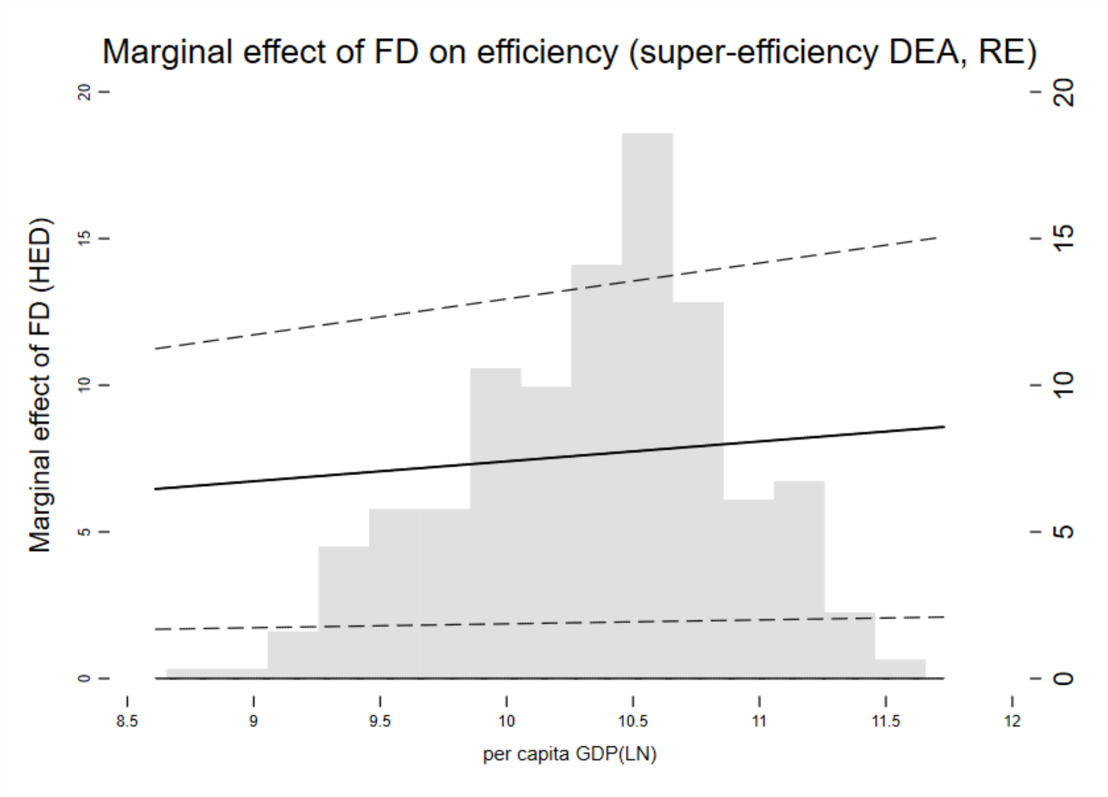


Figure 6.6: Marginal effect of HED on efficiency contingent on relative wealth (random effects model, Super-efficiency DEA)



6.1.3 Summary

Key findings in terms of the effects of FD and the moderating effects of relative wealth on the FD-efficiency relationship are summarised in Table 6.6. It shows that all three types of FD are positively related to healthcare efficiency. Further investigation of the moderating effects shows that as relative wealth (i.e., government capacity) grows, HED may bring more benefits for healthcare efficiency. However, TED and RD fail to bring such greater contributions. In the next section, the individual impact of FD on healthcare effectiveness – the second key aspect of PSP and the moderating effects of relative wealth on the FD-effectiveness relationship will be presented.

Table 6.6: Summary of key findings for the impact of FD on healthcare efficiency

Effects	FD indicators	BCC-DEA efficiency		Super-efficiency DEA	
		Fixed effects models	Tobit models	Fixed effects models	Random effects models
Effects of FD on efficiency	HED	(+)***	(+)***	(+)***	(+)***
	TED	(+)*	(+)***	N.S.	N.S.
	RD	(+)***	(+)***	(+)***	(+)***
Moderating effects of relative wealth	HED	(+)*	(+)**	(+)*	(+)**
	TED	N.S.	N.S.	N.S.	N.S.
	RD	N.S.	N.S.	N.S.	N.S.

(+) means a positive effect of FD on healthcare efficiency or a positive moderating effect of relative wealth on the FD-efficiency relationship, (-) means a negative effect of FD on healthcare efficiency or a negative moderating effect of relative wealth on the FD-efficiency relationship; *** $p < 0.01$, ** $p < 0.05$, * $p < 0$. N.S. means insignificant

6.2 FD and healthcare effectiveness

In line with the structure of the last section, summary statistics and correlation analysis are given at the beginning of this section. Then, regression results are reported and explained.

6.2.1 Summary statistics and correlation analysis

Table 6.7 reports the number of observations, min values, max values, mean and median values, and standard deviations of the variables. The summary statistics for the three FD indicators and the healthcare effectiveness indicator (PSR) have been given in the Methodology chapter (5.4.1 and 5.4.2). Also, the summary statistics for most control variables in the FD-effectiveness models have been reported in the last section (6.1.1) as they were also included in the FD-efficiency models. Thus, only the summary statistics for *birthrate* (birth rate per 1,000 people), *healthcare institutions* (number of public medical institutions per 1,000 people), and *beds* (number of hospital beds per 1,000 people) are reported here. Moreover, Figure 6.7 reports the yearly variations in the average values of *birthrate*, *healthcare institutions*, and *beds*, in which values are non-dimensioned to ensure comparability. For *birthrate*, its max value (17.89), min value (5.36), median value (11.98), and mean value (11.66) show a low level of birth rate in China, which is even lower than the global average birth rate in 2018 (18.63, World Bank 2018).

The summary statistics for *Healthcare institutions* show that on average, there are only 7.3 public medical institutions for 100,000 persons. As illustrated by the value of the standard deviation and Figure 6.6, the yearly variation of this indicator remains relatively stable from 2006 to 2017. Nevertheless, people have witnessed a significant increase in terms of the number of hospital beds: from 2006 to 2017, the average number of beds increased by more than 100%. The summary statistics table shows that for every 1,000 people there are 4.22 hospital beds in China, with the maximum and minimum values equal to 7.21 and 1.87, respectively. The average value of 4.22 is significantly higher than the average values for the world (2.89) and for upper middle-income countries (3.49) in 2018 (World Bank 2018), showing the increasing efforts made by the Chinese government to promote healthcare infrastructures. Table 6.8 presents the correlation matrix for the selected variables and VIF values, in which statistically significant correlations can be identified between most variables. The VIF values in all FD-effectiveness models are smaller than 10, thus, multicollinearity is still not a serious concern.

Figure 6.7: Yearly variations of 3 control variables (all variables are non-dimensionalised)

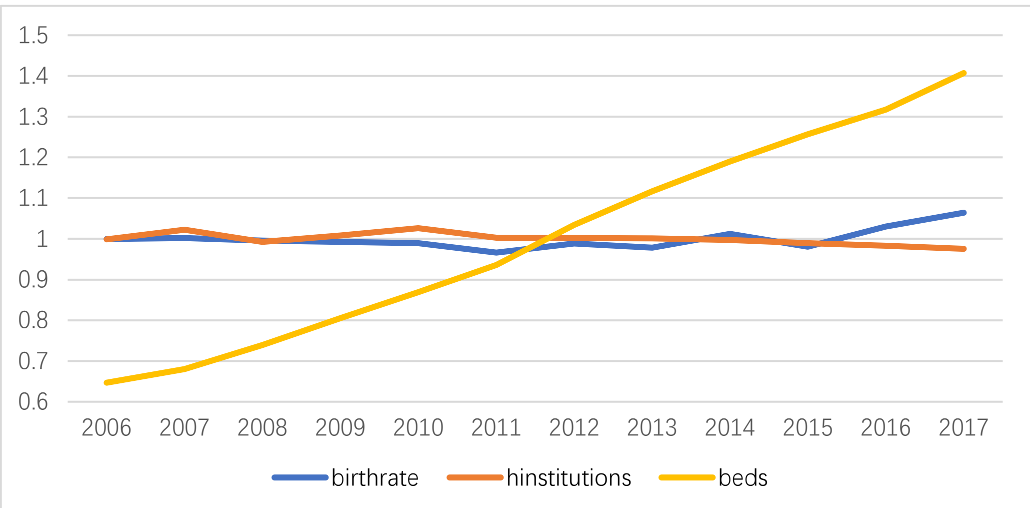


Table 6.7: Summary statistics of variables for FD-effectiveness models

Variables	Obs.	Min	Max	Mean	Median	Standard deviation
PSR (%)	312	98.023	99.719	99.260	99.316	0.314
Birthrate (%)	312	5.360	17.890	11.657	11.980	2.538
Hinstitutions	312	0.038	0.153	0.073	0.064	0.028
Beds	312	1.870	7.214	4.217	4.019	1.048

Abbreviations: Hinstitutions, number of public medical institutions per 1,000 persons; Beds, number of hospital beds per 1,000 persons.

Table 6.8: Correlation Matrix for FD-effectiveness models

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. PSR	1															
2. HED	0.5658*	1														
3. TED	0.6516*	0.7470*	1													
4. RD	0.3587*	0.5049*	0.6324*	1												
5. GDP per capita (ln)	0.5993*	0.4896*	0.5514*	0.4350*	1											
6. Population (ln)	0.4750*	0.6712*	0.7164*	0.5078*	0.1605*	1										
7. Population density	0.5630*	0.3636*	0.6235*	0.5718*	0.3479*	0.6675*	1									
8. Urbanisation	0.4734*	0.2931*	0.4717*	0.3365*	0.8585*	0.1132*	0.3666*	1								
9. Education	0.3150*	0.3240*	0.2743*	0.2755*	0.7905*	-0.0845	0.0311	0.7113*	1							
10. Birthrate	-0.1996*	-0.1628*	-0.2454*	-0.0587	-0.2369*	-0.2716*	-0.0741	-0.4252*	-0.2707*	1						
11. Health institutions	-0.2623*	-0.3597*	-0.3051*	-0.2071*	0.0645	-0.4820*	-0.3970*	0.0666	0.2423*	-0.0669	1					
12. Beds	0.2851*	0.4160*	0.1890*	0.1913*	0.6706*	-0.0587	-0.1511*	0.4116*	0.7589*	-0.0738	0.1542*	1				
13. Healthcare spending per capita (ln)	0.4291*	0.3661*	0.2130*	0.1108*	0.7267*	-0.1920*	-0.1192*	0.4396*	0.6885*	0.0909	0.1597*	0.8213*	1			
14. Fiscal solvency	0.3316*	0.2515*	0.5135*	0.6150*	0.4303*	0.5173*	0.7675*	0.5376*	0.1780*	-0.2064*	-0.1697*	-0.1414*	-0.1731*	1		
15. Fragmentation	-0.5045*	-0.5720*	-0.6302*	-0.3779*	-0.1364*	-0.7947*	-0.5972*	-0.1627*	0.0746	0.2821*	0.5507*	0.1396*	0.2056*	-0.5198*	1	
16. Minority	-0.4605*	-0.1419*	-0.2276*	-0.0693	0.0034	-0.3663*	-0.4020*	-0.0800	0.1004*	0.2633*	0.1705*	0.1101*	0.0858	-0.2142*	0.3250*	1
VIF (HED controlled)		4.06			8.88	6.74	4.89	5.92	5.38	2.34	1.81	5.18	6.70	4.78	3.60	1.54

* represents the 10% significance level

VIF values are from on the models where DEA is the dependent variable and HED is the key independent variable. For VIF values from the models controlling for TED and RD, see Appendix 7.

6.2.2 Regression results: FD and healthcare effectiveness

Estimation results for the impact of FD on healthcare service effectiveness (measured by PSR) as well as the impact of other control variables are reported in the following paragraphs. It is noticeable that for all regression models, the p-values (0.000) of the Sargan-Hansen statistics favour the fixed effects models over the random effects models. However, for completeness, the findings from random effects models are presented as well. As for Section 6.1, the effects of FD and the moderating effects of relative wealth on the FD-effectiveness relationship are discussed separately.

6.2.2.1 Effects of FD on healthcare effectiveness

Regarding the FD-effectiveness relationship, columns 1 to 3 and columns 4 to 6 in Table 6.9 present fixed effects and random effects estimations from the three models including HED, TED, and RD[28]. As shown in columns 1 to 3, for the three fixed effects models, the relationship between HED and effectiveness (measured by PSR) cannot be statistically confirmed. While the second and third indicators TED and RD, in contrast to the expectation, have negative relationships with healthcare effectiveness at the 5% and 10% significance levels, respectively. Considering the values of FD indicators are located between 0 (fully centralised) and 1 (fully decentralised), the above coefficient also means that, if TED and RD grow by 0.01, the value of PSR will have an decrease of 0.00458 and 0.00384 units. Similar findings can also be seen in random effects models (columns 4 to 7 of Table 6.9): the

[28] Hausman test results still favour fixed effects analysis.

relationship between HED and PSR cannot be confirmed, while the positive relationship between TED and PSR and between RD and PSR are statistically identified at the 10% and 5% significance level, respectively.

Table 6.9: FD's effects on healthcare effectiveness (Perinatal survival rate)

Dep: PSR	(1)	(2)	(3)	(4)	(5)	(6)
		Fixed effects			Random effects	
HED	0.046			0.150		
	(0.219)			(0.187)		
TED		-0.458**			-0.294*	
		(0.175)			(0.175)	
RD			-0.384*			-0.342**
			(0.191)			(0.149)
Lngdppc	0.219***	0.262***	0.218***	0.241***	0.287***	0.251***
	(0.057)	(0.069)	(0.058)	(0.074)	(0.073)	(0.071)
Lnpop	0.394	0.529	0.102	-0.018	0.009	0.001
	(0.751)	(0.723)	(0.769)	(0.067)	(0.067)	(0.065)
Popden	-0.003**	-0.004**	-0.003**	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Urban	0.005	0.006	0.005	-0.002	-0.002	-0.003
	(0.006)	(0.006)	(0.006)	(0.003)	(0.003)	(0.003)
Edu	0.007*	0.007*	0.007*	0.005	0.005	0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Birthrate	-0.031***	-0.030***	-0.030***	-0.024***	-0.023***	-0.022***
	(0.010)	(0.009)	(0.009)	(0.006)	(0.006)	(0.006)
Hinstitutions	2.604***	2.436***	2.721***	2.048***	1.989***	2.077***
	(0.741)	(0.653)	(0.694)	(0.516)	(0.515)	(0.513)
Beds	0.008	0.009	0.009	0.035**	0.032*	0.037**
	(0.034)	(0.031)	(0.031)	(0.016)	(0.016)	(0.016)
Lnhexppc	0.029	0.027	0.044	0.047	0.041	0.059*
	(0.039)	(0.042)	(0.035)	(0.032)	(0.032)	(0.033)
Solvency	-0.085	-0.218	-0.016	-0.114	-0.201*	-0.048
	(0.114)	(0.130)	(0.084)	(0.099)	(0.106)	(0.103)
Fragmentation	-1.843	-1.685	-2.132	-4.553***	-4.790***	-4.607***
	(2.360)	(2.487)	(2.212)	(1.528)	(1.521)	(1.505)
Minority regions				-0.270***	-0.270***	-0.260***
				(0.092)	(0.092)	(0.091)
Time	0.009	0.007	0.010			
	(0.016)	(0.014)	(0.015)			
Constant	90.643***	88.399***	95.867***	96.863***	96.369***	96.720***
	(12.886)	(12.355)	(13.208)	(1.291)	(1.280)	(1.254)
F-test	133.90	87.16	114.03			
Wald chi2 test				1499.78	1513.22	1523.74
Prob	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-sq within	0.866	0.870	0.869			
R-sq overall				0.599	0.595	0.600

N=312. Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Regarding control variables, first, per capita GDP is positively related to healthcare effectiveness in all six fixed effects and random effects models at the 1% significance level. A higher level of economic development, therefore, seems to promote healthcare service effectiveness in Chinese provinces. This finding corroborates the argument that economic development brings greater technical, resource, and organisational advantages to the public sector, leading to better healthcare service effectiveness (Hillestad et al. 2005; Burns et al. 2012). Second, the anticipated positive relationship between population size and healthcare effectiveness cannot be statistically confirmed. This finding suggests that consistent with the absence of a population-efficiency relationship (see the previous section), in the Chinese socio-political background, the greater healthcare demands caused by a larger population size may make limited contributions to healthcare performance.

For population density and urbanisation, it is found that population density is negatively associated with healthcare effectiveness in the fixed effects models at the 5% significance level but has no significant effects in random effects models. This finding indicates that the growth of population density within provinces may bring excessive burdens and more complicated needs for the healthcare sector, leading to worse healthcare quality and effectiveness. This finding is in contrast with several previous studies which claim that population density proxies for "urban advantages" (Eberhardt et al. 2001; Hartley 2004), thus contributing to scale economies and better healthcare service effectiveness (Magadi et al. 2003; Matthews et al. 2010). However, in this research, a more direct indicator of urban advantages, i.e., the level of urbanisation, is included in regression models but is statistically insignificant. Moreover, different from the western setting, in developing countries such as China, many underdeveloped and rural areas with worse healthcare infrastructures and economic situations might be more densely populated and have higher birth rates (Hathi et al. 2017). In these areas, the growth of population density does not necessarily contribute to greater urban advantages. Thus, this variable is negatively related to healthcare effectiveness.

Regarding education and birth rate, two variables reflecting local demographic features, education has positive coefficients in all six models and is statistically significant (at 10%) in the three fixed effects models. This finding is consistent with the expectation that a higher level of education proxies for people's healthier lifestyle, better living conditions, and a more capable healthcare sector (Arends 2017; Jiménez-Rubio 2011b), which contributes to healthcare effectiveness. The negative coefficients for *birthrate* are statistically significant (at 1%) in all six models. It is likely that a higher level of birth rate increases the burden on the healthcare sector. Also, in China, the practice of favouring fertility is more deeply rooted in the clan culture of rural areas where healthcare services are underprovided (Gao et al. 2010; Hershatter 2019). Thus, the negative impact of a high birth rate on healthcare effectiveness – measured by a lower PSR – appears to be particularly significant.

All three healthcare inputs have positive coefficients in the six models: the coefficients for *healthcare institutions* are positive and statistically significant at the 1% level in all six models. The coefficients for *Beds* are positive but are only statistically significant in the random effects models. As discussed above, the increased availability of hospital beds in all provinces may explain this result. *Lnhexppc* is only statistically significant (and has a negative coefficient) in one of the random effects models (the RD estimates) at the 10% significance level, suggesting that a higher level of per capita healthcare expenditure, to a certain extent, might improve healthcare effectiveness.

The coefficient for fiscal solvency on healthcare effectiveness is negative but only statistically significant in the random effects model focusing on TED, showing that a higher level of financial independence may undermine the effectiveness of healthcare services. The coefficient for *fragmentation* – is negative and statistically significant in all of the random effects models. This suggests that a greater level of fragmentation within Chinese provinces may fail to motivate local governments to attract population inflow by promoting healthcare effectiveness. On the one hand, this phenomenon might be explained by the fact that the scale of intra-provincial immigration is much smaller than that between provincial jurisdictions (Zhao et al., 2018). On the other hand, it also reveals that the residential sorting mechanism (i.e., vote with feet), as a potential factor that makes intra-provincial FD work on healthcare performance, may not be as effective as expected. The year trend dummy (*time*) is statistically insignificant. It suggests that the unobservable time fixed effects which are invariant over the observations (i.e., provinces) but changing over the years, such as nationwide policies, have no significant impact on healthcare effectiveness. Finally, the coefficient for *Minority regions* is negative and statistically significant, indicating that PSR-measured healthcare effectiveness is worse in the Provincial Autonomous Regions for ethnic minorities. This is consistent with the fact that all of these regions are located in remote areas of China and are socioeconomically underdeveloped.

6.2.2.2 The moderating effects of relative wealth on the FD-effectiveness relationship

To investigate the impact of relative wealth (i.e., per capita GDP in a natural logarithm form) on the FD-effectiveness relationship, three interaction terms *HED*GDP*, *TED*GDP*, and *RD*GDP* are entered into the models. Estimation results are reported in Table 6.10, in which columns 1 to 3 present fixed effects estimates, and columns 4 to 6 report random effects estimates. Different from the FD-efficiency models, none of the moderators is statistically significant. Thus, the moderating effects of relative wealth on the relationship between FD and healthcare effectiveness cannot be confirmed. This finding will be explored in the Discussion chapter.

Table 6.10: Moderating effects of relative wealth on the FD-effectiveness relationship

Dep: PSR	(1)	(2)	(3)	(4)	(5)	(6)
		Fixed effects			Random effects	
FD*GDP	-0.353	-0.637	0.402	-0.301	-0.302	0.501
	(0.212)	(0.377)	(0.463)	(0.368)	(0.384)	(0.444)
HED	0.017			0.117		
	(0.243)			(0.191)		
TED		-0.556***			-0.344*	
		(0.159)			(0.187)	
RD			-0.358*			-0.305**
			(0.180)			(0.152)
Lngdppc	0.211***	0.246***	0.219***	0.232***	0.278***	0.253***
	(0.057)	(0.069)	(0.058)	(0.075)	(0.074)	(0.071)
Lnpop	0.432	0.570	0.076	-0.012	0.011	-0.006
	(0.738)	(0.720)	(0.778)	(0.068)	(0.065)	(0.062)
Popden	-0.003**	-0.004**	-0.003**	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Urban	0.005	0.007	0.004	-0.001	-0.001	-0.003
	(0.006)	(0.006)	(0.006)	(0.003)	(0.003)	(0.003)
Edu	0.007*	0.007*	0.007*	0.005	0.005	0.005
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Birthrate	-0.031***	-0.030***	-0.030***	-0.023***	-0.022***	-0.021***
	(0.010)	(0.009)	(0.008)	(0.006)	(0.006)	(0.006)
Hinstitutions	2.540***	2.401***	2.766***	2.005***	1.912***	2.068***
	(0.744)	(0.646)	(0.702)	(0.523)	(0.519)	(0.514)
Beds	0.013	0.013	0.006	0.040**	0.037**	0.031*
	(0.033)	(0.032)	(0.032)	(0.018)	(0.018)	(0.017)
Lnhexppc	0.015	0.003	0.054	0.038	0.036	0.071**
	(0.040)	(0.042)	(0.042)	(0.034)	(0.033)	(0.034)
Fiscal solvency	-0.078	-0.197	-0.027	-0.112	-0.202*	-0.065
	(0.112)	(0.134)	(0.081)	(0.099)	(0.106)	(0.104)
Fragmentation	-1.582	-1.215	-2.248	-4.406***	-4.698***	-4.748***
	(2.281)	(2.334)	(2.261)	(1.553)	(1.508)	(1.460)
Minority region				-0.270***	-0.268***	-0.255***
				(0.094)	(0.089)	(0.086)
Time	0.011	0.012	0.010			
	(0.015)	(0.013)	(0.015)			
Constant	90.093***	87.951***	96.290***	96.877***	96.443***	96.779***
	(12.660)	(12.236)	(13.341)	(1.311)	(1.259)	(1.206)
F-test	323.15	99.41	107.97			
Wald chi2 test				1508.32	1499.31	1499.20
Prob	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-sq within	0.866	0.871	0.869			
R-sq overall				0.590	0.594	0.619

N=312. Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

6.2.3 Summary

Key findings from the analysis of the impact of FD on healthcare efficiency in Chinese provinces are summarised in Table 6.11. First, the impact of HED – the indicator directly measuring healthcare decentralisation cannot be confirmed, showing the limited contribution of decentralising healthcare responsibilities to healthcare service effectiveness. Second, TED is found to be negatively related to healthcare effectiveness in both random effects and fixed effects models (represented by a lower PSR). This result, in some way, reflects sub-provincial governments' relative neglect of healthcare effectiveness (Hao et al. 2021). In this case, decentralising the power for all sorts of government responsibilities to the sub-provincial governments fails to benefit healthcare effectiveness in the short run. Finally, RD is also negatively related to healthcare effectiveness in both fixed effects and random effects models, suggesting that under the current institutional background in China, allocating greater fiscal resources *per se* to sub-provincial jurisdictions does not appear to motivate local governments to reduce perinatal mortality rates. Finally, the moderating effect of relative wealth on the relationship between FD and healthcare effectiveness cannot be statistically confirmed. Detailed explanations about these findings will be given in the second section (7.2) of the Discussion chapter.

Table 6.11: Summary of key findings for the impact of FD on healthcare effectiveness

Effects	FD indicators	Results (relationship between FD and effectiveness)	
		Fixed effects models	Random effects models
Effects of FD on effectiveness	HED	N.S.	N.S.
	TED	(-)**	(-)*
	RD	(-)*	(-)**
Moderating effects of relative wealth	HED	N.S.	N.S.
	TED	N.S.	N.S.
	RD	N.S.	N.S.

(+) means a positive effect of FD on healthcare effectiveness, (-) means a negative effect of FD on healthcare effectiveness; *** $p < 0.01$, ** $p < 0.05$, * $p < 0$. N.S. means insignificant

CHAPTER 7 DISCUSSION

The impact of FD on healthcare service performance is an issue of timely and constant importance due to the lack of empirical evidence for decentralised policymakers and the general public's increasing concern for better healthcare (Channa and Faguet 2012; Arends 2017). Aiming to address this concern, this study focuses on intra-provincial FD in Chinese provinces between 2006 and 2017 and empirically investigates three key research questions: 1) the impact of FD on healthcare service efficiency; 2) the impact of FD on healthcare service effectiveness; and 3) the role of relative wealth on the relationship between FD and healthcare service performance. In Chapter 6, key findings for the above research questions and results related to the control variables are reported. Three FD indicators, namely the HED, TED, and RD, have positive effects on healthcare service efficiency. Moreover, the positive moderating effect of relative wealth (measured by per capita GDP) on FD's relationship with healthcare service efficiency can be statistically confirmed for HED. That is to say, the benefit of HED for healthcare service efficiency is larger with the growth of wealth/government capacity. For healthcare effectiveness measured by PSR, the results show that a positive impact of HED on healthcare service effectiveness cannot be confirmed. Furthermore, TED and RD are negatively related to healthcare effectiveness. Finally, the moderating effect of relative wealth on the FD-effectiveness relationship cannot be statistically confirmed in all models.

In this chapter, key features of the Chinese-style FD, i.e., the mechanism of upward accountability in Chinese provinces and the enhancement of healthcare services as a means to achieve performance legitimacy are introduced in Section 7.1. Then, Section 7.2 starts from the discussions of the target setting system (TRS) and the top-down PMS – two key underpinnings of upward accountability. This is followed by the explanations in regard to the impact of three types of FD (HED, TED, and RD) on healthcare efficiency. After that, in Section 7.3, the effects of FD on healthcare effectiveness are explored. Finally, Section 7.4 examines the effects of relative wealth on the relationship between FD and healthcare performance.

7.1 Reviewing the institutional background

As summarised in the Context chapter, intra-provincial FD in China has two key features: first, similar to FD reforms elsewhere, it authorises sub-provincial units to formulate and implement policies for local services. Second and different from FD in western settings, it is built on China's hierarchical and top-down administrative system where all governmental organisations at each administrative level are upwardly accountable to the superior government (Zheng 2006; Chien 2010). Over the past decades, such a system of upward accountability has been consolidated and become the key feature and underpinning of FD in China (Edin 2003; Chan and Gao 2008; Du and Yi 2022). In the healthcare sector, over the past decades, healthcare responsibilities have been increasingly decentralised to HCs at the sub-provincial level (Meng et al. 2012; Hao et al. 2021). Aligning with the current directions of upward accountability (see Figure 4.1), HCs (e.g., the prefecture-level HC), like most of the functional departments, are primarily accountable to the local government (e.g., the prefecture-level government) (Hipgrave et al. 2012; Ma 2017b; Tsai and Liao 2020). Then, the prefecture-level government is accountable to the provincial government, and so on. Such a link of upward accountability ensures that there is a consistency of the major purpose across multiple levels of government.

In recent years, the top party and state leadership of China has moved away from the GDP-centred pattern of development to focus more on other key governmental functions, including the improvement of healthcare services (Li 2011; Zeng 2014; Dickson et al. 2016). This greater emphasis on healthcare services and healthcare performance is not only demonstrated by the CPC's top policy agendas such as the "Scientific Outlook on Development", "Modernising Governance System and Capacity", and the "Healthy China 2030 Plan" initiated by the former and current state leaders (Central Committee of the CPC and State Council 2016; Tan et al. 2017; Tan et al. 2018), but also has been continuously delivered to provincial and sub-provincial governments and healthcare departments (Zhang 2020).

The aforementioned greater emphasis on healthcare "from the top", as described in the CPC's political discourse and the Chinese culture (Gore 2019), comes from the thought of "Mandate from Heaven" (*Tianming*) – a Chinese expression of legitimacy: "*…… the legitimacy of rulers is inherited from Tianming rather than from elections.*

But at the same time, the rulers had to keep meeting the needs of the people.". If the rulers failed to satisfy people's enduring needs, the complaints of the people will be "heard" by heaven, and consequently, the ruler will "*forfeit the source of his legitimacy and with it his divinely ordained entitlement to rule*". Finally, they will be overthrown by others entitled with the new Mandate (Weatherley and Magee 2018, p.53). Such ultimate failure of losing Tianming or legitimacy, in the political language of the CCP's top leadership, is described as *wangdang wangguo* (death of the party and death of the country) (Zhu 2011). Institutionally, the party-state's claim to legitimacy does not derive from democratic means such as elections – the CCP's ruling position originated from its success in the Chinese revolution and has been stipulated in the PRC's Constitution (Holbig and Gilley 2010). Instead, the main source of legitimacy for the party-state of China comes from better performance in satisfying people's changing needs (Zhao 2009). Since the early 2000s, people's discontent with inaccessible and low-quality healthcare services resulted in growing chaos in society, which has eroded the foundation of the CCP's regained legitimacy after the Reform and Openness period (Holbig 2009). In response to these legitimacy challenges, top leaders of the CPC and the Chinese government have been paying greater attention to the performance in providing essential public services, including healthcare (Holbig and Gilley 2010; Duckett and Munro 2022).

7.2 Making decentralisation work for healthcare efficiency: regulating upward accountability with target setting and top-down PMS for cadre evaluation

The above section provides a more detailed review that the party-state leadership in China, generally, are paying greater attention to healthcare and healthcare service performance. While more specifically, recognising that healthcare is a service with strong regional features in China, it is the sub-provincial governments that undertake the dominant responsibility for formulating and implementing local healthcare policies (Carrillo and Duckett 2011; Hipgrave et al. 2012). As mentioned by seminal studies in FD (e.g. Oates 1972), one of the assumptions of FD is that the local government has a greater understanding of local demands and a greater capacity to manage local resources and organisations. Accordingly, the advantages associated with decentralisation enable local authorities to improve PSP. This argument regarding local advantages, to a large extent, can be generalised to the Chinese context - due to the large landscapes, population, and socioeconomic differences, the higher-level government appears to have lower "legibility"[29] than their local counterparts (Shi and Ni 2017; Bulman and Jaros 2020). This is especially the case for healthcare services which are highly related to contextual factors such as socioeconomic status and demographic structures. Thus, in practice, higher-level governments in China not only decentralise administrative and fiscal power to local jurisdictions for healthcare, but also appoint a high proportion of "localists" as leaders

[29] As explained by Scott (2008), legibility refers to the extent to which the government can clearly "see" into their jurisdictions and manipulate local resources.

in charge of local healthcare policies (Hipgrave et al. 2012; Bulman and Jaros 2020). Thus, the assumption of FD that local officials know more about local healthcare demands also appears to be generalisable in the context of Chinese provinces.

However, previous studies suggest that compared with other government responsibilities, such as promoting infrastructure and attracting foreign investments, healthcare performance is less likely to produce economic benefits for local cadres (Lin and Liu 2000; Luo and Chen 2010; Shen et al. 2012). Thus, despite a greater emphasis on healthcare, the decentralised local governments still prefer to devote their efforts and advantages to “money-generating” activities instead of promoting healthcare performance (Shen et al. 2012). This raises the question of how the sub-provincial governments and healthcare agencies can be regulated to be accountable for the central government’s emphasis and requirements about healthcare without sacrificing their local advantages. In China, there are two major solutions to this question: target-setting and the top-down PMS.

Target-setting is accomplished by the TRS through which the higher-level government formulates target contracts jointly with the lower-level governments or functional departments[30]. In this way, the upper-level government’s policy agendas and requirements can be highlighted in the targets of the lower-level governments, while

[30] For example, the provincial government sets contracts with the governments of the affiliated prefecture-level cities, and the prefecture-level city government sets contracts with the prefectural HC.

lower-level governments are granted sufficient autonomy to develop detailed policies and approaches to attain the targets – with their information advantages regarding local status, demands, and resources (Gao 2009; Ho 2010). Then, the second key policy tool, i.e., the top-down PMS, ensures that the benefits (e.g. promotion, rewards, and bonuses) of healthcare departments, leaders, and staff are all determined by the top-down performance evaluation. For leaders of local HCs (e.g., the prefecture-level city's HC), performance evaluation is undertaken by the local government and CPC Committee. Within the healthcare sector, performance evaluation is also conducted in a top-down manner – cadres who fail to meet the performance targets could be punished by demotion, dismissal, or cutting down the funding of their departments (Qiu and Macnaughton 2017). In other words, the relationship of upward accountability also exists within governmental departments (e.g. the HC) as an institutional arrangement to regulate the behaviour of individuals.

In summary, a powerful system of target-setting (via TRS) and top-down PMS play an effective role in consolidating upward accountability in Chinese provinces. Under such arrangements, FD would not motivate the local governments and HCs to deviate from the higher-level government. Instead, they will fully utilise their greater responsibilities and revenue to satisfy their superiors with their initiatives (Chien 2010). For example, they are able to formulate and implement policies by fully considering local contexts, which leads to better healthcare performance (Shi and Ni 2017). Meanwhile, with decentralisation, the upper-level government gradually withdraw from local affairs but pay greater attention to monitoring and evaluating lower-level agents, as argued by Gao (2009) and Walker and Wu (2010), leading to a stronger state capacity by which the TRS and the top-down PMS can be consolidated, which may further promote healthcare service provision.

The above analysis, seemingly, entails that FD would consistently promote healthcare performance in Chinese provinces. However, empirical findings this study suggest that in the context of consolidated upward accountability, FD is beneficial for healthcare efficiency, but in most cases, it fails to benefit the effectiveness of healthcare services. To further understand these contradictory findings, we now separately examine the positive impact of three intra-provincial FD settings on healthcare service efficiency. Then, in Section 7.3, the negative effects of three FD indicators on healthcare effectiveness will be discussed in detail.

- **HED and healthcare service efficiency**

The first FD indicator, HED, refers to the percentage of healthcare expenditure spent by sub-provincial jurisdictions in a province. A greater value of HED shows that the sub-provincial HCs have to undertake greater responsibilities in local healthcare services. In this case, sub-provincial HCs have to be more accountable to the sub-provincial government's increasing requirements for healthcare performance, especially improved efficiency (Zhou et al. 2017; Yip et al. 2019). Meanwhile, given the sub-provincial HCs' greater informational, managerial, and organisational advantages regarding local healthcare services, a higher level of healthcare decentralisation also provides more room and opportunities for HCs to satisfy their superiors by fully utilising the above advantages. This will lead to better healthcare service efficiency (Millar et al. 2015). In practice, sub-provincial HCs' utilisation of their advantages under a decentralised setting is reflected by the policy innovations emerging in Chinese grassroots jurisdictions over the past years.

As mentioned by Shi (2012) and Heilmann (2008), China's progress in almost all socioeconomic aspects can be attributed to the policy innovations and experiments made by localities in a context where political and personnel power remain centralised but administrative and fiscal power are decentralised. Healthcare is an area where local innovations are particularly noticeable (Xiao et al. 2018). With healthcare decentralisation, sub-provincial jurisdictions in China are the centre of most healthcare innovations that aim to promote the performance and day-to-day management of the healthcare sector (Millar et al. 2015; Xiao et al. 2018). In fact, innovation is not only voluntarily carried forward by the local healthcare department under a decentralised setting, but also, it is a behaviour for which sufficient space and discretionary power are given by the upper-level governments (Ho 2010; Husain 2017). As explained by Millar et al. (2015), Yip and Hsiao (2015), and Yip et al. (2019), the geographical and socioeconomic variations in Chinese local jurisdiction make it impossible to set one-size-fits-all healthcare plans without considering local situations and mobilising local initiatives. Moreover, for unprecedented attempts in healthcare (e.g. the public hospital reform in 2009), both the central and local governments have no idea about "what works" (Husain 2017). Therefore, in a decentralised setting, initiating policy innovations in healthcare appears to be a natural choice for the local healthcare department.

For decades, most healthcare innovations in local jurisdictions of China can be summarised into a pattern that following general agendas or “open” policy frameworks given by the upper-level government, local healthcare staff formulate and implement specific policies with a detailed consideration of local situations and needs (Bowe and Ball 1992; Husain 2017). The upper-level governments will not intervene in local innovations unless their key agendas are seriously challenged (Li and Fu 2017). With a greater level of intra-provincial healthcare decentralisation, sub-provincial HCs gain greater autonomy and need to undertake more responsibilities. In this case, to satisfy the requirements of local government leaders for better healthcare, more innovations would be initiated at the sub-provincial level, which is expected to promote healthcare efficiency – a key concern highlighted by the sub-provincial government leaders (Chung 2000; Li and Fu 2017; Yip et al. 2019). As explained by Mei and Wang (2017), those grassroots bureaucrats who initiated remarkable and outstanding innovation stories, apart from being reward through the top-down PMS, might be further recognised by senior political leaders of China via informal channels such as the coverage of state media, which, in a long run, would bring greater political benefits for their career.

- **TED and healthcare service efficiency**

The second FD indicator TED represents the ratio of all sub-provincial expenditures to the total fiscal expenditures of the province. A higher value of TED means that in a province, greater responsibilities for all local affairs, as a whole, are decentralised to the sub-provincial bureaucracy. As mentioned above, sub-provincial governments, nowadays, are also under the upper-level government and the CPC's increasing pressure for better healthcare efficiency. Thus, supported by TED, sub-provincial governments may have greater space and willingness to perform as the political leader, monitor, and evaluator of the local healthcare department (Gao 2009; Chien 2010; Walker and Wu 2010), so as to win yardstick competitions and opportunities for promotion and other high-powered benefits. Specifically, this is achieved by addressing key performance concerns when setting targets, better monitoring the local HC's undertaking of responsibilities, and evaluating the performance of the local HC against stricter requirements. Such indirect efforts may promote healthcare service efficiency as well.

Moreover, local HC's responsibilities in healthcare, a service highly related to local socioeconomic situations, cannot be fulfilled without the cooperation of other local departments (Beaglehole 2004). For example, in Chinese sub-provincial jurisdictions, the project of building or expanding a public hospital needs to be approved by the local Commission of Development and Reform and the Bureau of Housing and Urban-Rural Development, while the local HC's scheme of recruiting talent needs support from the local Bureau of Human Resources and Social Security (Qiu et al. 2020). Without support from other jurisdictional departments, the local healthcare departments and institutions cannot process any day-to-day responsibilities, not to mention the improvement of service performance.

Over the past years and especially after President Xi came into power, the central government has been increasingly highlighting the importance of inter-departmental collaboration within the local bureaucracy (Central Leading Group for Comprehensively Deepening Reforms 2014; State Council 2016; Mu et al. 2019). In particular, Xi's political agenda "Health China 2030" explicitly demonstrates that the provision of healthcare services should not only be undertaken by the healthcare department, but also requires collaborations and supports from other governmental agencies (State Council 2016). Accordingly, a greater level of TED means that other sub-provincial bureaucracies, not just the HCs, are likely to have more responsibilities in local affairs. While similar to the HC, most functional departments in sub-provincial jurisdictions such as the prefecture-level city's Education Bureau and the Environmental Protection Bureau are mainly accountable to the local government (Ma 2017b). Thus, given a greater level of TED, leaders of the sub-provincial government may have greater willingness and capability to perform as "coordinators", i.e., ensuring the duties of the HC can be effectively supported by other local departments (Tsai and Liao 2020; Gao and Zhang 2021). Thus, local HC's costs in inter-departmental communications may be reduced, leading to better service efficiency.

- **RD and healthcare service efficiency**

RD, the third FD indicator, measures the proportion of fiscal revenues shared by the provincial and sub-provincial governments. With the growth of RD, sub-provincial governments and departments have a greater say in the revenue of their local jurisdictions. In this case, as analysed by Brennan and Buchanan (1980), the local government, as a "Leviathan", tends to maximise its economic benefits. In the Chinese context, a higher RD means that sub-provincial cadres control a larger proportion of local revenues and have a larger tax base, so they might pay greater attention to the efficient usage of their locally collected money (Wang and Herd 2013), thus generating greater benefits for themselves. Also, with a larger tax base, to maximise personal benefits in the long term, sub-provincial governments may have a stronger incentive to participate in tax competitions, i.e., initiating policies to reduce the tax burden of individuals and enterprises (Brueckner 2004; Weingast 2009; Choi 2009). Tax competitions can potentially promote economic development and bring greater economic and political capital to local leaders because economic development is still a key consideration for their promotion and other visible benefits. In this case, fiscal revenues and the money given to local healthcare services might be reduced (Qian et al. 2019), which requires local HCs to fulfil their responsibilities with higher efficiency.

Furthermore, apart from the egoistic features, with RD, the top-down exerted pressure for both a better economy and improved healthcare may also mobilise sub-provincial governments to fully take advantage of their greater capacity in managing and allocating local funds, via policy tools such as performance budgeting (Niu 2013; Li 2019). In this case, a certain level of output can be achieved with a lower level of input, leading to an increase in healthcare service efficiency.

7.3 Understanding the negative impact of FD on healthcare service effectiveness: the role of gaming in target setting and implementation under upward accountability

As explained in the above section, sub-provincial governments and functional agencies (such as the city's HC), supported by the TRS and top-down PMS, are accountable to the upper-level government. This relationship of upward accountability ensures that FD – no matter whether the revenue or expenditure side, motivates sub-provincial governments and HCs to work for better healthcare service efficiency. However, empirical analysis cannot confirm the positive impact of HED on healthcare effectiveness, while the other two FD indicators (TED and RD) have negative effects on healthcare effectiveness (measured by a lower PSR). These results might be because the final effectiveness of a public service is more sensitive to socioeconomic factors that the government cannot fully control (Bohte and Meier 2000). For example, effectiveness indicators such as the mortality/survival rates of infants, compared with input-output efficiency, might be more reliant on less controllable factors such as people's lifestyles and habits. But more importantly, these results suggest that compared with efficiency, less attention might be paid by sub-provincial governments and HCs to the effectiveness of healthcare. In particular, the divergent findings for the relationship between FD and effectiveness and efficiency – two major dimensions of PSP indicate the potential co-existence of upward accountability and gaming behaviours in Chinese local governments. Thus, by looking into the gaming behaviours of the decentralised sub-provincial cadres in an institutional context of upward accountability, this section will provide a more detailed discussion about the impact of FD on healthcare service effectiveness in Chinese provinces.

7.3.1 A typology of gaming

Gaming in the public sector refers to strategic actions that aim to bring greater advantages for individuals and/or public organisations in performance evaluation (Bevan and Hood 2006; Kelman and Friedman 2009; Gao 2015; Taylor 2021). Gaming behaviours can be classified into three categories: gaming in target-setting, gaming in work (achieving targets), and gaming in performance reporting (Liu et al. 2021; Taylor 2021).

In the target-setting stage, popular gaming behaviours include choosing favourable targets and setting lower performance goals (i.e., the ratchet effects, see Bevan and Hood 2006; Soss et al. 2011). In the target achievement stage, gaming can be performed by only focusing on the measured targets and criteria (Bevan and Hood 2006; Gibbons 1998), "cutting corners" (e.g. reducing quality, see Bohte and Meier 2000; Gao 2015), or "storming" (e.g. attaining targets with a flurry of radical activities in a short period, see Gao 2021). Finally, in the performance reporting stage, most gaming behaviours are achieved by hiding, manipulating, and even fudging data (Bevan and Hood 2006). Some of the above gaming practices, according to their perniciousness, can be treated as cheating (e.g. reporting fudged data), while others with lower risks (e.g. the ratchet effects) are more likely to be "creative" responses towards the pressure from performance evaluation (Benaine and Kroll 2020; Taylor 2021). Also, it appears that cheating types of gaming usually happen in the reporting stage, while creative and "benign" gaming practices are mostly seen in the stages of setting and achieving targets. In recent years, with the progress in regulations and oversight, in many countries including China, cheating has been increasingly replaced by gaming behaviours in legal and creative forms (Hood 2006). Thus, the following paragraphs will look into the plausibly "legal" gaming practices in the target setting and implementation stages under the institutional background of upward accountability, so as to help understand the impact of FD on healthcare service effectiveness in Chinese provinces.

7.3.2 Intra-provincial FD, gaming in target setting and implementation, and healthcare effectiveness

As demonstrated in the last section, the top-down PMS system in China is associated with high-powered incentives: the results of performance evaluation are directly linked with benefits such as promotion and funding (Ma 2016a). Moreover, relative performance, i.e., the performance of a cadre in comparison with his/her counterparts, is also considered, leading to fierce inter-jurisdictional competition (Edin 2003; Gao 2009; Gao 2015). As explained by de Bruijn (2006), a performance evaluation scheme with high stakes is usually associated with a greater level of motivation for gaming. Such motivation is further strengthened in Chinese sub-provincial jurisdictions where the governments are undertaking increasing roles in most local responsibilities with various targets and performance criteria (Chan and Gao 2009; Yu and Ma 2015).

Despite the presence of strong incentives, it might be difficult to carry out gaming practices without opportunities (Taylor 2021). However, with the TRS, as explained in the last section, the policy preferences and key performance concerns of the upper-level governments are clarified and step-by-step transmitted to lower-level governments, including the territorial governments and functional agencies. This, unintentionally, provides sub-provincial cadres with opportunities to strategically select the most effective way to win performance competitions, that is, prioritising those responsibilities and performance dimensions emphasised more by the upper-level governments when setting targets (Courty and Marschke 2004; Rutherford and Meier 2015; Chen and Jia 2021; Liu et al. 2021). This type of gaming in target-setting at the sub-provincial level might be intensified with FD.

At the revenue side of FD, local governments with greater revenue autonomy (RD) might be reluctant to make efforts towards those responsibilities and performance targets that are relatively ignored by the upper-level government, as their contributions have less impact on the top-down performance evaluation (Gao 2010; Shen et al. 2014; Qian et al. 2019). Instead, through advantages in revenue sharing, local cadres may gain the confidence to highlight the key responsibilities and performance goals favoured by their political leaders when setting targets and eventually achieving them. At the expenditure side, there might be an even stronger connection between FD and the above gaming behaviour: ED brings greater responsibilities, in which case local cadres may have no choice but to focus on performance goals of greatest interest to upper-level governments at the expense of those of less interest (Bohte and Meier 2000; Li 2015).

FD's intensified impact on a specific gaming behaviour may indicate that only those targets and performance goals favoured by upper-level governments are emphasised by sub-provincial cadres when setting targets. For healthcare services, in line with the above analysis, the national and provincial-level governments in China have specific "tastes". That is, compared with quality and effectiveness indicators measuring the final and long-term outcomes of healthcare services, currently, the top-level policymakers in China still pay more attention to healthcare output and efficiency indicators (Ma et al. 2019; Li et al. 2020). For most provincial jurisdictions, the key and enduring issue is to ensure the universal coverage of basic healthcare services with reasonable quality in all areas (Yip et al. 2019), instead of satisfying all performance dimensions. In particular, with the slowing of economic growth and pressure on public finances, increased emphasis has been given to the input-output efficiency of healthcare services. In fact, the Chinese central government has increasingly highlighted that governments at all levels should "tighten the belt" and "truly improve efficiency" (Li 2020), and all unnecessary and "ahead-of-time" spending should be avoided (Gong et al. 2019; Li 2020; Liu 2021).

Apart from the concern with costs, the upper-level governments' preference for output and efficiency indicators is also determined by the technical difficulties of performance evaluation. Compared with effectiveness – a performance dimension affected by various socioeconomic factors, output and input-output efficiency indicators can be more easily clarified and quantitatively measured (Bohte and Meier 2000). Furthermore, effectiveness is usually related to the long-term contributions/outcomes of a specific service, while output and efficiency can be easily evaluated in a short period of time (e.g. by month/year) and thus can be adopted for regular performance evaluation activities (Chan and Gao 2012; Li 2015). Thus, for upper-level government leaders with limited expertise in a specific local public service such as healthcare, output and input-output efficiency indicators appear to be more readable and reliable (Bohte and Meier 2000).

China's upper-level governments' preference for efficiency indicators of healthcare performance has been continuously transmitted to sub-provincial governments and sub-provincial HCs via the TRS. The TRS, along with the top-down PMS, brings greater incentives and opportunities for gaming by highlighting the key targets and performance goals favoured by upper-level superiors. Moreover, as explained in the above section and the Context chapter, the TRS is not just a one-way channel for the top-down delivery of upper-level government's commands, preferences, and goals. Instead, subordinates also have the opportunity to shape the targets and performance goals with local features (Ho 2010; Liu et al. 2021). In other words, a two-way communication channel exists within the TRS, which provides opportunities for local governments to influence key targets and performance goals in line with their personal preferences (Zheng 2007; Ahn et al. 2018; Liu et al. 2021).

Regarding this study, despite the influence from superiors, major targets and performance goals for local healthcare services within a sub-provincial jurisdiction are set by the sub-provincial government and HC (Xiao et al. 2018). As both the target setter and implementor, the local HC, under the pressure of performance evaluation and high-powered incentives (e.g. funding and promotion), tends to prioritise achievable and measurable performance indicators (Kelman and Friedman 2009; Liu 2011). Compared with effectiveness, output and input-output efficiency indicators can be more easily controlled, quantified, and achieved by the healthcare department (e.g. via improving internal management and resource allocation) in a shorter period of time (Li 2015). Instead, there might be more difficulties in achieving healthcare effectiveness, as effectiveness focuses more on the final contribution of a specific service which is more time-consuming and is more sensitive to less-controllable socioeconomic factors such as the demographic structure and people's lifestyle. Also, effectiveness indicators might contribute less to the annual-based performance evaluations as they possibly need several years to be realised. Thus, due to the above “limitations”, local healthcare departments may share a similar taste with their superiors for efficiency and thus may perform gaming practices by highlighting efficiency but pay less attention to effectiveness when setting healthcare targets (Li et al. 2020; Zhu et al. 2022). Moreover, because the personnel evaluation of sub-provincial government leaders, to a large extent, is determined by the fulfilment of local public service targets, gaming behaviours of local HCs and healthcare organisations in highlighting achievable targets might even be tacitly approved by the sub-provincial government (Yip and Hsiao 2009; Ramesh et al. 2014). This is because compared with achieving easier targets, failing to attain difficult targets is a worse thing for both the healthcare department and the top leaders of the government.

In summary, the above arguments suggest that in China, the central government, local governments, and local HCs may have reached a consensus of favouring quantity and efficiency while paying comparatively less attention to the effectiveness of local healthcare services (Meng et al. 2019). From the perspective of sub-provincial governments and healthcare departments (HCs), such preferences with high-powered incentives (i.e., benefits such as promotion, and the measurability of quantity/efficiency indicators) cannot be easily changed with intra-provincial FD. Instead, sub-provincial governments and HCs may take greater advantage of their increased responsibilities and revenue autonomy to perform a specific type of gaming, that is, intentionally satisfying the priorities of their upper-level superiors by highlighting the quantity and efficiency indicators of healthcare performance (Van Dijk and Van Den Ende 2002).

Moreover, with the greater responsibilities authorised via FD, the sub-provincial government and HC will potentially participate more in target and goal setting, which, as suggested by Liu et al. (2021), provide opportunities for them to game by shaping the target list in favour of their personal benefits – that is, highlighting efficiency but ignoring the long-term effectiveness of healthcare services. In China's authoritarian setting, the gaming behaviours of decision-makers cannot be punished at the ballot box by ordinary citizens (Ramesh et al. 2014). Although residential sorting appears to be a plausible mechanism that supports local accountability (Li and Li 2015; Zhang et al. 2017), due to information asymmetry and a lack of expertise, citizens may have little choice but to focus on input and output indicators which are more publicly visible (Ramesh et al. 2014). This, in turn, may encourage healthcare providers to game by intentionally neglecting the less important effectiveness indicators. Also, for healthcare services, as reflected by Sun (2021), the off-site settlement system of Basic Medical Insurance has been significantly improved in recent years and now covers more than 92% of Chinese county-level jurisdictions. In other words, people can easily access medical services in other jurisdictions. This positive change, along with the aforementioned reasons, might undermine the role of residential sorting in motivating local healthcare decision-makers and providers to improve service effectiveness under decentralised settings. In this case, the gaming behaviour of emphasising efficiency but ignoring effectiveness might be intensified. To have a better understanding of the above arguments, the following paragraphs provide more detailed explanations of the impact of 3 FD indicators (HED, TED, RD) on the effectiveness of healthcare services in Chinese provinces.

- **HED and healthcare service effectiveness**

As clarified in the above sections and chapters, a greater value of HED means that more responsibilities in local healthcare services have been transferred to sub-provincial HCs. In this situation, under greater pressure from the top-down performance evaluation, the fierce “promotion championship”, and the competition for other high-powered incentives such as extra funding, a sub-provincial HC may gain a stronger incentive to game by prioritising targets favoured by their superiors as well as those that are more achievable and measurable. These preferred targets, as explained above, are likely to be quantity and efficiency indicators. Such gaming practices, supported by the advantages of utilising local resources, managing local organisations, and initiating local innovations under a greater level of HED, lead to better healthcare service efficiency. As a result, the effectiveness of healthcare may be comparatively neglected in the aforementioned gaming practices, which means that even if there is a high level of healthcare decentralisation, the sub-provincial HC’s local advantages may not necessarily lead to better effectiveness (Jin and Sun 2011; Brock et al. 2015).

- **TED and healthcare service effectiveness**

An increase in TED means that the sub-provincial government and its functional agencies, as a whole, need to undertake a greater role in all local affairs. On the one hand, such inter-governmental arrangements ensure the sub-provincial cadres can fully utilise their local advantages in fulfilling their responsibilities. On the other hand, it brings greater burdens to the sub-provincial bureaucracy, which may exacerbate the gaming practice of emphasising those targets and goals favoured by their superiors but ignoring others. Accordingly, in terms of the performance of local healthcare services, the sub-provincial government leaders, in line with their superiors, may consider healthcare effectiveness as a less important issue and place it at a lower position in their “to-do” list. This explains the negative relationship between TED and healthcare effectiveness. Moreover, a decentralised sub-provincial government may have greater capability to foster coordination between different departments as well as between different targets. However, if the responsibilities and targets are in irreconcilable conflict – which in fact is an enduring problem for the public sector (Boyne 2003; Nielsen 2014), the sub-provincial government leaders might tacitly tolerate a kind of gaming practice that achieves those key/favourable targets at the expense of others (Rutherford and Meier 2015; Chen and Jia 2021). Such “pernicious gaming” also explains why TED is positively related to healthcare efficiency but leads to worse healthcare effectiveness.

- **RD and healthcare service effectiveness**

A greater level of RD means that sub-provincial governments and departments, in comparison to their provincial leaders, have a greater say over their local revenues. However, with the top-down PMS and the preferences for quantity and efficiency indicators, despite a high level of RD, sub-provincial cadres may still be reluctant to promote healthcare effectiveness with their greater discretionary power on the revenue side. Moreover, in the current context of fiscal tightening (Li 2020; Liu 2021), the local government might be motivated to allocate their funds to money-generating and money-saving activities (Shen et al. 2014; Qian et al. 2019). In this case, as shown by the empirical results of this study, the long-term effectiveness of healthcare services might be sacrificed.

7.4 Effects of relative wealth on the relationship between FD and healthcare service performance

The above two sections provide a detailed discussion regarding the effects of FD on healthcare service efficiency and effectiveness. Nevertheless, relative wealth might bring moderating effects on the relationship between FD and healthcare service performance. That is, the impact of FD on healthcare performance might be greater in a wealthier region. As explained in the Methodology chapter, the level of relative wealth determines the resources that can be employed by a government to build up its capacity (van den Bergh 2009). This capacity, as defined by Ingraham and Donahue (2000, p. 294), refers to the ability to produce, marshal, direct, and control local human, physical, and information capital to achieve policy targets. Clearly, in a wealthier jurisdiction, local governments tend to have a greater capacity to perform their policy targets, preferences, and performance goals. As indicated by Prud'homme (1995), FD's benefits on public services still depend, at least partially, on the local government's capacity in dealing with local demands and resources. If local governments have significant disadvantages in terms of this capacity, FD might even bring negative effects on PSP (Tanzi 1995; Bello-Gomez 2020). Given that China is still a developing country with huge regional disparities in relative wealth, it is hypothesised that the benefits of FD on healthcare service performance are greater in those wealthier regions where local governments and the healthcare sector have a higher capacity to achieve their policy targets and performance goals.

In this research, a part of the above arguments is empirically confirmed. Both random and fixed effects analysis find that with the increase of relative wealth, the benefits of HED (healthcare decentralisation) on healthcare efficiency become greater. This finding suggests that a higher level of relative wealth may indicate sub-provincial HCs' greater capacity to manage local healthcare resources and initiate local healthcare innovations. This higher capacity thus directly supports sub-provincial HCs to perform better in promoting service efficiency with a given level of healthcare decentralisation (HED). However, relative wealth appears to make no difference to the TED-efficiency relationship. As discussed above (Barankay and Lockwood 2007; Kang et al. 2012), TED is an aggregated FD indicator – in a province, the increase of

TED also means the sub-provincial governments and functional agencies, as a whole, get greater responsibilities in their own fields. Accordingly, the benefits of higher capacity brought by more relative wealth might be distributed evenly across all sub-provincial agencies, rather than being dominantly given to the healthcare sector. Once other sub-provincial agencies are decentralised with greater responsibilities, they may utilise their greater capacity to work better on their key responsibilities and targets which are less relevant to healthcare. Even though the performance improvement in other services might ultimately contribute to the efficiency of local healthcare departments, such benefits might be indirect and less foreseeable.

Relative wealth also appears to have no moderating effect on the RD-efficiency relationship. Greater RD means the sub-provincial government have more say over the local money for localised responsibilities including but not limited to healthcare (Wang 2010; Gao et al. 2014). Accordingly, with the growth of local wealth, the sub-provincial government's better capacity in managing local funds and resources will apply to all responsibilities accomplished by local agencies, while only a limited part of such benefits would go to healthcare services and the sub-provincial HCs. In this case, the incremental impact of RD on healthcare efficiency along with the growth of relative wealth (capacity) might be negligible. Moreover, in a wealthier region, despite the greater capacity of the local government, RD may be more susceptible to the soft budget constraint, resulting in various problems such as corruption and the waste of funding. In other words, the growth of relative wealth may fail to help RD contribute more to healthcare efficiency (Tanzi 1995; Qian and Roland 1996; Fisman and Gatti 2000).

There is no evidence to support the presence of the moderating effect of relative wealth on the FD-effectiveness relationship. For moderating effects on the EDT-effectiveness and RD-effectiveness relationships, these insignificant results still could be explained by the fact that the decentralisation of total responsibilities and revenue autonomy would support sub-provincial governments and agencies to work better on all local services, not just healthcare. Moreover, the insignificant effect of relative wealth on the relationship between FD – no matter what indicators – and healthcare effectiveness could also be attributed to the institutional arrangements for the Chinese-style FD and the gaming behaviours under a decentralised setting. That is, decentralising healthcare responsibilities (HED), total responsibilities (TED), and revenues (RD) would not motivate sub-provincial cadres to use their stronger capacity (measured by relative wealth) to work better on those performance dimensions that have limited contributions to their visible benefits such as promotion, bonus, and extra funding. Particularly for sub-provincial HCs, as explained above, healthcare effectiveness compared with efficiency is relatively ignored by the upper-level governments, and sub-provincial HCs also have greater difficulties in achieving and measuring effectiveness responsibilities. Thus, even if capacity grows, decentralised sub-provincial HCs may not utilise that capacity to improve healthcare effectiveness.

7.5 Summary

This chapter aims to discuss the empirical results by drawing on theoretical literature and the viewpoints of real-world practitioners. It is found that in Chinese provinces, the behaviours of sub-provincial governments and healthcare agencies under FD settings seem to be deeply affected by the upward accountability mechanism. That is to say, despite a high level of FD, with upward accountability, local cadres are still influenced by their superiors at higher administrative levels. Such relationships of upward accountability in China rely on two key underpinnings: the TRS and the top-down PMS. TRS helps upper-level governments clarify and communicate their key policy agendas and preferences downward to local jurisdictions (Chan and Gao 2009; Burns and Zhou 2010), while the top-down PMS links the fulfilment of the above agendas and preferences with high-powered incentives such as promotion and funding (Chien 2010). Thus, with FD, subordinates are pushed to make greater efforts on the key agendas, preferences, and performance requirements of their superiors, supported by their FD-functionalised local advantages in various aspects, such as utilising local resources, initiating local policies, and promoting inter-departmental collaborations.

In terms of healthcare, this chapter as well as the background chapter of this thesis highlighted that the national and provincial leaders of the Chinese government and the CPC, for the sake of regime legitimacy, have been paying increasing attention to the performance of healthcare services (Ratigan 2022; Duckett and Munro 2022). In particular, given the lack of basic healthcare products in underdeveloped regions and evaluation difficulties, the quantity of output and input-output efficiency for basic healthcare services are still a key concern of China's top policymakers. This emphasis on healthcare efficiency is communicated via the TRS to sub-provincial governments and HCs and is underpinned by the top-down PMS. Supported by FD, the sub-provincial governments and HCs fully utilise their local advantages to achieve improvements in healthcare service efficiency. Moreover, the higher capacity brought by the growth of relative wealth can help sub-provincial healthcare agencies improve service efficiency to a greater extent at a given level of decentralisation for healthcare (HED).

However, the strong pressure from the top-down PMS and lower-level cadres' wide participation in TRS, unintentionally, appears to provide incentives and opportunities for sub-provincial cadres to game in target setting and performing processes (Chen and Jia 2021). For this research, two types of gaming behaviours may be important. First, sub-provincial government and healthcare leaders may deliberatively prioritise those key targets favoured by their upper-level superiors over those perceived to be less important. Second, sub-provincial governments and HCs, via TRS, could shape the target list by increasing the weight of achievable and measurable targets. As FD provides sub-provincial governments and the healthcare sector with greater local responsibilities/revenue autonomy and more opportunities to participate in TRS, the above two gaming practices might be intensified. Both senior policymakers and sub-provincial cadres in China, as explained in this chapter, may pay greater attention to efficiency than effectiveness. Thus, in a decentralised setting, the gaming behaviour of favouring healthcare efficiency but relatively ignoring effectiveness might be exacerbated. Even the higher capacity brought by the growth of relative wealth fails to motivate the decentralised sub-provincial government and HCs to work harder on improving healthcare effectiveness.

In summary, both FD's benefits for healthcare efficiency and its' apparent costs for healthcare effectiveness, in essence, are the results of China's upward accountability system, located in a political context lacking mechanisms for healthcare service users (i.e., citizens) and other external stakeholders (e.g. media and NGOs) to have an influential say in local issues (Lam 2010). Although residential sorting appears to be a plausible mechanism for people to impact on the local government (Chen 2014), people's foot-voting decisions, in practice, can be easily distorted by information asymmetry and the lack of expertise in understanding various healthcare performance indicators. Also, the off-site settlement system of the Basic Medical Insurance potentially undermines the function of residential sorting in motivating healthcare policy-makers and providers (Sun 2021). Therefore, unless people's demand for better healthcare outcomes can be fully articulated and become a key concern of China's top leadership, it might be difficult for FD to truly promote the effectiveness of healthcare services under the current arrangement of upward accountability.

CHAPTER 8 CONCLUSION

Focusing on the background of Chinese provinces, this doctoral study investigated the impact of intra-provincial FD on healthcare efficiency and effectiveness – two key aspects of healthcare service performance. In addition, the impact of relative wealth – a proxy for governance capacity – on the relationship between FD and healthcare efficiency and effectiveness was explored. With the theoretical specifications and empirical results applied here, this study is expected to bring contributions to both academic and non-academic readers.

8.1 Academic contributions

For academic readers, the contributions of this doctoral study are centred on the empirical and theoretical sides. Empirical contributions revolve around the addressed research gaps. Firstly, this research addressed an obvious gap in the existing literature that hardly any empirical studies focused on the relationship between FD – especially FD between the provincial and sub-provincial level – and healthcare service performance in Chinese provinces. Moreover, the special attention of this research on healthcare can be seen as a remedy for the previous FD-PSP empirical studies which are mainly concerned with the FD's effect on the performance (particularly efficiency) of all major services but lack deep insight into the impact of FD on a single but important public service. Also, by discussing the impact of FD on both healthcare efficiency and effectiveness, the lack of a simultaneous investigation of FD's impact on multiple performance dimensions in existing studies is addressed. Accordingly, with the aforementioned gaps being covered, various empirical contributions have been brought to the field of FD-PSP research.

In addition to its empirical contributions, the analysis and empirical findings of this research could contribute to the ongoing debates surrounding the NPM, FD, and the enhancement of PSP. First, this study focused on FD within the unique context of China and found positive relationships between intra-provincial FD – measured from the expenditure and revenue sides – and healthcare efficiency. These findings provide real-world evidence that FD, as an NPM-inspired reform, could improve public service efficiency – one of the key concerns of NPM theories and research (Boyne et al. 2003).

Furthermore, the second part of this research demonstrated a negative relationship between FD and healthcare service effectiveness. This result aligns with the growing theoretical and practical critics regarding the future of NPM. Over the past years, the value of NPM has been increasingly debated. Researchers argue that traditional NPM theories excessively prioritise economic dimensions such as input, quantity, and cost-efficiency, which leads to the neglect of broader outcomes such as the formal objectives of public service provision, equity among different social groups, and democratic values like citizenship and procedural justice (Balfour and Grubbs 2000; Lynn 2001; O'Flynn 2007; Çolak 2019). With the negative findings regarding FD and healthcare effectiveness, this thesis provided timely justifications for these critical discussions in the current NPM theoretical landscape. This could contribute to the growing call for the evolvement of NPM[31] theories towards a more universal performance framework that incorporates more considerations of public value (Bryson et al. 2014; Anderson et al. 2016).

Secondly, this study contributed to the federalism theory, which serves as the theoretical foundation for FD practices. Specifically, this study comprehensively reviewed the key explanations of FD in China and, at the sub-provincial level, justified the generalisability of the de facto fiscal federalism theory (Zheng 2007) in conceptualising the key features of the Chinese-style FD. Regarding the relationship between FD and PSP, most previous theories developed in the western context (e.g. Oates 1972; Qian and Weingast 1997) suggest that the aim of winning in the competitions for ballots and population inflow motivates the decentralised government to promote performance. However, given the sharp socio-political differences, such western-based analysis cannot be directly applied to support research based in China. Thus, the third theoretical contribution of this doctoral study is the clarification of the key underpinnings of the Chinese-style FD that motivate sub-provincial cadres to improve their performance, which refer to upward accountability, the target-setting system (TRS) and the top-down PMS. With these

[31] Some researchers consider those critics towards NPM as post-NPM theories (e.g., Reiter and Klenk 2019). However, as argued by Lapuente and Van de Walle (2020), there is no substantial difference between these two concepts. Post-NPM is more like a complement rather than an alternative to NPM.

underpinnings, cadres at each administrative level are held accountable to the government located one administrative level above. Meanwhile, the higher-level government is able to communicate their key agendas and requirements about healthcare to the local governments (via TRS) and regulates them with performance-linked cadre evaluation (i.e. the top-down PMS).

Furthermore, supported by empirical findings, this doctoral research discussed the decentralised cadres' diverse behaviours towards different dimensions of performance, which could be insightful for the existing PSP literature. As analysed, intra-provincial FD provides more space for sub-provincial governments and HCs to fully utilise their advantages in managing and utilising local information, resources, and organisations. However, these advantages primarily work the key targets highlighted by the higher-level government, which, for healthcare services, refer to efficiency. For less important targets such as effectiveness, the decentralised local cadres are reluctant to address it, as it is less related to their personal benefits. Moreover, they may game the performance management rules by prioritizing the "achievable" targets and key requirements of their superiors at the expense of the less important ones, which leads to a deterioration of healthcare service effectiveness in Chinese provinces.

Finally, the analysis of moderating effects shows that the greater the capacity of the sub-provincial healthcare sector (measured by the level of relative wealth – per capita

GDP), the more benefits healthcare decentralisation would have for healthcare efficiency. This finding contributed to FD theories by empirically justifying the theoretical argument that the benefits of FD are contingent upon the extent of the local government's relative advantage (Prud'homme 1995; Arends 2017). Previous empirical studies often took the local advantages for granted, whereas this thesis illustrated that if these advantages are undermined (indicated by a lower level of capacity), no matter how strong the motivation from performance-based cadre management (or elections and residential sorting) is, FD will have limited benefits for public service efficiency[32].

8.2 Practical contributions

Apart from influencing the research field of FD and PSP, this study can also bring practical contributions by informing public policymakers and practitioners with the following evidence-based recommendations:

First, recognising the positive relationship between intra-provincial FD on healthcare efficiency, provincial governments valuing improved resource utilisation should continue to decentralise revenues and responsibilities to sub-provincial governments and healthcare departments. This can enable the sub-provincial bureaucracies to fully utilise their advantages in managing local resources and demands. Thus, healthcare services can be provided in more cost-efficient ways.

[32] For healthcare effectiveness, as explained, this is a less important target. Therefore, no matter how strong/weak the capacity/local advantage is, local cadres would not work better/worse on it with a certain degree of FD.

Second, recognising the positive role of government capacity in enhancing the FD-efficiency relationship, sub-provincial governments – particularly those located in underdeveloped areas should continue to promote economic development as a means to building their capacity for managing local public organisations and utilising local resources. In doing so, FD could make greater contributions to healthcare efficiency.

Third, empirical findings indicate a negative effect of FD on healthcare service effectiveness. Based on the analysis of the upward accountability mechanisms, senior policymakers at the national and provincial levels could address this issue by paying greater attention to healthcare effectiveness in their target-setting and performance-management practices. On the one hand, this research acknowledges that with the system of upward accountability, gaming behaviour at the sub-provincial level, i.e., satisfying the key requirements/agendas of the superiors but ignoring the less important ones, cannot be eradicated. On the other hand, the optimal and presumably more applicable option for senior policymakers is to treat healthcare efficiency and effectiveness as equally as possible. An increasing emphasis on healthcare effectiveness could inform sub-provincial governments and HCs via the two existing mechanisms of upward accountability – TRS and the top-down PMS. In this case, FD would be more likely to improve both healthcare efficiency and effectiveness.

8.3 Limitations

Despite the academic and practical contributions, this doctoral study still suffers from limitations which could be addressed by future research. The first limitation relates to the indicators of FD. As explained in the Context chapter, nowadays in China, responsibilities in formulating and implementing policies in healthcare and other major local services have been largely decentralised to the sub-provincial governments (Donaldson 2017b), which can be reflected by the three FD indicators of this doctoral research. However, another popular indicator to measure intra-provincial FD named expenditure autonomy (Psycharis et al. 2016; Alonso and Andrews 2019), i.e., the percentage of sub-provincial healthcare/total expenditure covered by the provincial transfers, is not adopted in this research due to the lack of data (particularly the data for provincial healthcare transfers). In recent years, to ensure that transfer grants can be flexibly allocated to fulfil local demands and responsibilities, higher-level governments in China usually avoid directly interposing the usage of transfer grants (Dollar and Hofman 2006; Niu 2013; Huang et al. 2017). Nevertheless, a higher ratio of expenditure autonomy, to some extent, still indicates that the provincial government may have a certain say in sub-provincial policies and thus partially reflects the level of decentralisation. If data for this indicator were accessible, more robustness checks would be performed to improve the validity of the main findings of this research.

The second limitation relates to the indicators of effectiveness which represent the attainment of the formal objective of healthcare services (Boyne 2002). As justified in the Methodology chapter, following most previous studies (e.g. Habibi et al. 2003; Asfaw et al. 2007; Kang et al. 2012), healthcare effectiveness is measured by the perinatal survival rate (PSR, calculated based on the perinatal mortality rate PMR). To the researcher's knowledge, for now, this is the most applicable choice. Other indicators such as life expectancy, compared with PSR, rely more on estimated or modelled data (e.g. people's assumed lifespan) and thus potentially suffer more from data manipulation (Jiménez-Rubio and Smith 2005). Also, there is a lack of yearly data for LE in Chinese provinces. However, although to a large extent, PSR is a comprehensive indicator of the overall effectiveness of the healthcare sector, it is still likely that discrepancies exist between what is reflected by PSR and the effectiveness of other healthcare activities. Thus, if data are accessible, more indicators of effectiveness, e.g. hospital readmission rate and surgery mortality rate (Quentin et al. 2019), should be adopted.

The third limitation is about the measurement of relative wealth, which refers to per capita GDP. It should be noticed that the positive effects of this variable on the relationship between FD and efficiency may be underestimated, because the three economically developed province-level municipalities[33] with high levels of per capita GDP were excluded from the dataset. Furthermore, the use of relative wealth as a

[33] They refer to Beijing, Shanghai, and Tianjin, while Chongqing is considered relatively underdeveloped compared to the above three municipalities and other ordinary provinces.

proxy for government capacity is not without its criticisms. For example, in resource-dependent countries and regions, a high per capita GDP may coexist with low-capacity public sector agencies (Renzio et al. 2005; Barma et al. 2012). In this case, relative wealth may not accurately proxy for the level of government capacity. Although such wealthy and resource-dependent provinces do not exist in China, it is still valuable to consider supplementary indicators such as the educational level of local cadres when data becomes accessible. Moreover, with accessible dataset, integrated indices developed by international organisations, such as the Quality of Government Institute's capacity index and the World Economic Forum's public institutional factor, are also inspirational for measuring government capacity in China (Honadle 2001; Im and Choi 2018).

Finally, this research follows previous public sector studies (e.g. Balaguer-Coll et al. 2007; Fonchamnyo and Sama 2016) to evaluate productive efficiency with the BCC-DEA approach and super-efficiency DEA. The advantages of such non-parametric approaches over parametric approaches such as the SFA have been explained in the Methodology chapter. Nevertheless, other DEA-based techniques, particularly the Malmquist index which decomposes the variations of productivity into changes in pure technical efficiency, scale efficiency, and technology improvement (Wang and Lan 2011), could be employed to improve the robustness of evaluation results.

In terms of data estimation, fixed-effect, Tobit, and random-effect analysis are adopted, which are considered as robust for econometric estimations and have been widely used in previous studies. However, it is important to acknowledge the potential presence of endogeneity and its implications on the significance of the coefficients. While controlling for the time effects in this research helps to mitigate endogeneity caused by missing variables, incorporating a broader set of socio-economic control variables would further address this limitation (Wooldridge 2001). Also, regarding the endogeneity arising from the reverse causality between healthcare performance and FD, theoretical discussions in this thesis suggest a limited possibility of such a problem. This is because the higher-level governments in China tend to punish local cadres who failed to improve PSP through the top-down PMS, rather than directly withdrawing the decentralised power. Nevertheless, this limitation still could be better addressed by employing suitable instrumental variables and appropriate estimation approaches such as the generalized method of moments (GMM) (Baum et al. 2002).

8.4 Directions for further research

Recognising the aforementioned contributions and limitations, several directions for further research are identified. First, as mentioned above, future research should focus on the impact of intra-provincial FD on other public services in China. Accordingly, the impact of FD on the performance of all major local services could also be investigated. In this case, the two indicators TED and RD which measure the responsibilities and revenues of the sub-provincial government and agencies as a whole could play a key role.

Second, this doctoral study revolves around healthcare efficiency and effectiveness – two key aspects of PSP highlighted in the '3Es' model and the IOO model (Boyne 2002). However, the relationship between FD and other user-based PSP indicators, such as citizen satisfaction, long-term impact, and the equity of service provision among people in different social groups, should be discussed in future studies. Especially, when looking into the perceptions of service users, data can be collected by the highly structured survey approach and the unstructured interview approach. Thus, not only quantitative methods can be adopted to investigate causal effects, but also research questions such as people's experiences in a decentralised setting can be addressed by qualitative approaches.

Third, this doctoral research focuses on the background of Chinese provinces to discuss the impact of intra-provincial FD on healthcare performance. FD in this study measures the extent to which responsibilities and revenues are decentralised from the provincial to the sub-provincial level (i.e., prefecture-level cities) where three tiers of governments are included: the prefectural level, the county level, and the township level. In the above three tiers of governments, the township governments are directly controlled by the counties above them (Donaldson 2017b) and thus could be considered as 'branches' of the county authorities. However, a certain level of decentralisation exists between the other two tiers, i.e., the prefecture-level and county-level governments. While this research mainly focuses on the FD between the provincial government and the two lower tiers of government due to the lack of county-level statistical data, future studies could perform deeper investigations into the FD between the prefectural level and county level once data are accessible. In this way, to which tier (i.e., prefecture or county) and to what degree FD can benefit the performance of healthcare and other services can be further clarified.

Furthermore, by focusing on the healthcare sector in smaller jurisdictions at the prefectural or county level, further research could investigate into the effects of a specific healthcare reform or policy innovation on the performance of the local healthcare sector and/or healthcare institutions under a decentralised setting. This nationwide PhD research did not look into this aspect due to the absence of nationally unified healthcare reform across provinces. However, with the healthcare innovations being continuously initiated by the decentralised healthcare authorities, pre-post analysis of such policy interventions in small jurisdictions would be of timely importance. This could be achieved by methods such as the Difference-in-Difference estimation (Abadie 2005).

Finally, due to the inaccessibility of the newest data, this research focuses on the impact of FD on healthcare performance from 2006 to 2017. At the beginning of 2020, China became the first epicentre of the COVID-19 pandemic, which, in the following two years, has been posing severe challenges to China and the world.

Although at the beginning of the outbreak, the Chinese government was criticised for its poor response and the lack of transparency, after the early-stage chaos, China has been effectively containing the pandemic in a short period. Over the past three years, despite a high vaccination rate, China strictly followed a zero-tolerance policy (until late December 2022), while most countries around the world have already largely released their covid-19 restrictions (Burki 2022). In such a context, it is of timely importance to empirically discuss the impact of the pandemic on the Chinese-style FD as well as the impact of the Chinese-style FD on healthcare performance during the pandemic.

For the first question, there are several studies suggesting that the top-down pattern of decentralisation in China enabled the central government to take back the power from local governments at the early stage of the outbreak, so as to impose nationwide covid-containing policies and ensure effective mobilisation of resources (e.g. Xing et al. 2021; Qu and Lv 2021; Lv et al. 2022). For the second question, it appears that the impact of FD on healthcare performance in the pandemic setting could still be explained with the theory of upward accountability highlighted in this research. That is, after the first wave of the pandemic, the role of regular COVID-19 prevention and control has been taken by sub-provincial governments (Gao et al. 2020). However, as the top leadership of the Chinese government and CPC still favours 'zero tolerance', decentralised local governments have to be fully accountable for this requirement by locating the containment of COVID-19 as a central policy and adopting measures which are more radical than those recommended by the higher-level governments (Mei 2020). This behaviour, on one hand, ensures the control of the pandemic – fulfilling the top agenda of China's top party-state leadership. On the other hand, it threatens the provision of other healthcare and social services (Xiao et al. 2021). For example, patients requiring hemodialysis have not been treated on time due to the covid-19 restrictions and the lack of medical resources for non-COVID-19 services (Yuan 2022; Zhou et al. 2022). Such a phenomenon reflects the potential negative effects of FD on the overall level of healthcare performance under COVID-19.

In summary, the above two questions regarding FD and the impact of FD on healthcare performance during the pandemic, seemingly, could be explained by the key features of the Chinese-style FD summarised in this doctoral research. Moreover, in the final months of 2022, the negative effects of FD on the healthcare sector and other socioeconomic aspects became increasingly unacceptable to the public and led to nationwide protests, which strongly threatened the legitimacy of the Communist party. In response to such an unprecedented challenge, the top leadership of China finally decided to release the zero-covid policy. After that, people have been witnessing various initiatives employed by local governments to recover the economy and normal healthcare services. This dramatic change, as explained by the upward accountability theories, could be seen as a natural behaviour of the decentralised local governments to satisfy the new requirement of their superiors. Nevertheless, deeper empirical research would be useful to provide tenable evidence and interpretations regarding the ongoing impact of FD on the performance of healthcare and other services in the post-covid era.

REFERENCES

Abadie, A. 2005. Semiparametric Difference-in-differences estimators. *The Review of Economic Studies* 72(1), pp. 1-19. doi: 10.1111/0034-6527.00321

Abimbola, S., Baatiema, L. and Bigdeli, M. 2019. The impacts of decentralization on health system equity, efficiency and resilience: a realist synthesis of the evidence. *Health Policy and Planning* 34, pp. 605-617. doi: 10.1093/heapol/czz055

Adam, A., Delis, M. D. and Kammas, P. 2014. Fiscal decentralization and public sector efficiency: evidence from OECD countries. *Economics of Governance* 15(1), pp. 17-49. doi: 10.1007/s10101-013-0131-4.

Adler, N., Friedman, L. and Sinuany-Stern, Z. 2003. Review of ranking methods in the data envelopment analysis context. *European Journal of Operational Research* 140(2), pp. 249-265. doi: 10.1016/S0377-2217(02)00068-1

Afonso, A. and Fernandes, S. 2008. Assessing and explaining the relative efficiency of local government. *The Journal of Socioeconomics* 37, pp.1946-1979. doi: 10.1016/j.socec.2007.03.007

Afonso, A. and Hauptmeier, S. 2009. Fiscal behaviour in the European Union: rules, fiscal decentralization and government indebtedness. *ECB Working Paper, No. 1054*. Frankfurt, Germany: European Central Bank (ECB).

Afonso, A. and Aubyn, M. 2006. Cross-country efficiency of secondary education provision: a semi-parametric analysis with non-discretionary inputs. *Economic Modelling* 23, pp. 476-491. doi:10.1016/j.econmod.2006.02.003

Afonso, A., Schuknecht, L. and Tanzi, V. 2006. Public sector efficiency: evidence for new EU member states and emerging markets. *ECB Working Paper No. 581*. Frankfurt: European Central Bank (ECB).

Ahmad, E. and Craig, J. 1997. Intergovernmental transfers. In: Ter-Minassian, T. ed. *Fiscal federalism in theory and practice*. Washington: International Monetary Funds, pp. 73-107.

Ahmad, E., Hewitt, D. and Ruggieero, E. 1997. Assigning expenditure responsibilities. In: Ter-Minassian, T. eds. *Fiscal federalism in theory and practice*. Washington: International Monetary Funds, pp. 25-48.

Ahmad, E., Brosio, G. and Tanzi, V. 2008. Local service provision in selected OECD countries: do decentralized operations work better? *IMF Working Paper 08/67*. Washington: International Monetary Funds.

Ahn, T.S., Choi, Y.S., Hwang, I. and Hyeon, J. 2018. The effect of information asymmetry and participative target setting on target ratcheting and incentives. *Public*

Performance and Management Review 41(3), pp. 439–468. doi: 10.1080/15309576.2018.1456940.

Akin, J. S., Dow, W. H., Lance, P. M., and Loh, C. P. A. 2005. Changes in access to health care in China, 1989–1997. *Health Policy and Planning* 20(2), pp. 80–89.

Alesina, A. 1999. Too large and too small governments. In: Tanzi, V., Chu, K., and Gupta, S. Eds. *Economic policy and equity*. Washington, DC: International Monetary Fund, pp. 216–234.

Alfada, A. 2019. Fiscal decentralization and government expenditure efficiency in Indonesia: a Malmquist Productivity Index. *Expert Journal of Economics* 7(1), pp. 45-57.

Allin, S., Veillard, J., Wang, L. and Grignon, M. 2015. How can health system efficiency be improved in Canada? *Healthcare Policy* 11(1), pp. 33–45.

Alonso, J. M. and Andrews, R. 2019. Fiscal decentralisation and local government efficiency: does relative deprivation matter? *Environment and Planning C: Politics and Space* 37(2), pp. 360–381. Doi: 10.1177/2399654418784947.

Alonso, J.M., Clifton, J. and Díaz-Fuentes, D. 2015. Did new public management matter? an empirical analysis of the outsourcing and decentralization effects on public sector size. *Public Management Review* 17(5), pp. 643–660. doi: 10.1080/14719037.2013.822532.

Andersen, L.B., Boesen, A. and Pedersen, L.H. 2016. Performance in public organizations: clarifying the conceptual space. *Public Administration Review* 76, pp. 852–862. doi: 10.1111/puar.12578

Andre, R. 2010. Assessing the accountability of government-sponsored enterprises and quangos. *Journal of Business Ethics* 97(2), pp. 271–89.

Andrews, R. and Entwistle, T. 2010. Does cross-sectoral partnership deliver? An empirical exploration of public service effectiveness, efficiency, and equity. *Journal of Public Administration Research and Theory* 20(3), pp. 679-701.

Andrews, R. and Entwistle, T. 2013. Four faces of public service efficiency. *Public Management Review* 15(2), pp. 246-264. doi: 10.1080/14719037.2012.725760.

Andrews, R. and Martin, S. 2010. Regional variations in public service outcomes: the impact of policy divergence in England, Scotland and Wales. *Regional Studies* 44(8), pp. 919-934. doi: 10.1080/00343400903401592

Andrews, R., Beynon, M. J. and McDermott, A. 2019. Configurations of New Public Management reforms and the efficiency, effectiveness and equity of public healthcare systems: a fuzzy-set qualitative comparative analysis. *Public Management Review* 21(8), pp. 1236-1260. doi: 10.1080/14719037.2018.1561927

Andrews, R., Boyne, G. A., Moon, M. J. and Walker, R. M. 2010. Assessing organizational performance: exploring differences between internal and external measures. *International Public Management Journal* 13(2), pp.105-129. Doi: 10.1080/10967491003766533

Andrews, R., Boyne, G.A. and Walker, R.M. 2006. Subjective and objective measures of organizational performance: an empirical exploration. In Boyne, G.A., Meier, K.J., O'Toole, L.J. Jr. and Walker, R.M. eds. *Public service performance: perspectives on measurement and management.* Cambridge, UK: Cambridge University Press, pp. 14-34.

Andrews, R. and van de Walle, S. 2013. New public management and citizens' perceptions of local service efficiency, responsiveness, equity and effectiveness. *Public Management Review* 15(5), pp. 762–783. doi: 10.1080/14719037.2012.725757.

Angrist, J., Bettinger, E. and Kremer, M. 2006. Long-term educational consequences of secondary school vouchers: evidence from administrative records in Colombia. *The American Economic Review* 96(3), pp. 847-862

Arcelus, F.J., Arocena, P., Cabasés, F. and Pascual, P. 2015. On the cost-efficiency of service delivery in small municipalities. *Regional Studies* 49(9), pp. 1468-1480. doi: 10.1080/00343404.2013.837872

Arends, H. 2017. More with less? Fiscal decentralisation, public health spending and health sector performance. *Swiss Political Science Review* 23(2), pp. 144-174. doi: 10.1111/spsr.12242.

Arrow, K. J. 1971. Political and economic evaluation of social effects and externalities. In: Intriligator, M. D. eds. *Frontiers of quantitative economics*. Amsterdam: North Holland, pp. 3–23.

Asfaw, A., Frohberg, K. James, K. S. And Jutting, J. 2007. Fiscal decentralization and infant mortality: empirical evidence from rural India. *The Journal of Developing Areas* 41(1), pp. 17-35. doi: 10.1353/jda.2008.0026.

Aucoin, P. and Heintzman, R. 2000. The dialectics of accountability for performance in public management reform. *International Review of Administrative Sciences* 66(1), pp. 45-55. doi: 10.1177/0020852300661005.

Azfar, O., Kahkonen, S., and Meagher, P. 2001. Conditions for effective decentralized governance: A synthesis of research findings. *IRIS Center Working Paper No. 256.* College Park, MD: University of Maryland.

Bahl, R. and Wallace, S. 2007. Intergovernmental Transfers: the vertical sharing dimension. In Martinez-Vazquez, J. and Searle, B. eds. *Fiscal Equalization: challenges in the design of intergovernmental transfers*. New York: Springer, pp. 205-249.

Bahl, R., Martinez-Vazquez, J. and Wallace, S. 2002. State and local government choices in fiscal redistribution. *National Tax Journal* 55(4), pp. 723-742. doi: 10.17310/ntj.2002.4.04.

Bai, Z. and Liu, J. 2020. China's governance model and system in transition. *Journal of Contemporary East Asia Studies* 9(1), pp, 65-82. doi: 10.1080/24761028.2020.1744229

Baicker, K. 2001. The spillover effects of state spending. *NBER Working Paper No. 8383*. Cambridge, MA: National Bureau of Economic Research.

Balaguer-Coll, M.T., Prior, D. and Tortosa-Ausina, E. 2007. On the determinants of local government performance: a two-stage nonparametric approach. *European Economic Review* 51(2), pp. 425-451. doi: 10.1016/j.euroecorev.2006.01.007.

Balaguer-Coll, M.T., Prior, D. and Tortosa-Ausina, E. 2010. Decentralization and efficiency of local government. *The Annals of Regional Science* 45(3), pp. 571-601. doi: 10.1007/s00168-009-0286-7

Balfour, D.L. and Grubbs, J.W. 2000. Character, corrosion and the civil servant: the human consequences of globalization and the new public management. *Administrative Theory & Praxis* 22(3), pp. 570-584. doi: 10.1080/10841806.2000.11643472

Balli, H.O. and Sørensen, B.E. 2013. Interaction effects in econometrics. *Empirical Economics* 45, pp. 583–603. doi: 10.1007/s00181-012-0604-2.

Baltagi, B. 2008. *Econometric analysis of panel data* 4th ed. UK: John Wiley & Sons Ltd.

Baltagi, B. and Moscone, F. 2010. Health care expenditure and income in the OECD reconsidered: Evidence from panel data. *Economic Modelling* 27(4), pp, 804-811. doi: 10.1016/j.econmod.2009.12.001

Banker, R. D. and Gifford, J. L. 1988. *A relative efficiency model for the evaluation of public health nurse productivity.* Pittsburgh: School of Urban and Public Affairs, Carnegie Mellon University.

Banker, R. D., Charnes, A. and Cooper, W. W. 1984. Models for estimating technical and scale efficiencies in data envelopment analysis. *Management Science* 30(9), pp. 1078-1092.

Banker, R. D., Charnes, A., Cooper, W. W., Swarts, J. and Thomas, D. A. 1989a. An introduction to data envelopment analysis with some of its models and their uses. *Research in Governmental and Nonprofit Accounting* 5, pp. 125-163.

Barankay, I. and Lockwood, B. 2007. Decentralization and the productive efficiency of government: evidence from Swiss cantons. *Journal of Public Economics* 91(5–6), pp. 1197–1218. doi: 10.1016/j.jpubeco.2006.11.006.

Barberis, P. 1998. The new public management and a new accountability. *Public Administration* 76(3), pp. 451–70.

Bardhan, P. and Mookherjee, D. 2005. Decentralizing antipoverty program delivery in developing countries. *Journal of Public Economics* 89(4), pp. 675–704.

Barma, N.H., Kaiser, K., Le, T.M. and Vinuela, L. 2012. *Rents to riches? The political economy of natural resource–led development.* Washington, DC: The World Bank.

Barrett, M. 2001. A stakeholder approach to responsiveness and accountability in non-profit organisations. *Social Policy Journal of New Zealand* 17, pp. 36-51.

Baskaran, T. 2010. On the link between fiscal decentralization and public debt in OECD countries. *Public Choice* 145(3-4), pp. 351-378.

Batley, R.A. and Larbi, G.A. 2006. Capacity to deliver? Management, institutions and public services in developing countries. In: Bangura, Y. and Larbi, G. eds. *Public sector management reform in developing countries*. Basingstoke: Palgrave Macmillan.

Bauhr, M., Czibik, A., de Fine Licht, J. and Fazekas, M. 2020. Lights on the shadows of public procurement: transparency as an antidote to corruption. *Governance* 33(3), pp. 495-523.

Baum, C.F., Schaffer, M.E., and Stillman, S. 2002. Instrumental variables and GMM: estimation and testing. *Boston College Economics Working Paper, No 545*. Massachusetts, USA: Boston College.

Beaglehole, R. 2004. Public health in the new era: improving health through collective action. *Lancet* 363, pp. 2084-2086

Bedard, K. and Frech, H.E. 2009. Prison health care: is contracting out healthy? *Health Economics* 18(11), pp. 1248–1260. doi: 10.1002/hec.1427.

Beetham, D. 2012. Political legitimacy. In Amenta, E., Nash, K. and Scott, A. eds. *Political Sociology*. Chichester, UK: Blackwell Publishing Ltd.

Bello-Gomez, R.A. 2020. Interacting capacities: the indirect national contribution to subnational service provision. *Public Administration Review 80*, pp. 1011-1023. doi: 10.1111/puar.13192

Benaine, S. L., and Kroll, A. 2020. Explaining effort substitution in performance systems: the role of task demands and mission orientation. *Public Management Review* 22(6), pp. 813-835. doi: 10.1080/14719037.2019.1604794

Bergman, M.A., Johansson, P., Lundberg, S. and Spagnolo, G. 2016. Privatization and quality: Evidence from elderly care in Sweden. *Journal of Health Economics* 49, pp. 109–119. doi: 10.1016/j.jhealeco.2016.06.010.

Bertrand. M., Duflo, E. and Mullainthan, S. 2004. How much should we trust differences-in-differences estimates. *Quarterly Journal of Economics* 119(1), pp. 249-275.

Besley, T. and Case, A. 1995. Incumbent behavior: vote seeking, tax setting and yardstick competition. *American Economic Review* 85 (1), pp. 25-45.

Besley, T. and Coate, S. 2003. Central versus local provision of public goods: a political economy analysis. *Journal of Public Economics* 87(12), pp. 2611–2637. doi: 10.1016/S0047-2727(02)00141-X

Bevan, G. and Hood, C. 2006. What's measured is what matters: targets and gaming in the English public health care system. *Public Administration* 84(3), pp. 517–538. doi: 10.1111/j.1467-9299.2006.00600.x.

Bhaskar, R. 1989. *Reclaiming Reality: a Critical Introduction to Contemporary Philosophy.* London: Verso.

Blaikie, N. 2003. *Designing social research*. Cambridge: Polity Press.

Blöchliger, H. 2006. Fiscal autonomy of sub-central governments. *Network on Fiscal Relations Across Levels of Government Working Paper* 2. Paris: OECD.

Bloom, G. 2011. Building institutions for an effective health system: lessons from China's experience with rural health reform. *Soc Sci Med* 72, pp. 1302–1309. doi: 10.1016/j.socscimed.2011.02.017

Blumenthal D. and Hsiao, W.C. 2005. Privatization and its discontents – the evolving Chinese health care system. *New England Journal of Medicine* 353, pp. 1165-1170. doi: 10.1056/NEJMhpr051133

Boadway, R. and Shah, A. 2009. Introduction to federalism and the role of government in federal economies. In Boadway, R. and Shah, A. eds. *Fiscal federalism: principles and practice of multiorder governance*. New York: Cambridge University Press, pp. 4-60.

Boadway, R. and Tremblay, J. 2012. Reassessment of the Tiebout model. *Journal of Public Economics* 96(11), pp. 1063-1078. doi: 10.1016/j.jpubeco.2011.01.002

Bodman, P. and Hodge, A. 2010. What drives fiscal decentralisation? Further assessing the role of income. *Fiscal Studies* 31(3), pp. 373–404. doi: 10.1111/j.1475-5890.2010.00119.x

Boetti, L., Turati, G. and Piacenza, M. 2012. Decentralization and local governments' performance: how does fiscal autonomy affect spending efficiency? *FinanzArchiv Public Finance Analysis* 68(3), pp. 269-302. doi: 10.1628/001522112x653840.

Bohte, J. and Meier, K. J. 2000. Goal displacement: assessing the motivation for organizational cheating. *Public Administration Review* 60(2), pp. 173–182. doi: 10.1111/0033-3352.00075.

Bordignon, M., Cerniglia, F. and Revelli, R. 2004. Yardstick competition in intergovernmental relationships: theory and empirical predictions. *Economics Letters* 83(3), pp. 325–333. doi: 10.1016/j.econlet.2003.11.014

Borge L.E., Falch, T. and Tovmo, P. 2008. Public sector efficiency: the roles of political and budgetary institutions, fiscal capacity, and democratic participation. *Public Choice* 136(34), pp. 475–495

Borins, S. 1994. *Government in transition: a new paradigm in public administration.* Report on the Inaugural Conference of CAPAM. Toronto: CAPAM.

Bouchard, T. J., Jr. 1976. Field research methods: interviewing, questionnaires, participant observation, systematic observation, unobtrusive measures. In Dunnette, M. D. eds. *Handbook of Industrial and Health National Psychology.* Chicago: Rand McNally, pp. 363– 413.

Bouckaert, G. 1993. Measurement and meaningful management. *Public Productivity and Management Review 17*(1), pp. 31–43.

Bovens, M. 2005. Public accountability. In: Ferlie, E., Lynne, L. and Pollitt, C. eds. *The Oxford Handbook of Public Management*. Oxford: Oxford University Press, pp. 182–208.

Bovens, M. 2007. Analysing and assessing accountability: a conceptual framework. *European Law Journal* 13 (4), pp. 447–468. doi: 10.1111/eulj.2007.13.issue-4

Bowe, R. and Ball S. J. 1992. *Reforming education and changing schools: case studies in policy sociology*. London: Routledge

Bowlin, W.F. 1998. Measuring performance: an introduction to data envelopment analysis (DEA). *The Journal of Cost Analysis* 15, pp. 3–27. doi: 10.1080/08823871.1998.10462318

Boyne, G.A. 1995. Population size and economies of scale in local government. *Policy & Politics* 23, pp. 213–222.

Boyne, G.A. 1996. *Constraints, choices and public policies*. London: JAI Press.

Boyne, G.A. 1998. *Public choice theory and local government: a comparative analysis of the UK and the USA*. London: Macmillan Press Ltd.

Boyne, G.A 2002. Theme: local government: concepts and indicators of local authority performance: an evaluation of the statutory frameworks in England and Wales. *Public Money and Management* 22(2), pp. 17-24. doi: 10.1111/1467-9302.00303.

Boyne, G.A. 2003. Sources of public service improvement: a critical review and research agenda. *Journal of Public Administration Research and Theory* 13(3), pp. 367-394. doi: 10.1093/jpart/mug027

Boyne, G., Farrell, C., Law, J., Powell, M. and Walker, R. 2003. *Evaluating public management reforms: principles and practice*. Buckingham: Open University Press.

Brambor, T., Clark, W.R. and Golder, M. 2006. Understanding interaction models: improving empirical analyses. *Political Analysis* 14(1), pp. 63–82. doi: 10.1093/pan/mpi014.

Brennan, G. and Buchanan, J. 1980. *The power to tax: analytical foundations of a fiscal constitution,* Cambridge: Cambridge University Press.

Breuss, F. and Eller, M. 2004. The optimal decentralisation of government activity: normative recommendations for the European Constitution. *Constitutional Political Economy* 15(1), pp. 27–76.

Brignall, S. And Modell, S. 2000. An institutional perspective on performance measurement and management in the “New Public Sector.” *Management Accounting Research* 11(3): pp. 281-306

Brinkerhoff, D. W. and Wetterberg, A. 2013. Performance-based public management reforms: experience and emerging lessons from service delivery improvement in Indonesia. *International Review of Administrative Sciences* 79(3), pp. 433-457. doi: 10.1177/0020852313491059

Brock, G., Jin, Y. and Zeng, T. 2015. Fiscal decentralization and China’s regional infant mortality. *Journal of Policy Modeling* 37, pp. 175-188. doi: 10.1016/j.jpolmod.2015.03.001

Broms, R., Dahlström, C. and Nistotskaya, M. 2020. Competition and service quality: evidence from Swedish residential care homes. *Governance* 33(3), pp. 525-543.

Brouhle, K. et al. 2005. Local residential sorting and public goods provision: a Classroom Demonstration. *The Journal of Economic Education* 36(4), pp. 332-341. doi: 10.3200/jece.36.4.332-344.

Brueckner, J. 1977. Quasi-public goods in a Tiebout model. *College of Commerce and Business Administration faculty working papers*. Urbana-Champaign, IL: University of Illinois at Urbana-Champaign.

Bruno, G. 2004. Limited dependent panel data models: a comparative analysis of classical and Bayesian inference among econometric packages. *Computing in Economics and Finance* 41, Society for Computational Economics. Available at: http://fmwww.bc.edu/RePEc/sce2004/up.24654.1076487923.pdf [Accessed: 11 November 2022]

Bryman, A. 2012. *Social research methods* 4th ed. Oxford: Oxford University Press.

Bryman, B., and Bell, E. 2011. *Business research methods* 3rd ed. Oxford: Oxford University Press.

Bryson, J.M., Crosby, B.C. and Bloomberg, L. 2014. Public value governance: moving beyond traditional public administration and the New Public Management. *Public Administration Review* 74(4), pp. 445-456. doi: 10.1111/puar.12238.

Bulman, D.J. and Jaros, K.A. 2020. Loyalists, localists, and legibility: the calibrated control of provincial leadership teams in China. *Politics and Society* 48(2), pp. 199–234. doi: 10.1177/0032329220908691.

Burki, T. 2022. Dynamic zero COVID policy in the fight against COVID. *The Lancet.* doi: 10.1016/ S2213-2600(22)00142-4

Burns, J. and Zhou, Z. 2010. Performance management in the government of the People's Republic of China. *OECD Journal on Budgeting* 10(2), pp. 1-28. doi: 10.1787/budget-10-5km7h1rvtlnq.

Burns, L.R. and Huang, Y. Z. 2017. History of China's healthcare system. In Burns, L. R. and G. G. Liu. Eds. *China's Healthcare System and Reform.* Cambridge: Cambridge University Press. doi:10.1017/9781316691113

Burns, L.R., Bradley, E.H. and Weiner, B.J. 2012. The management challenge of delivering value in health care: global and US perspectives. In: Burns, L.R., Bradley, E.H. and Weiner, B.J. eds. *Shortell & Kaluzny's Health Care Management: Organization Design and Behavior* 6th ed. Clifton Park, NY: Delmar Cengage Learning, pp. 2-32.

Burr, V. 2015. *Social constructionism*. London: Routledge.

Butler, J. and Moffit, R. 1982. A computationally efficient quadrature procedure for the one factor multinomial Probit model. *Econometrica* 50, pp. 761-764.

Cai, F. and Wang, D. W. 2008. Impacts of internal migration on economic growth and urban development in China. In: DeWind, J. and Holdaway, J. eds. *Migration and Development within and across borders: research and policy perspectives on internal and international migration*. Geneva and New York: International Organization for Migration (IOM) and Social Science Research Council (SSRC), pp. 245-272.

Caldeira, E. 2010. Yardstick competition in a Federation: theory and evidence from China. *China Economic Review* 23(4), pp. 878-897. doi: 10.1016/j.chieco.2012.04.011

Caldeira, E. 2012. Yardstick competition in a federation: theory and evidence from China. *China Economic Review* 23, pp. 878-897. doi: 10.1016/j.chieco.2012.04.011

Camanho, A. and Dyson, R. 2006. Data envelopment analysis and Malmquist indices for measuring group performance. *Journal of Productivity Analysis* 26(1), pp. 35-49. doi: 10.1007/s11123-006-0004-8.

Campbell, A. 1997. Stakeholders: the case in favour. *Long Range Planning* 30(3), pp. 446–49.

Cantarero, D. and Pascual, M. 2008. Analysing the impact of fiscal decentralization on health outcomes: empirical evidence from Spain. *Applied Economics Letters* 15(2), pp. 109-111. doi: 10.1080/13504850600770913

Caribbean Policy Research Institute. 2011. *Towards public sector reform in Jamaica: what local and international experiences tell us about successful public sector reform.* Available at: https://www.capricaribbean.org/documents/towards-public-sector-reform-jamaica [Accessed: 11 November 2022]

Carrillo B, and Duckett J. 2011. *China's changing welfare mix: local perspectives.* London: Routledge.

Cartlidge, P. and Stewart, J. 1995. Effect of changing the stillbirth definition on evaluation of perinatal mortality rates. *The Lancet* 346(8973), pp. 486-488. doi: 10.1016/s0140-6736(95)91327-0.

Castells, M. 1992. Four Asian tigers with a dragon head: a comparative analysis of the state, economy, and society in the Asian Pacific Rim. In Richard P. A and Henderson, J. (eds). *State and development in the Asia-Pacific Rigion*. 52–75. London: Sage.

Caulfield, J.L. 2006. Local government reform in China: a rational actor perspective. *International Review of Administrative Sciences* 72(2), pp. 263-267.

Cavalieri, M. and Ferrante, L. 2016. Does fiscal decentralization improve health outcomes? Evidence from infant mortality in Italy. *Social Science & Medicine* (164), pp. 74-88. doi: 10.1596/1813-9450-2565

Central Committee of the CPC. 1978. *Communiqué of the Third Plenary Session of the 11th Central Committee of the CPC.* Available at: http://cpc.people.com.cn/GB/64162/64168/64563/65371/4441902.html [Accessed: 19 Feb 2022]

Central Committee of the CPC. 1993. *Decisions on market reform*. Beijing: Central Committee of Chinese Communist Party.

Central Committee of the CPC. 2006. *Decision of the Central Committee of the Communist Party of China on several big issues on promoting the Socialist Harmonious Society.* Available at: http://www.gov.cn/gongbao/content/2006/content_453176.htm [Accessed: 02 Oct 2022]

Central Committee of the CPC and State Council. 2016. *Health China 2030 Blueprint.* Available at: http://www.gov.cn/zhengce/2016-10/25/content_5124174.htm [Accessed: 19 Feb 2022]

Central Leading Group for Comprehensively Deepening Reforms. 2014. *Communique of the third plenary session of the 18th Central Committee of the Communist Party*. Beijing: Central Leading Group for Comprehensively Deepening Reforms.

Chan, H.S. and Chow, K.W. 2007. Public management policy and practice in Western China: metapolicy, tacit knowledge, and implications for management innovation transfer. *The American Review of Public Administration* 37(4), pp. 479-498.

Chan, H.S. and Gao, J. 2008. Performance measurement in Chinese local governments. *Chinese Law and Government* 41(2–3), pp. 4–9. doi: 10.2753/CLG0009-4609410200

Chan, H.S. and Gao, J. 2009. Putting the cart before the horse: accountability or performance? *The Australian Journal of Public Administration* 68(S1), pp. S51–S61. doi:10.1111/j.1467-8500.2009.00621.x

Chan, H.S. and Gao, J. 2012. Death versus GDP! Decoding the fatality indicators on work safety regulation in post-Deng China. *China Quarterly* 210, pp. 355–77. doi: 10.1017/S0305741012000379

Chan, H.S. and Rosenbloom, D.H. 2010. Four challenges to accountability in contemporary public administration: lessons from the United States and China. *Administration and Society* 42(S1), pp. 11-33. doi: 10.1177/0095399710361851

Chan, K.W. and Buckingham, W. 2008. Is China abolishing the Hukou system? *The China Quarterly 195*, pp. 582-606. doi: 10.1017/s0305741008000787.

Chan, K.W. 2012. Migration and development in China: trends, geography and current issues, *Migration and Development* 1(2), pp. 187-205. doi: 10.1080/21632324.2012.739316

Channa, A. and Faguet, J. 2012. Decentralization of health and education in developing countries: a quality-adjusted review of the empirical literature. *Economic organisation and public policy discussion papers, EOPP 38*. London, UK: STICERD

Chao, J. Lu. B., Zhang, H., Zhou, L., Jin, H. and Liu, P. 2017. Healthcare system responsiveness in Jiangsu Province, China. *BMC Health Services Research* 17(1). doi: 10.1186/s12913-017-1980-2.

Charnes, A. and Cooper, W.W. 1962. Programming with linear fractional functionals. *Naval Research Logistics Quarterly* 15, pp. 517–522.

Charnes, A., Cooper, W.W. and Rhodes, E. 1978. Measuring efficiency of decision-making units. *European Journal of Operational Research* 2(6), pp. 429-444.

Chen, F., Ma, X. and Chen, S. 2014. Refined-scale panel data crash rate analysis using random-effects tobit model. *Accident Analysis & Prevention* 73, pp. 323-332. doi: 10.1016/j.aap.2014.09.025.

Chen, S. 2014. Research on the expression mechanism of urban public service demand: an analytical framework. *Fudan Public Administration Review* 12, pp. 110-128.

Chen, S. and Jia, K. 2021. How local governments prioritize multiple conflicting goals: Beyond the sole-goal perspective. *Public Administration*. doi: 10.1111/padm.12807.

Chen, Y. 2005. Measuring super-efficiency in DEA in the presence of infeasibility. *European Journal of Operational Research* 161(2), pp. 545-551. doi: 10.1016/j.ejor.2003.08.060.

Chen, Y., Li, H. and Zhou, L. 2005. Relative performance evaluation and the turnover of provincial leaders in China. *Economics Letters* 88(3), pp. 421-425. doi: 10.1016/j.econlet.2005.05.003.

Chen, Y. and Sherman, H. 2004. The benefits of non-radial vs. radial super-efficiency DEA: an application to burden-sharing amongst NATO member nations. *Socioeconomic Planning Sciences* 38(4), pp. 307-320. doi: 10.1016/j.seps.2003.10.007.

Cheol, G.G. 2012. *Developing composite indicators for fiscal decentralization: which is the best measure for whom*? MPRA Paper No. 43032.

Chien, S. 2008. Local responses to globalization in China a territorial restructuring process perspective. *Pacific Economic Review* 13(4), pp. 492-517. doi: 10.1111/j.1468-0106.2008.00414.x

Chien, S. 2009. Prefectures and prefecture-level cities: the political economy of administrative restructuring. In Chung, J.H. and Lam, T.C. eds. *China's Local Administration.* London: Routledge, Chapter 7.

Chien, S. 2010. Economic freedom and political control in post-Mao China: a perspective of upward accountability and asymmetric decentralization. *Asian Journal of Political Science* 18(1), pp. 69-89. doi: 10.1080/02185371003669379.

Cho, S. And Duggan, J. 2009. Bargaining foundations of the median voter theorem. *Journal of Economic Theory* 144(2), pp. 851-868. doi: 10.1016/j.jet.2008.06.011.

Cho, Y.N. 2009. *Local People's Congresses in China: development and transition*. Cambridge, New York: Cambridge University Press.

Choi, E.K. 2009. Informal tax competition among local governments in China since the 1994 tax reforms. *Issues and Studies* 45, pp. 159-183.

Christensen, T., Dong, L. and Painter, M. 2008. Administrative reform in China's central government: how much learning from the West? *International Review of Administrative Sciences* 74(3), pp. 351–371. doi: 10.1177/0020852308095308

Christensen, T. and Laegreid, P. 2012. Performance and accountability – a theoretical discussion and an empirical assessment. *Public Organization Review* 15(2), pp. 207-225. doi: 10.1007/s11115-013-0267-2

Chu, Y.H. 2013. Sources of regime legitimacy and the debate over the Chinese model. *China Review* 13(1), pp. 1–42.

Chun, Y.H. and Rainey, H.G. 2005. Goal ambiguity and organizational performance in U.S. Federal Agencies. *Journal of Public Administration Research and Theory* 15(4), pp. 529–558. doi: 10.1093/jopart/mui030

Chung, J.H. 2000. *Central control and local discretion in China.* Oxford: Oxford University Press.

Clark, P. 2008. *The Chinese Cultural Revolution, a history*. Cambridge: Cambridge University Press.

Coelli, T.J., Rao, D.S.P., O'Donnell, C.J. and Battese, G.E. 2005. *An introduction to efficiency and productivity analysis.* Boston, MA: Springer.

Cohen, L. and Manion, L. 1994. *Research methods in education.* 4th ed. London: Routledge.

Collins, C. and Green, A. 1994. Decentralization and primary healthcare: some negative implications in developing countries. *International journal of health services: planning, administration, evaluation* 24(3), pp. 459.

Common, R.K. 1998. Convergence and transfer: a review of the globalisation of new public management. *International Journal of Public Sector Management* 11(6), pp. 440-450.

Confucius. 1999. *The analects of Confucius: a philosophical translation.* New York: Ballantine Books.

Cook, W. D., Liang, L., Zha, Y. and Zhu, J. 2009. A modified super-efficiency DEA model for infeasibility. *Journal of the Operational Research Society* 60(2), pp. 276-281. doi: 10.1057/palgrave.jors.2602544

Cook, W. D., Tong, K., and Zhu, J. 2014. Data envelopment analysis: prior to choosing a model. *Omega* 44, pp. 1-4. doi: 10.1016/j.omega.2013.09.004

Cooper, W. W., Seiford, L. M. and Tone, K. 2007. *Data envelopment analysis: a* *comprehensive text with models, applications, references and DEA-Solver Software* 2nd ed. New York: Springer.

Courty, P. and Marschke, G. 2004. An empirical investigation of gaming responses to explicit performance incentives. *Journal of Labor Economics* 22(1), pp. 23–56. doi: 10.1086/380402.

Creswell, J.W. 2009. *Research design: qualitative, quantitative, and mixed methods approaches*. 3rd ed. Thousand Oaks, CA: SAGE.

Creswell, J.W., and Plano Clark, V.L. 2007. *Designing and conducting mixed methods research*. Thousand Oaks, CA: SAGE.

Creswell, J.W., and Plano Clark, V.L. 2011. *Designing and conducting mixed methods research.* 2nd ed. Thousand Oaks: Sage Publications.

Creswell, J.W., Plano Clark, V., Gutmann, M., and Hanson, W. 2003. Advanced mixed methods designs. In Tashakkori, A. and Teddlie, C. eds. *Handbook of mixed method research in the social and behavioral sciences.* Thousand Oaks, CA: SAGE, pp. 209-240.

Crowley, G.R. and Sobel, R.S. 2011. Does fiscal decentralization constrain leviathan? New evidence from local property tax competition. *Public Choice* 149 (1), pp. 5–30. doi:10.1007/s11127-011-9826-7

Cruz, N.F. and Marques, R.C. 2014. Revisiting the determinants of local government performance. *Omega* 44, pp. 91–103. doi: 10.1016/j.omega.2013.09.002.

Curran-Everett, D. 2018. Explorations in statistics: the log transformation. *Adv Physiol Educ* 42, pp. 343-347. doi:10.1152/advan.00018.2018.

Çolak, Ç.D. 2019. Why the New Public Management is obsolete: an analysis in the context of the post-New Public Management trends. *Croatian and Comparative Public* 19(4), pp. 517–536. doi: 10. 31297/hkju.19.4.1

Dabla-Norris, E. 2005. Issues in Intergovernmental Fiscal Relations in China. *IMF working paper 05/30*. Washington: IMF.

Dada, E.A. 2015. Fiscal decentralization and social services in Nigeria. *AERC Research Paper 291*. Nairobi, Kenya: African Economic Research Consortium.

Dagher, R. 2018. Legitimacy and post-conflict state-building: the undervalued role of performance legitimacy. *Conflict, Security & Development* 18(2), pp. 85-111. doi: 10.1080/14678802.2018.1447860.

Davis, M.L. and Hayes, K. 1993. The demand for good government. *Review of Economics and Statistics* 75, pp. 148–152.

Day, P. and Klein, K. 1987. *Accountability. Five public services*. London: Tavistock Publishers.

De Borger, B. and Kerstens, K. 1996. Cost efficiency of Belgian local governments: a comparative analysis of FDH, DEA, and econometric approaches. *Regional Science and Urban Economics* 26(2), pp. 145–170.

De Borger., Kerstens, K., Moesen, W. and Vanneste, J. 1994. Explaining differences in productive efficiency: an application to Belgian municipalities. *Public Choice* 80(3-4), pp. 339-358. doi: 10.1007/bf01053225.

De Bruijn, H. 2006. *Managing Performance in the Public Sector*. New York: Routledge.

De Mello, L. 2011. Does fiscal decentralisation strengthen social capital? Cross-country evidence and the experiences of Brazil and Indonesia. *Environment and Planning C: Government and Policy* 29, pp. 281–296. doi: 10.1068/c10102r

De Walle, S. V. and Bouckaert, G. 2003. Public service performance and trust in government: the problem of causality. *International Journal of Public Administration* 26(8-9), pp. 891-913. doi: 10.1081/PAD-120019352

Deng, F. 2018. Stakes, stakeholders and urban governance: a theoretical framework for the Chinese city. *Eurasian Geography and Economics* 59(3-4), pp. 291-313. doi: 10.1080/15387216.2019.1570298

Desai, V. and Potter, R.B. 2006. Interviewing. In: Desai, V. and Potter, R. B. eds. *Doing development research.* London: Sage Publications, pp. 142–152.

Development Research Center. 2005. *An assessment of health care reforms in China*. Beijing: Development Research Centre of the State Council.

Dickson, B., Landry, P., Shen, M. and Yan, J. 2016. Public goods and regime support in urban China. *The China Quarterly* 228, pp. 859-880. doi: 10.1017/s0305741016001156.

Ding, J., Hu, X., Zheng, X., Shang, L., Yu, M. and Chen, H. 2018. Equity and efficiency of medical service systems at the provincial level of China's mainland: a comparative study from 2009 to 2014. *BMC Public Health* 18, pp. 1–14. doi: 10.1186/s12889-018-5084-7.

Dollar, D. and Hofman, B. 2006. Intergovernmental fiscal reforms, expenditure assignment, and governance. *Paper Presented at the Roundtable Conference on Public Finance for a Harmonious Society*. Beijing, 28-27 June, 2006.

Dollery, B. and Fleming, E. 2006. A conceptual note on scale economies, size economies and scope economies in Australian local government. *Urban Policy and Research* 24(2), pp. 271-282. doi: 10.1080/08111140600704111.

Donaldson, J.A. 2017a. Introduction: understanding central-local relations in China. In Donaldson, J. A. eds. *Assessing the Balance of Power in Central-Local Relations in China.* New York: Routledge, pp. 1-18.

Donaldson, J.A. 2017b. China's administrative hierarchy: the balance of power and winners and losers within China's levels of government. In Donaldson, J. A. eds. *Assessing the Balance of Power in Central-Local Relations in China*. New York: Routledge, pp. 105-137.

Dong L., Cui, Q. and Christensen, T. 2015. Local public services provision in China – an institutional analysis. *Croatian and Comparative Public Administration* 15(3), pp. 617-644.

Dou, G., Wang, Q. and Ying, X. 2018. Reducing the medical economic burden of health insurance in China: achievements and challenges. *BioScience Trends* 12(3), pp. 215-219. doi: 10.5582/bst.2018.01054.

Dowding, K. and Mergoupis, T. 2003. Fragmentation, fiscal mobility, and efficiency. *The Journal of Politics* 65(4), pp. 1190-1207. doi: 10.1111/1468-2508.t01-1-00132

Downs, A. 1957. An economic theory of political action in a democracy. *The Journal of Political Economy* 65(2), pp.135-150.

Du, J. and Yi, H. 2022. Target-setting, political incentives, and the tricky trade-off between economic development and environmental protection. *Public Administration* 100(4), pp. 923-941. doi: 10.1111/padm.12768

Duan, H. and Zhan, J.V. 2011. Fiscal transfer and local public expenditure in China: a case study of Shanxi Province. *China Review* 11(1), pp. 57-88.

Duckett, J. and Munro, N. 2022. Authoritarian regime legitimacy and health care provision: survey evidence from contemporary China. *Journal of Health Politics, Policy and Law* 47(3), pp. 375-409. doi: 10.1215/03616878-9626894

Duckett, J. and Wang, G. 2017. Why do authoritarian regimes provide public goods? Policy communities, external shocks and ideas in China's rural social policy making. *Europe-Asia Studies* 69(1), pp. 92–109. doi: 10.1080/09668136.2016.1274379

Dziobek, C., Mangas, C. G. and Kufa, P. 2011. Measuring fiscal decentralization – exploring the IMF's databases. *IMF working paper 11/126*. Washington: IMF.

Ebel, R.D. Yilmaz, S. 2002. On the measurement and impact of fiscal decentralization. *Policy Research Working Paper 2809.* Washington: the World Bank.

Eberhardt, M. S., Ingram, D. D., Makuk, D. M., Pamuk, E. R., Freid, V. M., Harper, S. B. and Xia, H. 2001. *Health, United States, 2001: urban and rural health chartbook.* Hyattsville, MD: National Center for Health Statistics.

Edin, M. 2003. Remaking the communist party-state: the cadre responsibility system at the local level in China. *China* 1(1), pp. 1-15.

Eggleston, K., Li, L., Meng, Q., Lindelow, M. and Wagstaff, A. 2008. Health service delivery in China: a literature review. *Health Econ* 17, pp. 149–65

Enikolopov, R.and Zhuravskaya, E. 2003. Decentralisation and political institutions. *Discussion paper 3857*. London: Centre for Economic Policy Research.

Evans, D.B., Tandon, A., Murray, C.J.L. and Lauer, J.A. 2001. Comparative efficiency of national health systems: cross national econometric analysis. *BMJ* 323(7308), pp. 307-310. doi: 10.1136/bmj.323.7308.307.

Farrell, C. and Jones, J. 2000. Evaluating stakeholder participation in public services – parents and schools. *Policy and Politics* 28(2), pp. 251-262. doi: 10.1332/0305573002500956.

Feng, C. 2014. Log-transformation and its implications for data analysis. *Shanghai Archives of Psychiatry* 26(2), pp. 105-109. doi: 10.3969/j.issn.1002-0829.2014.02.009

Feng, X., Ljungwall, C., Guo, S. and Wu, A.M. 2013. Fiscal Federalism: a refined theory and its application in the Chinese context. *Journal of Contemporary China* 22(82), pp, 573-593. doi: 10.1080/10670564.2013.766381

Ferlie, E., Ashburner, L., Fitzgerald, L. and Pettigrew, A. 1996. *The new public management in action*. Oxford: Oxford University Press.

Fernandez-Val, I. and Weidner, M. 2016. Individual and time effects in nonlinear panel models with large N, T. *Journal of Econometrics* 192(1), pp. 291-312. doi: 10.1016/j.jeconom.2015.12.014

Fisman, R. and Gatti, R. 2000. *Decentralization and corruption: evidence across countries.* Policy Research Working Paper No. 2290. Washington DC. The World Bank.

Fonchamnyo, D.C. and Sama, M. C. 2016. Determinants of public spending efficiency in education and health: evidence from selected CEMAC countries. *Journal of Economics and Finance* 40, pp. 199-201. doi: 10.1007/s12197-014-9310-6

Fredriksen, K. 2013. Decentralisation and economic growth – part 3: decentralisation, infrastructure investment and educational performance. *OECD Working Papers on Fiscal Federalism* No. 16. Paris: OECD.

Freeman, R.E. 1984. *Strategic management: a stakeholder approach*. Cambridge, UK: Cambridge University Press.

Freinkman, L. and Plekhanov, A. 2009. FD and the quality of public services in Russian regions. *EBRD Working Paper No. 111*. London: European Bank for Reconstruction and Development.

Frenkiel, E. 2021. Participatory budgeting and political representation in China. *Journal of Chinese Governance* 6, pp. 58–80.

Friedman, J. 2011. *Sticking to the numbers: performance monitoring in South Africa, 2009–2011*. Princeton, NJ: Princeton University.

Fukasaku, K. 1999. FD and macroeconomic governance: Asian perspectives. In: Fukasaku, K. and De Mello, L. R. eds. *FD in emerging economies.* Paris: OECD, pp. 101-120.

Funck, E.K. and Karlsson, T.S. 2020. Twenty-five years of studying new public management in public administration: Accomplishments and limitations. F*inancial Accountability and Management* 36(4), pp. 347-375. doi: 10.1111/faam.12214

Gao, H., Tyson, A. and Cheng, G. 2021. Novel virus, novel response: Local discretion and responses to COVID-19 in Hebei Province, China. *Asia and The Pacific Policy Studies* 9(1), pp. 5-22. doi: 10.1002/app5.342

Gao, J. 2009. Governing by goals and numbers: a case study in the use of performance measurement to build state capacity in China. *Public Administration and Development* 29(1), pp. 21-31. doi: 10.1002/pad.514.

Gao, J. 2010. Hitting the target but missing the point: the rise of non-mission-based targets in performance measurement of Chinese local governments. *Administration and Society* 42(1S), pp. 56S–76S doi: 10.1177/0095399710361854.

Gao, J. 2015. Pernicious manipulation of performance measures in China's cadre evaluation system. *China Quarterly* 223, pp. 618–637. doi: 10.1017/S0305741015000806.

Gao, J. 2021. Mitigating pernicious gaming in performance management in China: dilemmas, strategies and challenges. *Public Performance and Management Review* 44(2), pp. 321–351. doi: 10.1080/15309576.2020.1775662.

Gao, J. and Zhang, P. 2021. China's public health policies in response to COVID-19: from an "authoritarian" perspective. *Frontiers in Public Health* 9:756677. doi: 10.3389/fpubh.2021.756677

Gao, S., Meng, X. and Zhang, L. 2014. Fiscal decentralization and life satisfaction: evidence from urban China. *Social Indicators Research* 119(3), pp. 1177–1194. doi: 10.1007/s11205-013-0552-z.

Gao, Y., Barclay, L., Kildea, Sue., Hao, M., and Belton, S. 2010. Barriers to increasing hospital birth rates in rural Shanxi Province, China. *Reproductive Health Matters*

18(36), pp. 35-45. doi: 10.1016/ S0968-8080(10)36523-2

Geis, J. P. and Holt, B. 2009. “Harmonious Society” rise of the new China. *Strategic Studies Quarterly* 3(4), pp. 75-94.

Geys, B. and Moesen, W. 2008. Exploring sources of local government technical inefficiency: evidence from Flemish municipalities. *Discussion Paper SP II 2008 – 18*. Berlin: Wissenschaftszentrum.

Geys, B. and Moesen, W. 2009. Measuring local government technical (in)efficiency. *Public Performance & Management Review* 32(4), pp. 499-513. doi: 10.2753/PMR1530-9576320401

Ghauri, P. and Gronhaug, K. 2002. *Research methods in business studies: a practical guide.* 2nd ed. Harlow; New York: Prentice Hall.

Ghobadian, A., Gallear, D., Viney, H., and O’Regan, N. 2007. Public sector performance improvement through private sector management practices: a satisfactory solution? *International Journal of Business Performance Management 9*(4). pp. 363-379. doi: 10.1504/IJBPM.2007.013360

Ghose, A.K. 2008. The growth miracle, institutional reforms and employment in China. *Economic and Political Weekly* 43(22), pp. 47–56.

Ghuman, B.S. and Singh, R. 2013. Decentralization and delivery of public services in Asia. *Policy and Society* 32(1), pp. 7–21. doi: 10.1016/j.polsoc.2013.02.001

Gibbons, R. 1998. Incentives in organizations. *Journal of Economic Perspectives* 12(4), pp. 115–32.

Giddings, L.S. and Grant, B.M. 2006. Mixed methods research for the novice researcher. *Contemporary Nurse* 23(1), pp. 3-11. doi: 10.5172/conu.2006.23.1.3

Gill, J. and Johnson, P. 2002. *Research methods for managers.* 3rd ed. London: SAGE.

Gill, N. and Rodríguez-Pose, A. 2012. Do citizens really shop between decentralised jurisdictions? Tiebout and internal migration revisited. *Space and Polity* 16(2), pp.175-195. doi: 10.1080/13562576.2012.721505

Gilley, B. 2006. The determinants of state legitimacy: results for 72 countries. *International Political Science Review* 27(1), pp. 47–71

Golany, B. and Roll, Y. 1989. An application procedure for DEA. *Omega* 17(3), pp. 237–250.

Golder, M. 2006. Presidential coattails and legislative fragmentation. *American Journal of Political Science* 34(1), pp. 34-48

Gong, G., Chen, Y., Gao, H., Su, D. and Chang, J. 2019. Has the Efficiency of China's healthcare system improved after healthcare reform? A network Data Envelopment Analysis and Tobit regression approach. *International Journal of Environmental Research and Public Health* 16(23), p. 4847. doi: 10.3390/ijerph16234847.

Gong, T. 2006. Corruption and local governance: the double identity of Chinese local governments in market reform. *The Pacific Review* 19(1), pp. 85-102.

Goodman, G.B. 2015. Local government fragmentation & the local public sector: a panel data analysis. *Public Finance Review* 43(1), pp. 82-107. doi: 10.1177/1091142113515048

Goodman, G.B. 2019. Local government fragmentation: what do we know? *State and Local Government Review* 51(2), pp. 134-144. doi: 10.1177/0160323X19856933

Gore, L. 2019. The Communist Party-dominated governance model of China: legitimacy, accountability, and meritocracy. *Polity* 51(1), pp. 161-194. doi: 10.1086/700615.

Grabowski, D.C. and Stevenson, D.G. 2008. Ownership conversions and nursing home performance. *Health Services Research* 43(4), pp. 1184–1203. doi: 10.1111/j.1475-6773.2008.00841.x.

Greene, W.H. 2012. *Econometric analysis* 7th ed. London: Pearson.

Greene, J.C., Caracelli, V.J. and Graham, W.F. 1989. Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis* 11(3), pp. 255-274.

Greene, W.H. 2004. Fixed effects and bias due to the incidental parameters problem in the Tobit model. *Econometric Reviews* 23(2), pp. 125-147.

Grix, J. 2002. Introducing students to the generic terminology of social research. *Politics*, 22(3), pp. 175-185.

Grossman, P. 1989. Fiscal decentralization and government size: an extension. *Public Choice* 62, pp. 63–69.

Grossman, P., Mavros, P. and Wassmer, R. 1999. Public sector technical inefficiency in large U.S. cities. *Journal of Urban Economics* 46 (2), pp. 278–299.

Gruening, G. 2001. Origin and theoretical basis of new public management. *International Public Management Journal* 4(1), pp. 1-25. doi: 10.1016/s1096-7494(01)00041-1.

Gu, E. and Zhang, J. 2006. Health care regime change in urban China: unmanaged marketization and reluctant privatization. *Pacific Affairs* 79(1), pp. 49–71. doi: 10.1016/j.mseb.2006.03.026.

Gu, G.C. 2012. Developing composite indicators for fiscal decentralization: which is the best measure For whom? *MPRA Working Paper No. 43032*. Munich: Munich Personal RePEc Archive.

Guba, E.G. and Lincoln, Y.S. 1994. Competing paradigms in qualitative research. *Handbook of Qualitative Research* 2(163-194), pp. 105-117.

Guest, G., MacQueen, K.M. and Namey, E.E. 2012. *Applied thematic analysis*. Thousand Oaks: Sage Publications.

Guo, H., Zhao, Y., Niu, T. and Tsui, K.L. 2017. Hong Kong Hospital Authority resource efficiency evaluation: via a novel DEA-Malmquist model and Tobit regression model. *PLoS One*12(9) e0184211.

Habibi, N. et al. 2001. Decentralization in Argentina. *Center discussion paper No. 825.* New Haven, CT: Yale University.

Habibi, N. et al. 2003. Decentralization and human development in Argentina. *Journal of Human Development and Capabilities* 4(1), pp. 73–101.

Hainmueller, J., Mummolo, J. and Xu, Y. 2019. How much should we trust estimates from multiplicative interaction models? simple tools to improve empirical practice. *Political Analysis* 27(2), pp. 163–192. doi: 10.1017/pan.2018.46.

Hammerschmid, G., van de Walle, S. V. and Andrews, R. 2018. New public management reforms in Europe and their effects: findings from a 20-country top executive survey. *International Review of Administrative Sciences* 85(3), pp. 399-418. doi: 10.1177/0020852317751632.

Hankla, C.R. 2009. When is fiscal decentralization good for governance? *Publius* 39(4), pp. 632-650

Hänni, M. 2017. Responsiveness – To whom? Why the primacy of the median voter alienates minorities. *Political Studies* 65(3), pp. 665-684. doi: 10.1177/0032321716680376.

Hao, Y., Chen, Y., Liao, H. and Wei, Y. 2020. China's fiscal decentralization and environmental quality: theory and an empirical study – Erratum. *Environment and Development Economics* 25(2), pp. 204-204. doi: 10.1017/s1355770x1900038x.

Hao, Y., Liu, J, Lu, Z. N., Shi, R. and Wu, H. 2021. Impact of income inequality and fiscal decentralization on public health: evidence from China. *Economic Modelling* 94, pp. 934-944. doi: 10.1016/j.econmod.2020.02.034

Harrison, J.A., Rouse, P. and de Villiers, C.J. 2012. Accountability and performance measurement: a stakeholder perspective. *Journal of Centrum Cathedra: The Business and Economics Research Journal* 5(2), pp. 243-258. doi: 10.7835/jcc-berj-2012-0077.

Harrison, M. 1996. Soviet agriculture and industrialisation. In Davis, J.A. and Mathias, P. eds. *Agriculture and Economic Growth from the Eighteenth Century to the Present*. Oxford: Blackwell, pp. 192-208.

Hartley, D. 2004. Rural health disparities, population health, and rural culture. *American Journal of Public Health* 94, pp. 1675–1678.

Hasenfeld, Y. 1983. *Human service organizations.* Englewood Cliffs, NJ: Prentice Hall.

Hathi, P., Haque, S., Pant, L., Coffey, D. and Spears, D. 2017. Place and child health: the interaction of population density and sanitation in developing countries. *Demography* 54, pp. 337-360. doi: 10.1007/s13524-016-0538-y

Hauner, D. and Kyobe, A. 2010. Determinants of government efficiency. *World Development* 38(11), pp. 1527-1542. doi: 10.1016/j.worlddev.2010.04.004

He, B. 2011. Civic engagement through participatory budgeting in China: three different logics at work. *Public Administration and Development* 31(2), pp. 122–133. doi:10.1002/pad.598

He, B. 2019. Deliberative participatory budgeting: a case study of Zeguo Town in China. *Public Administration and Development* 39(3), pp. 144-153. doi: 10.1002/pad.1853.

Hsieh, C.T. and Urquiola, M. 2006. The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of Public Economics* 90(8–9), pp. 1477–1503. doi: 10.1016/j.jpubeco.2005.11.002.

Heilmann, S. 2008. From local experiments to national policy: the origins of China's distinctive policy process. *The China Journal* 59, pp. 1-30.

Heinrich, C.J. and Lynn, L.E. Jr. 2000. *Governance and performance: new Perspectives.* Washington, DC: Georgetown University Press.

Hendrick, R.M., Jimenez, B.S. and Lal, K. 2011. Does local government fragmentation reduce local spending? *Urban Affairs Review* 47, pp. 467–510. doi: 10.1177/1078087411400379

Heng, Y. and Hong, Z. 2012. A study of productive expenditure bias in county-level finance in China. *Social Sciences in China* 33(1), pp. 127-147. doi: 10.1080/02529203.2012.650414.

Hershatter, G. 2019. *Women and China's revolutions*. Lanham: Rowman et Littlefield

Hillestad R., Bigelow, J., Bower, A., Girosi, F., Meili, R., and Scoville, R. 2005. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs* 24(5), 1103.

Hillier, S. and Shen, J. 1996. Health care systems in transition: People's Republic of China: Part I: an overview of China's health care system. *Journal of Public Health* 18(3), pp. 258-265. doi: 10.1093/oxfordjournals.pubmed.a024502.

Hines, J. and Thaler, R. 1995 Anomalies: the flypaper effect. *Journal of Economic Perspectives* 9, pp. 217–26. doi:10.1257/jep.9.4.217

Hipgrave, D. et al. 2012. Chinese-style decentralization and health system reform. *PLoS Medicine* 9(11), pp. 9–12. doi: 10.1371/journal.pmed.1001337.

Ho, C. 2010. Health reform and de facto federalism in China. *China: An International Journal* 8(1), pp. 33-62. doi: 10.1142/s021974721000004x.

Holbig, H. 2009. Ideological reform and political legitimacy in China. In Heberer, T. and Schubert, G. eds. *Regime Legitimacy in Contemporary China*. London: Routledge, pp. 13-34.

Holbig, H. and Gilley, B. 2010. Reclaiming legitimacy in China. *Politics and Policy* 38(3), pp. 395–422. doi: 10.1111/j.1747-1346.2010.00241.x.

Honadle, B.W. 2001. Theoretical and practical issues of local government capacity in an era of devolution. *The Journal of Regional Analysis and Policy* 31(1), pp. 77-90. doi: 10.22004/ag.econ.132195

Hood, C. 1991. A public management for all seasons. *Public Administration* 69, pp. 3-19.

Hope, K.R. 1998. Development policy and economic performance in Botswana: lessons for the transition economies in Sub-Saharan Africa. *Journal* of *International Development* 10(4), pp. 539-554.

Hope, K.R. and Chikulo, B.C. 2000. Decentralization, the New Public Management, and the changing role of the public sector in Africa. *Public Management an International Journal of Research and Theory* 2(1), pp. 25-42. doi: 10.1080/14719030000000002

Hsiao, W.C. 2007. The political economy of Chinese health reform. *Health Economics, Policy and Law* 2, pp. 241–249.

Hsiao, W.C. 1995. The Chinese health care system: lessons for other nations. *Social Science and Medicine* 41(8): pp. 1047-1055.

Hsu, Y. 2013. The efficiency of government spending on health: evidence from Europe and Central Asia. *The Social Science Journal 50*, pp. 665-673. doi: 10.1016/j.soscij.2013.09.005

Hu, J.T. 2007. *Report to the 17th National Congress of the CPC.* Beijing: Xinhua News Agency. Available at: https://www.chinadaily.com.cn/hqzg/2007-

10/31/content_6220592.htm. [Accessed: 02 Dec 2022].

Hu, R., Shen, C. and Zou, H. F. 2011. Health care system reform in China: issues, challenges and options. *CEMA Working Papers 517*. Beijing: China Economics and Management Academy, Central University of Finance and Economics.

Huang, J., Chen, L., Li, J. and Groot, W. 2017. Expenditure decentralization and citizen satisfaction with healthcare: evidence from urban China. *Social Indicators Research* 133(1), pp. 333–344. doi: 10.1007/s11205-016-1361-y.

Huang, Y. 1995. Administrative monitoring in China. *China Quarterly* 143, pp. 828−842.

Huang, Y. 2009. An institutional analysis of China's failed healthcare reform. In Wu, G. and Lansdowne, H. eds. *Socialist China, Capitalist China: Social Tension and Political Adaptation under Economic Globalization.* New York: Routledge.

Hughes, J.A. 1990. *The philosophy of social research*. 2nd ed. Harlow: Longman.

Hughes, N.P. and Edwards, M. 2000. Leviathan vs. Lilliputian: a Data Envelopment Analysis of government efficiency. *Journal of Regional Science* 40(4), pp. 649-669. doi: 10.1111/0022-4146.00192.

Husain, J. 2017. Policy experimentation and innovation as a response to complexity in China's management of health reforms. *Globalization and Health* 13(54), pp. 1-13.

Im, T. and Choi, Y. 2018. Rethinking national competitiveness: a critical assessment of governmental capacity measures. *Social Indicators Research* 135, pp. 515-532. doi: 10.1007/s11205-016-1494-z

Ingraham, P. W. and Donahue, A. K. 2000. Dissecting the black box revisited: characterizing government management capacity. In Heinrich, C. J. and Lynn, L. E. eds. *Governance and Performance: New Perspectives.* Washington, DC: Georgetown University Press.

International Monatary Funds. 2022. World Economic Outlook Database. Available at: https://www.imf.org/en/Publications/WEO/weo-database/2022/April [Accessed: 15 September 2022]

Iregui, A. 2005. Decentralised provision of quasi-private goods: the case of Colombia. *Economic Modelling* 22(4), pp. 683-706. doi: 10.1016/j.econmod.2005.04.002.

Irfan, S. and Nutley, S. 2016. Debate: using ideas from the west to improve education in Pakistan? *Public Money and Management* 36 (6), pp. 396–397. doi:10.1080/09540962.2016.1206728.

Jackson, P.M. and Palmer, A.J. 1988. The economics of internal organisations: the efficiency of parastatals in LDSc. In Cook, P. and Kirkpatrick, C. eds. *Privatisation in Less Developed Countries.* Brighton: Wheatsheaf Books.

Jiménez-Rubio, J.D. 2011a. The impact of fiscal decentralization on infant mortality rates: Evidence from OECD countries. *Social Science and Medicine* 73(9), pp. 1401–1407. doi: 10.1016/j.socscimed.2011.07.029

Jiménez-Rubio, J.D. 2011b. The impact of decentralization of health services on health outcomes: evidence from Canada. *Applied Economics* 43(26), pp. 3907-3917. doi: 10.1080/00036841003742579

Jiménez-Rubio, J.D. and Smith, P.C. 2005. *Decentralisation of health care and its impact on health outcomes*. Discussion Papers 05/10. New York: Department of Economics, University of York.

Jin, H., Qian, Y. and Weingast, B.R. 1999. Regional decentralization and fiscal incentives: federalism Chinese style. *Journal of Public Economics* 89(9-10), pp. 1719-1742.

Jin, J. and Zou, H. 2002. How does fiscal decentralization affect aggregate, national, and subnational government size? *Journal of Urban Economics* 52(2), pp. 270-293. doi: 10.1016/s0094-1190(02)00004-9.

Jin, Y. and Sun, R. 2011. Does fiscal decentralization improve healthcare outcomes? Empirical evidence from China. *Public Finance and Management* 11(3), pp. 234-261.

Jing, Y., Cui, Y. and Li, D. 2015. The politics of performance measurement in China. *Policy and Society* 34(1), pp. 49-61. doi: 10.1016/j.polsoc.2015.02.001

Joanis, M. 2014. Shared accountability and partial decentralization in local public good provision. *Journal of Development Economics* 107, pp. 28-37. doi: 10.1016/j.jdeveco.2013.11.002.

Johansson, O. 1997. Optimal Pigovian taxes under altruism. *Land Economics* 73(3), p. 297. doi: 10.2307/3147169.

Johnson, R.B., Onwuegbuzie, A.J. and Turner, L. A. 2007. Toward a definition of mixed methods research. *Journal of Mixed Methods Research 1*(2), pp. 112-133. doi: 10.1177/1558689806298224

Joshi, A. 2017. Legal empowerment and social accountability: complementary strategies toward rights-based development in health? *World Development* 99, pp.160-172. doi: 10.1016/j.worlddev.2017.07.008

Kalb, A., Geys, B. and Heinemann, F. 2011. Value for money? German local government efficiency in a comparative perspective, *Applied Economics* 44(2), pp. 201-218. doi: 10.1080/00036846.2010.502110

Kang, Y., Cho, W. and Jung, K. 2012. Does decentralization matter in health outcomes? Evidence from 22 OECD unbalanced panel data for 1995–2005. *International Review of Public Administration* 17(1), pp. 1-32. doi: 10.1080/12264431.2012.10805215.

Kaplan, H., Hooper, P.L., Stiglitz, J. and Gurven, M. 2015. The causal relationship between fertility and infant mortality. In. Kreager, P., Winney, B., Ulijaszek, S. and Capelli, C. eds. *Population in the Human Sciences: Concepts, Models, Evidence.* Oxford: Oxford University Press, pp. 361-379.

Kelman, S. and Friedman, J.N. 2009. Performance improvement and performance dysfunction: An empirical examination of distortionary impacts of the emergency room wait-time target in the English National Health Service. *Journal of Public Administration Research and Theory* 19(4), pp. 917–946. doi: 10.1093/jopart/mun028.

Kennedy, J. J. 2009. Maintaining popular support for the Chinese Communist Party: the influence of education and the state-controlled media. *Political Studies* 57(3), pp. 517–536. doi.org/10.1111/j.1467-9248.2008.00740.x

Khaleghian, P. 2003. *Decentralization and public services: the case of immunization.* World Bank Policy Research Working Paper No. 2989, Washington DC.

Kim, S. 2003. Research paradigms in organizational learning and performance: competing modes of inquiry. *Information Technology, Learning, and Performance Journal* 21(1), pp. 9-18.

Kim, N. and Cho, W. 2015. Agencification and performance: the Impact of autonomy and result-control on the performance of executive agencies in Korea. *Public Performance and Management Review* 38(2), pp. 214–233. doi: 10.1080/15309576.2015.983826.

Kluvers, R. 2003. Accountability for performance in local government. *Australian journal of public administration* 62(1), pp. 57-69.

Kothari, C.R. 2004. *Research methodology* 7th ed. New Delhi: New Age International (P) Ltd., Publishers

Kowal, P., Naidoo, N., Williams, S.R. and Chatterji, S. 2011. Performance of the health system in China and Asia as measured by responsiveness. *Health* 3(10), pp. 638-646. doi: 10.4236/health.2011.310108.

Krauss, S. E. 2005. Research paradigms and meaning making: a primer. *The Qualitative Report* 10(4), pp. 758-770.

Kudo, H. 2016. Still convinced by New Public Management or introducing somehow New Public Governance? How Japanese local governments are coping with these changes to deliver public services. *Journal for Public and Nonprofit Services* 39(1/2), pp. 77-88.

Lam, T. 2010. Central-provincial relations amid greater centralization in China. *China Information* 24(3), pp. 339-363. doi: 10.1177/0920203x10382338.

Lancaster, T. 2000. The incidential parameter problem since 1948. *Journal of Econometrics* 95, pp. 391-413.

Lapuente, V., Suzuki, K. and Van de Walle, S. 2020. Goats or wolves? Private sector managers in the public sector. *Governance*, 33(3), pp. 599-619. doi: 10.1111/gove.12462

Lapuente, V. and van de Walle, S. 2020. The effects of new public management on the quality of public services. *Governance* 33(3), pp. 461–475. doi: 10.1111/gove.12502.

Lather, P. 1986. Research as praxis. *Harvard Educational Review 56(*3), pp. 257-277.

Le Grand, J. 1982. *The strategy of equality.* London: George Allen & Unwin.

Lennart, B., Dong, H., Wang, K., Cai, W., and Vidon, D. 1996. The cost of coverage: rural health insurance in China. *Health Policy and Planning* 11(3), pp. 238-252.

Levi, M., Sacks, A. and Tyler, T. 2009. Conceptualizing legitimacy, measuring legitimating beliefs. *American Behavioural Scientists* 53(5), pp. 354–375.

Li, C. 2007 *The leadership of China's four major cities: a study of municipal party standing committees, China Leadership Monitor* 21, Washington D.C., Brookings Institution.Li, C. 2008. The philosophy of harmony in classical Confucianism. *Philosophy Compass* 3(3), pp. 423-435. doi: 10.1111/j.1747-9991.2008.00141.x

Li, H. and Zhang, J. 2007. Do high birth rates hamper economic growth? *Review of Economics and Statistics* 89 (1), pp. 110-117.

Li, H. and Zhou, L.A. 2005. Political turnover and economic performance: the incentive role of personnel control in China. *Journal of Public Economics* 89(9), pp. 1743–1762.

Li, J. 2015. The paradox of performance regimes: Strategic responses to target regimes in Chinese local government. *Public Administration* 93(4), pp. 1152–1167. doi: 10.1111/padm.12197.

Li, K.Q. 2020. *Report on the Work of the Government.* Beijing: State Council.

Li, L. 2010. Central-local relations in the People's Republic of China: trends, processes and impacts for policy implementation. *Public Administration and Development* 30(3), pp. 177-190. doi: 10.1002/pad.573.

Li, L. 2011. The challenges of healthcare reforms in China. *Public Health* 125(1), pp.

6-8. doi: 10.1016/j.puhe.2010.10.010.

Li, L., Chen, Q. and Powers, D. 2012. Chinese Healthcare Reform. *Modern China* 38(6), pp. 630-645. doi: 10.1177/0097700412457913.

Li, L. and Fu, H. 2017. China's health care system reform: progress and prospects. *International Journal of Health Planning and Management* 32, pp. 240-253. doi: 10.1002/hpm.2424

Li, M. and Zhu, A. 2004. China's public services privatization and poverty reduction: health care and education reform (privatization) in China and the impact on poverty. *United Nations Development Programme Policy Brief*. Available at: http://www.econ.utah.edu/~mli/CV/China%20Public%20Services%20Privatization%20and%20Poverty%20Reduction%20new.doc *[Accessed at: 8/5/2019].*

Li, P. 2016. *Replacement of business tax with a value-added tax - a countermeasure adopted by construction enterprises.* Proceedings of the 2nd International Conference on Economy, Management and Education Technology. Zhengzhou, China: Atlantis Press.

Li, S.K. and He, X. 2019. The impacts of marketization and subsidies on the treatment quality performance of the Chinese hospitals sector. *China Economic Review* 54, pp. 41–50.

Li, T. and Li, B. 2015. Factor analysis of cross-regional population flow in China: an empirical study with spatial econometrics models and panel data of 286 cities in China. *China Journal of Population Science* 2, pp. 73-83.

Li, W. 2019. The evolution of performance budgeting initiative in Guangdong, China. In Ho, A., de Jong, M. and Zhao, Z. eds. *Performance Budgeting Reform*. New York: Routledge.

Li, X. et al. 2017. The primary health-care system in China. *Lancet* 390(10112), pp. 2584–2594. doi: 10.1016/S0140-6736(17)33109-4.

Li, X. et al. 2020. Quality of primary health care in China: challenges and recommendations. *The Lancet* 395(10239), pp. 1802–1812. doi: 10.1016/S0140-6736(20)30122-7.

Li, Y. and Lee, C. 2010. Data envelopment analysis. *The Stata Journal* 10(2), pp. 267-280.

Liang, J. and Langbein, L. 2015. Performance management, high-powered incentives, and environmental policies in China. *International Public Management Journal* 18(3), pp. 346-385. doi: 10.1080/10967494.2015.1043167.

Lin, H. 2018. Privatising public hospitals: motivations and dilemmas. *Open Times* 2, pp. 178-196 (in Chinese).

Lin, J.Y. and Liu, Z. 2000. Fiscal decentralization and economic growth in China. *Economic Development and Culture Change* 49 (1), pp, 1-23.

Lin, R. and Chen, Z. 2018. Modified super-efficiency DEA models for solving infeasibility under non-negative data set. *INFOR: Information Systems and Operational Research* 56(3), pp. 265-285. doi: 10.1080/03155986.2017.1279898

Lind, D., Marchal, W. and Wathen, S. 2010. *Statistical techniques in business and economics.* 14 ed. New York, USA: McGraw-Hill.

Lindaman, K. and Thurmaier, J. 2002. Beyond efficiency and economy: an examination of basic needs and fiscal decentralization. *Economic Development and Cultural Change* 50(4), pp. 915–934.

Liu, C., Dou, X., Li, J. and Cai, L.A. 2020a. Analyzing government role in rural tourism development: an empirical investigation from China. *Journal of Rural Studies* 79, pp. 177-188.

Liu, D., Xu, C., Yu, Y., Rong, K. and Zhang, J. 2020b. Economic growth target, distortion of public expenditure and business cycle in China. *China Economic Review* 63, p. 101373. doi: 10.1016/j.chieco.2019.101373.

Liu, G.G., Vortherms, S.A. and Hong, X. 2017. China's health reform update. *Annual Review of Public Health* 38, pp. 431–448. doi: 10.1146/annurev-publhealth-031816-044247.

Liu, J. 2011. The difficulties and solutions: performance evaluation of primary health institutions. *Chinese Journal of Health Policy* 4(1), pp. 50-56.

Liu, K. 2021. Fostering the basic and substantial role of public finance in state governance (in Chinese). *Party-building Research* 12, pp. 8-10.

Liu, M., Shih, V. and Zhang, D. 2018. The fall of the old guards: explaining decentralization in China. *Studies in Comparative International Development* 53, pp. 379-403. doi: 10.1007/s12116-018-9267-0

Liu, M., Song, B. and Tao, R. 2006. Perspective on local governance reform in China. *China & World Economy* 14(2), pp. 16-31.

Liu, T., Li, Ji., Chen, J. and Yang, S. 2020c. Regional differences and influencing factors of allocation efficiency of rural public health resources in China. *Healthcare* 8(270), pp. 1-22. doi:10.3390/healthcare8030270

Liu, X. and Hsiao, W. 1995. The cost of escalation of social health insurance plans in China: its implications for public policy. *Social Science and Medicine* 41, pp. 1095–101.

Liu, Z., Yang, Y. and Wu, J. 2021. Participatory management, goal ambiguity, and gaming behaviors in performance management: evidence from Township Government Cadres in mainland China. *Public Performance and Management Review* 44(1), pp. 58–80. doi: 10.1080/15309576.2020.1768873.

Loikkanen, H.A. and Susiluoto, L. 2005. Cost efficiency of Finnish Municipalities in basic service provision 1994-2002. *Urban Public Economics Review* 4, pp. 39-63.

Lou, J. 2008. The reform of intergovernmental fiscal relations in China: lessons learned. In Lou, J. and Wang, S. eds. *Public Finance in China Reform and Growth for a Harmonious Society.* Washington: the Word Bank, pp. 155-170.

Luo, W. and Chen, S. 2010. Fiscal decentralization and public education provision in China. *Canadian Social Science* 6(4), pp. 28-41.

Lv, A., Luo, T. and Duckett, J. 2022. Centralization vs. decentralization in COVID-19 responses: lessons from China. *Journal of Health Politics, Policy and Law.* Doi: 10.1215/03616878-9626908

Lynn, L. E. 2001. The myth of the bureaucratic paradigm: what traditional public administration really stood for. *Public Administration Review* 61(2), pp. 144-160. doi: 10.1111/0033-3352.00016

Ma, L. 2016a. Performance feedback, government goal-setting and aspiration level adaptation: evidence from Chinese provinces. *Public Administration* 94(2), pp. 452–471. doi: 10.1111/padm.12225.

Ma, L. 2016b. Does super-department reform improve public service performance in China? *Public Management Review* 18(3), pp. 369-391. doi: 10.1080/14719037.2014.984624.

Ma, L. 2017a. Performance management and citizen satisfaction with the government: Evidence from Chinese municipalities. *Public Administration* 95(1), pp. 39-59. doi: 10.1111/padm.12275

Ma, L. and Christensen, T. 2020. Mapping the evolution of the central government apparatus in China. *International Review of Administrative Sciences* 86(1). doi: 10.1177/0020852317749025

Ma, X. et al. 2019. Realigning the incentive system for China's primary healthcare providers. *The BMJ* 365, pp. 8–11. doi: 10.1136/bmj.l2406.

Ma, Y. 2017b. Vertical environmental management: a panacea to the environmental enforcement gap in China?. *Chinese Journal of Environmental Law* 1(1), pp. 37-68. doi: 10.1163/24686042-12340004.

Mackenzie, N. and Knipe, S. 2006. Research dilemmas: paradigms, methods and methodology. *Issues In Educational Research* 16, pp. 1-11.

Magadi, M.A., Zulu, E.M. and Brockerhoff, M. 2003. The inequality of maternal health care in urban sub-Saharan Africa in the 1990s. *Population Studies* 57, pp. 347–366.

Man, J.Y. 2011. Local public finance in China: an overview. In Man, J.Y. and Hong, Y.H. eds. *China's Public Finance in Transition.* Hollis, New Hampshire: Puritan Press, pp. 3-20.

Mandl, U., Dierx, A. and Ilzkovitz, F. 2008. The effectiveness and efficiency of public spending. *Economic Papers 301*. Brussels, Belgium: European Commission

Martinez-Vazquez, J. 1999. The assignment of expenditure responsibilities. *World Bank Proceedings*. Washington, DC: World Bank Group.

Martinez-Vazquez, J. and McNab, R. 2003. Fiscal decentralization and economic growth. *World Development* 31(9), pp. 1597-1616. doi: 10.2139/ssrn.259281

Martinez-Vazquez, J. and Timofeev, A. 2009. Decentralization measures revisited. *International Studies Program Working Paper 09-13*. Atlanta, GA: Andrew Young School of Policy Studies, Georgia State University.

Marczyk, G., Dematteo, D., and Festinger, D. 2005. *Essentials of research design and methodology*. New Jersey: John Wiley and Sons, Inc.

Matthews, Z., Channon, A., Neal, S., Osrin, D., Madise, N. and Stones, W. 2010. Examining the "urban advantage" in maternal health care in developing countries. *PLoS Medicine* 7(9), e1000327. doi:10.1371/journal.pmed.1000327

May, T. 2011. *Social research: issues, methods and process*. 4th Edition. Maidenhead, Berks: Open University Press/Mc Graw-Hill.

McCandless, H.E. 2001. *A citizen's guide to public accountability. Changing the relationship between citizens and authorities*. Victoria B.C.: Trafford.

Mcloughlin, C. 2015. When does service delivery improve the legitimacy of a fragile or conflict-affected state? *Governance* 28(3), pp. 341-356. doi: 10.1111/gove.12091.

McLure, C.E. 1999. The tax assignment problem: conceptual and administrative considerations in achieving subnational fiscal autonomy. *World Bank Proceedings*. Washington, DC: World Bank Group.

Mei, C. and Wang, X. 2017. Political incentives and local policy innovations in China. *Journal of Chinese Political Science* (22), pp. 519-547. doi: 10.1007/s11366-017-9513-8

Meier, K.J. and O'Toole, L.J.Jr. 2002. Public management and organizational performance: the effect of managerial quality. *Journal of Policy Analysis and Management* 21 (4), pp, 629–43.

Meier, K.J., O'Toole, L., and Nicholson-Crotty, S. 2004. Multilevel governance and organizational performance: investigating the political-bureaucratic Labyrinth. *Journal of Policy Analysis and Management* 23(1). doi: 10.1002/pam.10177.

Meng, Q. et al. 2012. Trends in access to health services and financial protection in China between 2003 and 2011: a cross-sectional study. *Lancet* 379, pp. 805–814. doi: 10.1016/S0140-6736(12)60278-5

Meng, Q., Mills, A., Wang, L. and Han, Q. 2019. What can we learn from China's health system reform? *BMG* 365(l2349). doi: 10.1136/bmj.l2349

Mercier Ythier, J. 2010. Regular distributive efficiency and the distributive liberal social contract. *Journal of Public Economic Theory* 12(5), pp. 943-978. doi: 10.1111/j.1467-9779.2010.01481.x.

Mertens, D. M. 2005. *Research methods in education and psychology: integrating diversity with quantitative and qualitative approaches* 2nd ed. Thousand Oaks, CA: SAGE.

Millar, R., Jian, W., Mannion, R. and Miller, R. 2015. Healthcare reform in China: making sense of a policy experiment? *Journal of Health Organization and Management* 30(3), pp. 324-330. doi: 10.1108/JHOM-12-2014-0200

Milliken, J. and Krause, K. 2002. State failure, state collapse, and state reconstruction: concepts, lessons and strategies. *Development and Change* 33(5), pp. 753–774.

Montinola, G., Qian, Y. and Weingast, B.R. 1995. Federalism, Chinese style: the political basis for economic success in China. *World Politics* 48(1), pp. 50-81.

Morgan, D.L. 2007. Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research* 1(1), pp. 48-76. doi: 10.1177/2345678906292462

Morgan, W., Mutti, J. and Rickman, D. 1996. Tax exporting, regional economic growth, and welfare. *Journal of Urban Economics* 39(2), pp. 131-159. doi: 10.1006/juec.1996.0007.

Morin, J., Olsson, C., and Atikcan, E. O. 2021 *Research methods in the social sciences: an A-Z of key concepts*. Oxford: Oxford University Press.

Mosca, I. 2006. Is decentralisation the real solution? A three country study. *Health Policy* 77(1), pp. 113-120. doi: 10.1016/j.healthpol.2005.07.011

Mu, R., De Jong, M. and Koppenjan, J. 2019. Assessing and explaining interagency collaboration performance: a comparative case study of local governments in China. *Public Management Review* 21(4), pp. 581-605. doi: 10.1080/14719037.2018.1508607

Mueller, D. 1979. *Public choice*. Cambridge: Cambridge University Press.

Mulgan, R. 1997. The processes of public accountability. *Australian Journal of Public Administration* 56(1), pp. 25–36.

Muramatsu, M., Matsunami, J. 2003. The late and sudden emergence of new public management reforms in Japan. In Wollmann, H. ed. *Evaluation in Public-Sector Reform: Concepts and Practice in International Perspective*. Cheltenham: Edward Elgar, pp. 169–181.

Musgrave, R.M. 1959. *The Theory of Public Finance*. NY: McGraw-Hill.

Narayan, P. and Peng, X. 2006. An econometric analysis of the determinants of fertility for China, 1952–2000. *Journal of Chinese Economic and Business Studies* 4(2), pp. 165-183. doi: 10.1080/14765280600737039.

Nathan, A. J. 2003. Authoritarian resilience. *Journal of Democracy* 14, pp. 6–17.

Ngok, K. and Zhu, Y. 2010. In search of harmonious society in China: a social policy response. In Mok, K.H. and Ku, Y.W. eds. *Social Cohesion in Greater China: Challenges for Social Policy and Governance*. Singapore: World Scientific Publishing Company.

Nguyen, P.A., Kecskés, A., and Mansi, S. 2020. Does corporate social responsibility create shareholder value? The importance of long-term investors. *Journal of Banking and Finance* 112. doi: 10.1016/j.jbankfin.2017.09.013.

Nielsen, P.A. 2014 Learning from performance feedback: performance information, aspiration levels, and managerial priorities. *Public Administration* 92(1), pp. 142–160.

Niu, M. 2013. Fiscal decentralization in China revisited. *Australian Journal of Public Administration* 72(3), pp. 251–263. doi: 10.1111/1467-8500.12026.

Norregaard, J. 1997. Tax assignment. In: Ter-Minassian, T. ed. *Fiscal federalism in theory and practice*. Washington: International Monetary Funds, pp. 49-72.

O'Flynn, J. 2007. From New Public Management to public value: paradigmatic change and managerial implications. *The Australian Journal of Public Administration* 66(3), pp. 353-366.

O'Mahony, M. and Stevens, P. 2006. International comparisons of output and productivity in public service provision: a review. In Boyne, G.A., Meier, K.J., O'Toole, L.J. Jr. and Walker, R.M. eds. *Public service performance: perspectives on measurement and management*. Cambridge, UK: Cambridge University Press, pp. 233-253.

Oates, W. 1972. *Fiscal Federalism*. New York: Harcourt Brace Jovanovich.

Oates, W. 1999. An essay on fiscal federalism. *Journal of Economic Literature* 37(3), pp. 1120-1149. doi: 10.1257/jel.37.3.1120.

Oates, W. 2005. Towards a second-generation theory of fiscal federalism. *International Tax and Public Finance* 12(4), pp. 349–373.

Oates, W. 2006. On the theory and practice of fiscal decentralization. *IFIR Working Paper No. 2006-05.* Lexington, KY: Institute for Federalism and Intergovernmental Relations.

OECD. 2003. Fiscal relations across levels of government. *OECD Economic Outlook* 74(2), pp. 143-160.

OECD. 2005. *Modernising government: the way forward*. Paris: OECD

OECD. 2011. *Supporting state-building in situations of conflict and fragility: policy Guidance*. Paris: OECD.

OECD. 2016. *Fiscal federalism 2016: making decentralisation work.* Paris: OECD publishing.

O'Flynn, J. 2007. From New Public Management to public value: paradigmatic change and managerial implications. *The Australian Journal of Public Administration* 66(3), pp. 353-366. doi:10.1111/j.1467-8500.2007.00545.x

Onwuegbuzie, A. J. Teddlie, C. 2003. A framework for analyzing data in mixed methods research. In Tashakkori, A and Teddlie, C. eds. *Handbook of mixed methods in social and behavioral research.* Thousand Oaks, CA: SAGE, pp. 351-383.

Ostrom, V., Tiebout, C. and Warren, R. 1961. The organisation of government in metropolitan areas: a theoretical inquiry. *American Political Science Review* 55(4), pp. 831–42.

Otsuka, A., Goto, M. and Sueyoshi, T. 2014. Cost-efficiency of Japanese local governments: effects of decentralization and regional integration *Regional Studies, Regional Science* 1(1), pp. 207-220. doi: 10.1080/21681376.2014.953196

Palmer, A.J. 1993. Performance measurement in local government. *Public Money & Management* 13(4), pp. 31-6.

Patton. M. Q. 1990. *Qualitative evaluation and research methods*. 2nd ed. Newbury

Persson, T. and Tabellini, G. 2000. *Political economics: explaining economic policy*. Cambridge, MA: MIT Press.

Pollitt, C. 1988. Bringing consumers into performance measurement: concepts, consequences and constraints. *Policy and Politics* 16(2), pp. 77–87.

Pollitt, C. 2003. *The essential public manager*. London: Open university press/McGraw-Hill.

Pollitt, C. and Dan, S. 2011. *The impacts of the new public management in Europe: a meta-analysis.* COCOPS Research Report.

Porcelli, F. 2009. Effects of FD and electoral accountability efficiency evidence from the Italian health care sector. *Institu' d'Economia de Barcelona (IEB) working Papers 29.* Barcelona, Spain: IEB.

Prieto, D. C. and Saez, M. P. 2006. Decentralisation and health care outcomes: an empirical analysis within the European Union. *Studies on the Spanish Economy 220, FEDEA*. Cantabria, Spain: Department of Economics, University of Cantabria.

Provan, K. and Kenis, P. 2008. Modes of network governance: structure, management, and effectiveness. *Journal of Public Administration Research and Theory* 18(2), pp. 229-252. doi: 10.1093/jopart/mum015.

Prud'homme, R. 1995. The dangers of decentralization. *The World Bank Research Observer* 10(2), pp. 201–220. doi: 10.1093/wbro/10.2.201

Psycharis Y., Zoi M., and Ilipoulou, S. 2016. Decentralization and local government fiscal autonomy: evidence from the Greek municipalities. *Environment and Planning C: Government and Policy* 34, pp. 262–280

Putnam, R. D. 1993. *Making democracy work: civic traditions in modern Italy.* Princeton, NJ: Princeton University Press.

Qian, D., Pong, R.W., Yin, A., Nagarajan, K.V. and Meng, Q. 2009. Determinants of health care demand in poor, rural China: the case of Gansu Province. *Health Policy and Planning* 24(5), pp. 324-334. doi: 10.1093/heapol/czp016.

Qian, W., Cheng, X., Lu, G., Zhu, L. and Li, F. 2019. Fiscal decentralization, local competitions and sustainability of medical insurance funds: evidence from China. *Sustainability* 11(8), p. 2437. doi: 10.3390/su11082437.

Qian, Y. and Roland, G. 1996. The soft budget constraint in China. *Japan and the World Economy* 8, pp. 207-223.

Qian, Y. and Roland, G. 1998. Federalism and the soft budget constraint. *The American Economic Review* 88(5), pp. 1143-1162.

Qian, Y. and Weingast, B.R. 1996. China's transition to markets: market-preserving federalism, Chinese style. *Journal of Policy Reform* 1, pp. 149–185

Qian, Y. and Weingast, B.R. 1997. Federalism as a commitment to preserving market incentives. *Journal of Economic Perspectives* 11(4), pp. 83–92.

Qiu, S. and Macnaughton, G. 2017. Mechanisms of accountability for the realization of the right to health in China. *Health and Human Rights Journal* 19(1), pp. 279-292.

Qiu, W., Chu, C., Wang, K., Meng, Y., Yang, Y. and Mao, A. 2020. Collaboration and cooperation of public health service management in China. *Disaster Med Public Health Prep* 6, pp. 1-7. doi: 10.1017/dmp.2019.142.

Qu, J. and Lv, X. 2021. The response measures to the coronavirus disease 2019 outbreak in China. *Open Forum Infectious Diseases* 8 (2). doi: 10.1093/ofid/ofab014

Quentin, W., Partanen, V. M. and Brownwood, I. 2009. Measuring healthcare quality. In: Busse, R., Klazinga, N., Panteli, D. and Quentin, W. eds. *Improving healthcare quality in Europe*. Paris: OECD, pp. 31-62.

Ramesh, M. 2008. Autonomy and control in public hospital reforms in Singapore. *American Review of Public Administration* 38, pp. 62–79.

Ramesh, M., Wu, X. and He, A. J. 2014. Health governance and healthcare reforms in China. *Health Policy and Planning* 29(6), pp. 663–672. doi: 10.1093/heapol/czs109.

Rao, M. G. and Singh, N. 2007. The political economy of India's federal system and its reform. *Publius The Journal of Federalism* 37, pp. 26-44.

Raphael, G. 2013. A DEA-based Malmquist Productivity Index approach in assessing performance of commercial banks: evidence from Tanzania. *European Journal of Business and Management* 5(6), pp. 25-34.

Ratigan, K. 2022. Riding the tiger of performance legitimacy? Chinese villagers' satisfaction with state healthcare provision. *International Political Science Review* 43(2), pp. 259–278. doi: 10.1177/0192512120927482.

Rayp, G. and Van De Sijpe, N. 2007. Measuring and explaining government efficiency in developing countries. *Journal of Development Studies* 43(2), pp. 360–381. doi: 10.1080/00220380601125230

Reingewertz, Y. 2014. Fiscal decentralization - a survey of the empirical literature. *MPRA Working Paper No. 59889*. Munich: Munich Personal RePEc Archive.

Reiter, R. and Klenk, T. 2019. The manifold meanings of 'post-New Public Management' – a systematic literature review. *International Review of Administrative Sciences* 85(1), pp. 11-27.

Reiter, M. and Weichenrieder, A. 1999. Public goods, club goods, and the measurement of crowding. *Journal of Urban Economics* 46(1), pp. 69-79. doi: 10.1006/juec.1998.2112.

Remenyi, D., Williams, B., Money, A. and Swartz, E. 1998. *Doing research in business and Management: an introduction to process and method.* London: Sage.

Renzio, P., Gomez, P. and Sheppard, J. 2005. Budget transparency and development in resource-dependent countries. *Social Science Journal* 57(s1), pp. 57-69. doi: 10.1111/j.1468-2451.2009.00706.x

Rezk, E. 1999. Experiences of decentralisation and intergovernmental fiscal relations in Latin America. In: Fukasaku, K. and De Mello, L. R. eds. *Fiscal decentralisation in emerging economies.* Paris: OECD, pp. 101-120.

Richardus, J.H., Graafmans, W.C., Verloove-Vanhorick, S.P. and Mackenbach, J.P. 1998. The perinatal mortality rate as an indicator of quality of care in international Comparisons. *Medical Care* 36(1), pp. 54–66. doi: 10.1097/00005650-199801000-00007.

Rixon, D. 2010. Stakeholder engagement in public sector agencies: ascending the rungs of the accountability ladder. *International Journal of Public Administration* 33(7), pp. 347-356. doi: 10.1080/01900691003606006.

Rixon, D. and Ellwood, S. 2010. Reporting for public sector agencies: a stakeholder model. In Osborne, S.P. and Ball, A. eds. *Social Accounting and Public Management Accountability for the Public Good.* New York: Routledge, pp. 131-144.

Robalino, D.A., Picazo, O.F. and Albertus, V. 2001. Does fiscal decentralization improve health outcome? Evidence from a cross-country analysis. *Policy Research Working Paper 2565*. Washington: World Bank.

Rocha, F., Orellano, V.F. and Nishijima, M. 2016. Health spending autonomy and infant mortality rates: a matter of local administrative capacity? *The Journal of Developing Areas* 50(2), pp. 293-309.

Rodden, J. 2003. Reviving leviathan: fiscal federalism and the growth of government. *International Organization* 57(4), pp. 695–729. doi: 10.1017/s0020818303574021.

Rodden, J. 2004. Comparative federalism and decentralization: on meaning and measurement. *Comparative Politics* 36(4), pp. 481. doi: 10.2307/4150172.

Rodríguez-Pose, A. and Bwire, A. 2004. The economic (in)efficiency of devolution. *Environment and Planning A* 36(11), pp. 1907–1928

Rodríguez-Pose, A. and Gill, N. 2003. The global trend towards devolution and its implications. *Environment and Planning C: Government and Policy* 21(3), pp. 333–351.

Rodríguez-Pose, A. and Gill, N. 2005. On the 'economic dividend' of devolution. *Regional Studies* 39(4), pp. 405-420. doi: 10.1080/00343400500128390

Rossman, G. B. and Wilson. B. L. 1985. Numbers and words: combining quantitative and qualitative methods in a single large-scale evaluation study. *Evaluation Review* 9(5), pp. 627-643.

Rothstein, B. 2009. Creating political legitimacy: electoral democracy versus the quality of government. *American Behavioral Scientist* 53(3), pp. 311-330 doi: 10.2139/ssrn.1338615.

Rothstein, B. 2012. Political legitimacy for public administration. In Peters, B. G. and Pierre, J. eds. *The SAGE Handbook of Public Administration*. London: SAGE.

Rutherford, A. and Meier, K.J. 2015. Managerial goals in a performance-driven system: theory and empirical tests in higher education. *Public Administration* 93(1), pp. 17–33. doi: 10.1111/padm.12124.

Saich, T. 2000. Negotiating the State: The development of social organizations in China. *The China Quarterly* 161, pp. 124-141. doi: 10.1017/s0305741000003969.

Salman, Y. 2021. Public management reforms in Pakistan. *Public Management Review* 23(12), pp. 1725-1735. doi: 10.1080/14719037.2020.1850084

Salmon, P. 1987. Decentralisation as an Incentive Scheme. *Oxford Review of Economic Policy* 3(2): 24–43.

Samuelson, P. 1954. Pure theory of public expenditures. *Review of Economics and Statistics* 36(4). pp. 387–389.

Samuelson, P.A. and Nordhaus, W.D. 2005. *Economics* 18th ed. NY: McGraw-Hill.

Samut, K.P. and Cafrı, R. 2016. Analysis of the efficiency determinants of health systems in OECD countries by DEA and panel Tobit. *Social Indicators Research* 129(1), pp. 113-132. doi: 10.1007/s11205-015-1094-3.

Sandberg, J. 2005. How do we justify knowledge produced within interpretive approaches? *Organizational Research Methods* 8(1), pp. 41-68.

Sarkis, J. 2007. Preparing your data for DEA. In: Zhu, J. and Cook, W. D. eds. *Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis.* New York: Springer, pp. 305-320.

Saunders, M., Lewis, P. and Thornhill, A. 2007. *Research Methods for Business Students* 4th ed. Harlow: Pearson Education.

Saunders, M., Lewis, P. and Thornhill, A. 2016. *Research methods for business students*. 7th ed. Harlow: Pearson Education.

Schalk, J. 2011. Linking stakeholder involvement to policy performance: nonlinear and stakeholder-specific effects in Dutch local government policy making. *Paper to* *be presented at the 2011 Public Management Research Conference, Maxwell School at Syracuse University*. New York, USA, June 2-4, 2011.

Schaltegger, C.A. and Benno, T. 2007. Government accountability and fiscal discipline: a panel analysis using Swiss data. *Journal of Public Economics* 91(1-2), pp. 117–140.

Scott, J.C. 2008. *Seeing like a state*. New Haven: Yale University Press.

Seabright, P. 1996. Accountability and decentralization in government: an incomplete contracts model. *European Economic Review* 40(1), pp. 61–89. doi: 10.1016/0014-2921(95)00055-0

Sharma, M. and Sharma, P. 2019. Dynamics of employment and economic growth: a comparative study of China and India. *Millennial Asia* 10(3), pp. 417–437. doi: 10.1177/0976399619879861.

Shaw, R. 2012. La nouvelle gestion publique en Australie: passé, présent et futur (New public management in Australia: past, present and future). *Pouvoirs* 2, pp. 117-132.

Shen, C., Jin, J. and Zou, H. F. 2012. Fiscal decentralization in China: history, impact, challenges and next steps. *Annals of Economics and Finance* 13(1), pp. 1–51.

Shen, C., Zhao, X. and Zou, H. 2014. Fiscal decentralization and public services provision in China. *Annals of Economics and Finance* 15(1), pp. 135-160.

Shen, Y. 2001. Positive interaction: the new paradigm of the central-local relations and its policy implications. *Shanghai Administration Institute Journal* 2, pp. 50–56.

Shen, Y. and Li, B. 2020. Policy coordination in the talent war to achieve economic upgrading: the case of four Chinese cities. *Policy Studies* 43(3), pp. 443-463. doi: 10.1080/01442872.2020.1738368

Shi, P. and Ni, M. 2017. The allocation mechanism of authority and accountability in Chinese government hierarchies: a case study of the “Earmarked Project System”. *The Journal of Chinese Sociology* 4(1). doi: 10.1186/s40711-017-0051-8.

Shi, S.J. 2012. Social policy learning and diffusion in China: the rise of welfare regions. *Policy & Politics* 40(3), pp. 367-385.

Shi, T. 2001. Cultural values and political trust: a comparison of the People’s Republic of China and Taiwan. *Comparative Politics* 33(4), pp. 401–419.

Shih, V., Zhang, Q. and Liu, M. 2004. The ‘eating budget:’ determining fiscal transfers under predatory fiscal federalism. *Proceedings of the American Political Science Association Annual Conference.* Chicago, IL.

Sidel, V. W. and Sidel, R. 1975. The development of health care services in the People's Republic of China. *World Development* 3(7-8), pp. 539-549.

Silverman, J.M. 1992. *Public sector decentralization: economic policy and sector investment programs.* Technical Paper No. 188. Washington, DC: World Bank

Simon, H. A. 1976. *Administrative behavior: a study of decision-making processes in administrative organization* 3rd ed. London: Macmillan.

Sinclair, A. 1995. The chameleon of accountability: forms and discourses. *Accounting Organizations and Society* 20(2/3), pp. 219–37.

Singh, A. 2007. Health reforms in Rajasthan: an empirical study on agencification. *Policy and Society* 26(2), pp. 61-82. doi: 10.1016/S1449-4035(07)70108-X

Smith, D. and Brame, R. 2003. Tobit models in social science research. *Sociological Methods and Research* 31(3), pp. 364-388. doi: 10.1177/0049124102239080.

Smith, K.W. and Sasaki, M.S. 1979. Decreasing multicollinearity: a method for models with multiplicative functions. *Sociological Methods & Research* 8, pp, 35–56.

Smoke, P. 2001. Fiscal decentralization in developing countries: a review of current Concepts and practice. *Democracy, Governance and Human Rights Programme Paper Number 2.* Geneva: United Nations Research Institute for Social Development.

Snape, H. and Wang, W. 2020. Finding a place for the Party: debunking the "party-state" and rethinking the state-society relationship in China's one-party system. *Journal of Chinese Governance* 5(4), pp. 477–502. doi: 10.1080/23812346.2020.1796411.

So, B. 2014. Civic engagement in the performance evaluation of the public sector in China: building horizontal accountability to enhance vertical accountability. *Public Management Review* 16(3), pp. 341-357. doi: 10.1080/14719037.2013.770055

Song, Y. 2014. What should economists know about the current Chinese hukou system? *China Economic Review* 29, pp. 200-212. doi: 10.1016/j.chieco.2014.04.012.

Soss, J., Fording, R., and Schram, S. F. 2011. The organization of discipline: from performance management to perversity and punishment. *Journal of Public Administration Research and Theory*, 21(2), pp. 203–232.

Soto, V.E., Farfan, M.I. and Lorant, V. 2012. FD and infant mortality rate: the Colombian case. *Social Science & Medicine* 74(9), pp. 1426-1434. doi: 10.1016/j.socscimed.2011.12.051

Sow, M. and Razafimahefa, I.F. 2015. Fiscal decentralization and the efficiency of public service delivery. *IMF Working Paper 15/59.* Washington: International Monetary Funds.

Stastná, L. and Gregor, M. 2015. Public sector efficiency in transition and beyond: evidence from Czech local governments. *Applied Economics* 47(7), pp. 680-699. doi: 10.1080/00036846.2014.978077

State Council. 2016. *Opinions of the CPC central committee and the state council on further deepening the reform of the medical and healthcare system*. Beijing: the General Office of the State Council.

State Council. 2018. *Plan for the reform of the division of financial powers and expenditure responsibilities between Central and local governments in basic public services*. Beijing: the General Office of the State Council.

Stegarescu, D. 2005. Public sector decentralisation: measurement concepts and recent international trends. Fiscal Studies 26(3), pp. 301–333.

Stevens, P.A. 2005. Assessing the performance of local government. *National Institute Economic Review* 193, pp. 90–101.

Stewart, J. 1984. The role of information in public accountability. In Hopwood, A. and Tomkins, C. eds. *Issues in Public Sector Accounting*. Oxford: Philip Allan Publishers Limited, pp. 13 – 34.

Stone, B. 1995. Administrative accountability in the Westminster Democracies: toward a conceptual framework. *Governance* 8(4), pp. 505–26.

Strom, K. 2000. Delegation and accountability in parliamentary democracies. *European Journal of Political Research* 37(3), pp. 261-289. doi: 10.1111/1475-6765.00513

Su, M. and Zhao, Q. 2004. *China's fiscal decentralization reform*. Beijing: Research Institute for Fiscal Science, Ministry of Finance.

Sueyoshi, T. and Goto, M. 2013. Pitfalls and remedies in DEA applications: how to handle an occurrence of zero in multipliers by strong complementary slackness conditions. *Engineering* 5, pp. 29-34. doi: 10.4236/eng.2013.55a005.

Suits, D. 1957. Use of dummy variables in regression equations. *Journal of the American Statistical Association* 52(280), p. 548. doi: 10.2307/2281705.

Sulle, A. 2010. The application of New Public Management doctrines in the developing world: An exploratory study of the autonomy and control of executive agencies in Tanzania. *Public Administration and Development* 30(5), pp. 345-354. doi: 10.1002/pad.580

Sun, X. 2021. *Off-site settlement of medical insurance exceeds 10 million times*. People's Daily, 02 December 2021. https://cn.chinadaily.com.cn/a/202112/02/WS61a8684ba3107be4979fb0af.html [accessed 19 February 2022]

Tan, J. and Zhao, J. 2019. The rise of public–private partnerships in China: an effective financing approach for infrastructure investment? *Public Administration Review* 79(4), pp. 514-518. doi: 10.1111/puar.13046

Tan, X., Liu, X. and Shao, H. 2017. Healthy China 2030: a vision for health care. *Value in Health Regional Issues* 12, pp. 112-114. doi: 10.1016/j.vhri.2017.04.001.

Tan, X., Zhang, Y. and Shao, H. 2018. Healthy China 2030, a breakthrough for improving health. *Global Health Promotion* 26(4), pp. 96-99. doi:10.1177/1757975917743533

Tanzi, V. 1995. Fiscal federalism and decentralization: a review of some efficiency and macroeconomic aspects. In: World Bank ed. *Annual World Bank Conference on Development Economics* 1995. Washington, DC: World Bank.

Tashakkori, A. and Teddlie, C. 1998. *Mixed methodology: combining qualitative and quantitative approaches.* Thousand Oaks, CA: SAGE.

Taylor, J. 2021. Public officials' gaming of performance measures and targets: the nexus between motivation and opportunity. *Public Performance and Management Review* 44(2), pp. 272–293. doi: 10.1080/15309576.2020.1744454.

Ter-Minassian, T. 1997. Intergovernmental fiscal relations in a macroeconomic perspective: an overview. In: Ter-Minassian, T. eds. *Fiscal federalism in theory and practice*. Washington: International Monetary Funds, pp. 3-24.

Then, F. and Sussmuth-Dyckerhoff, C. 2017. China's healthcare reform: status and outlook. In Burns, L.B. and Liu, G.G. eds. *China's Healthcare System and Reform*. Singapore: Cambridge University Press, pp. 137-149.

Thomas, A.B. 2004. *Research skills for management studies.* London: Routledge.

Thompson, J.R. 2000. Reinvention as reform: assessing the National Performance Review. *Public management Review* 60(6), pp. 508-521.

Tian, C., Teng, Y. and Guo, H. 2020. Resource flow or yardstick competition: spatial dependence of educational expenditure among prefecture-level cities. *International Public Management Journal* 24(5), pp. 691-709. doi: 10.1080/10967494.2020.1724218.

Tiebout, C. M. 1956. A pure theory of local expenditures. *Journal of political economy* 64(5), pp. 416-424.

Tobin, J. 1958. Estimation of relationships for limited dependent variables. *Econometrica* 26, pp. 24-36

Tone, K. 2001. A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research* 130(3), pp. 498-509. doi: 10.1016/s0377-2217(99)00407-5.

Treisman, D. 2002. Decentralization and the quality of government. *Working paper, Department of Political Science*. Los Angeles: UCLA.

Tsen, W.H. 2010. Exports, domestic demand, and economic growth in China: Granger causality analysis. *Review of Development Economics* 14(3), pp. 625–639. doi: 10.1111/j.1467-9361.2010.00578.x.

Tsui K. and Wang, Y. 2008. Decentralization with political trump: vertical control, local accountability and regional disparities in China. *China Economic Review* 19(1), pp. 18-31. doi: 10.1016/j.chieco.2006.12.003.

Tsui, K. and Wang, Y. 2004. Between separate stoves and a single menu: fiscal decentralization in China. *China Quarterly* 177, pp. 71-90.

Tu, Z., Hu, T. and Shen, R. 2019. Evaluating public participation impact on environmental protection and ecological efficiency in China: evidence from PITI disclosure. *China Economic Review* 55, pp. 111-123.

Uchimura, H. and Jutting, J. 2009. Fiscal decentralization, Chinese style: good for health outcomes? *World Development* 37(12), pp. 1926-1934. doi: 10.1016/j.worlddev.2009.06.007.

Van den Bergh, J. C. 2009. The GDP paradox. *Journal of Economic Psychology* 30(2), pp. 117-135. doi: 10.1016/j.joep.2008.12.001

Van Dijk, C. and Van Den Ende, J. 2002. Suggestion systems: transferring employee creativity into practicable ideas. *R and D Management* 32(5), pp. 387–395. doi: 10.1111/1467-9310.00270

Varabyova, Y. and Schreyögg, J. 2013. International comparisons of the technical efficiency of the hospital sector: Panel data analysis of OECD countries using parametric and non-parametric approaches. *Health Policy* 112(1-2), pp. 70-79. doi: 10.1016/j.healthpol.2013.03.003.

Verbeek, M. 2004. *A guide to modern econometrics.* Chichester: John Wiley and Sons Press.

Verbeeten, F. and Spekle, R. 2015. Management control, results-oriented culture and public sector performance: empirical evidence on New Public Management. *Organization Studies* 36(7), pp. 953–978. doi: 10.1177/0170840615580014

Vogel, J. P. et al. 2014. Maternal complications and perinatal mortality: findings of the World Health Organization multicountry survey on maternal and newborn Health. *BJOG: an international journal of obstetrics and gynaecology* 121(S1), pp. 76–88. doi: 10.1111/1471-0528.12633.

Vrangbaek, K. 2007. Towards a typology for decentralization in health care. In Saltman, R. B., Bankauskaite, V. and Vrangbaek, K. eds. *Decentralization in Health Care*. New York: Open University Press, pp. 44–62.

Walker, H. 1937. *Public administration in the United States*. New York: Farrar & Reinhart.

Walker, R.M., Boyne, G.A. and Brewer, G. A. 2010. *Public management and performance: research directions*. Cambridge: Cambridge University Press.

Walker, R.M. and Wu, J. 2010. Future prospects for performance management in Chinese city governments. *Administration & Society* 42(s1), pp. 34–55. doi:10.1177/0095399710361853

Wang, Y.M. and Lan, Y.X. 2011. Measuring Malmquist productivity index: a new approach based on double frontiers data envelopment analysis. *Mathematical and Computer Modelling* 54, pp, 2760-2771. doi: 10.1016/j.mcm.2011.06.064

Wang, Y. and Wang, W. 2007. Escalation in health expenditure and difficult access to care in China. In: Chen J. and Wang, Y. eds. *Development of Social Security in China: Health Services and Protection in Transformation*. Beijing: Social Science Academic Press.

Wang, A.L. 2013. The search for sustainable legitimacy: environmental law and bureaucracy in China. *Harvard Environmental Law Review* 37(2), pp. 365–440. doi: 10.2139/ssrn.2128167.

Wang, E. 2010. Fiscal decentralization and revenue/expenditure disparities in China. *Eurasian Geography and Economics* 51(6), pp. 744-766. doi: 10.2747/1539-7216.51.6.744.

Wang, J. 2006. *Political economy of village governance in contemporary China.* PhD Thesis, Indiana University.

Wang, W., Liu, W. and Mingers, J. 2015. A systemic method for organisational stakeholder identification and analysis using Soft Systems Methodology (SSM). *European Journal of Operational Research* 246(2), pp. 562-574. doi: 10.1016/j.ejor.2015.05.014.

Wang, W., Zheng, X. and Zhao, Z. 2011. Fiscal reform and public education spending: a quasi-natural experiment of fiscal decentralization in China. *Publius: The Journal of Federalism* 42(2), pp. 334-356. doi: 10.1093/publius/pjr039.

Wang, X. and Herd, R. 2013. The system of revenue sharing and fiscal transfers in China. *OECD Economics Department Working Papers No. 1030.* Paris: OECD.

Wang, Y. and Wang, W. 2007. Escalation in health expenditure and difficult access to care in China. In: Chen, J. and Wang, Y. eds. *Development of Social Security in*

China: Health Services and Protection in Transformation. Beijing: Social Science Academic Press (in Chinese).

Wang, Z. and Ma, L. 2014. Fiscal decentralization in China: a literature review. *Annals of Economics and Finance* 15(2), pp. 751-770.

Wang, Z., Zhou, L. and Peng, F. 2018. Fiscal expenditure, scale comparison and public service satisfaction: an analysis based on county-level medical data. *Chinese Public Administration* 14, pp. 49-54.

Ware, G.T., Shaun Moss, J., Campos, E., and Noone, G.P. 2007. Corruption in public procurement: a perennial challenge. In Campos, J.E. and Pradhan, S. eds. *The many faces of corruption: Tracking vulnerabilities at the sector level*. Washington, DC: The World Bank.

Weatherley, R. and Magee, C. 2018. Using the past to legitimise the present: the portrayal of good governance in Chinese history textbooks. *Journal of Current Chinese Affairs* 47(1), pp. 41-69. doi: 10.1177/186810261804700102

Weingast, B.R. 2009. Second generation fiscal federalism: the implications of fiscal incentives. *Journal of Urban Economics* 65(3), pp. 279-293. doi: 10.1016/j.jue.2008.12.005

Whaites, A. 2008. *States in development: understanding state-building*. London: Department for International Development.

Whiting, S.H. 2001. *Power and wealth in rural China: the political economy of institutional change*. Cambridge, UK: Cambridge University Press.

Widmer, P. and Zweifel, P. 2012. Fiscal equalization, Tiebout competition, and incentives for efficiency in a federalist country. *Public Finance Review* 40(1), pp. 3-29. doi: 10.1177/1091142111422437.

Widstrom, E., Linna, M. and Niskanen, T. 2004. Productive efficiency and its determinants in the Finnish public dental service. *Community Dent Oral Epidemiol* 32, pp. 31-40.

Wiesner, E. 2003. *Fiscal federalism in Latin America from entitlements to markets*. Washington, DC: Inter-American Development Bank.

Willis, J. W. 2007. *Foundations of qualitative research: interpretive and critical approaches*. Thousand Oaks, CA: SAGE.

Wilson, J. 1986. A theory of interregional tax competition. *Journal of Urban Economics* 19(3), pp. 296-315. doi: 10.1016/0094-1190(86)90045-8.

Wilson, J. 2010. *Essentials of business research: a guide to doing your research project*. UK: Sage Publications.

Wingender, P. 2018. Intergovernmental Fiscal Reform in China. *IMF Working Paper 18/88*. Washington: International Monetary Funds.

Wollmann, H. 2003. *Evaluation in public-sector reform: concepts and practice in international perspective.* Cheltenham: Edward Elgar Publishing

Wong, C. 1991. Central–local relations in an era of fiscal decline: the paradox of fiscal decentralization in post-Mao China. *The China Quarterly* 128, pp. 691-715.

Wong, C. 2010. *Public sector reforms toward building the harmonious society in China.* Paper prepared for the China Economic Research and Advisory Programme. Oxford, UK: University of Oxford.

Wong, C. and Bird, R. 2008. China's fiscal system: a work in progress. In Brandt, L. and Rawski, T. eds. *China's Great Transformation: Origins, Mechanism and Consequences of the Post-Reform Economic Boom*. New York: Cambridge University Press, pp. 429-466.

Wooldridge, J.M. 2001. *Econometric analysis of cross section and panel data.* Cambridge, Mass: MIT Press.

World Bank. 1999. Decentralization: rethinking government. *World Development Report 1999/2000.* Washington DC: World Bank.

World Bank. 2003. *World development report 2004. Making services work for poor people.* Washington, DC: World Bank.

World Bank. 2018. *World Bank open data.* Available at: https://data.worldbank.org/ [Accessed: 04 August 2022].

Worthington, A.C. 2000. Cost efficiency in Australian local government: a comparative analysis of mathematical programming and econometrical approaches. *Financial Accountability and Management* 16(3), pp. 201–223. doi: 10.1111/1468-0408.00105.

Wu, F., Xu, J. and Yeh, A.G. 2006. *Urban development in post-reform China*. London: Routledge.

Wu, J. and Walker, R.M. 2020. Public management in China: reform, innovation and governance. *International Public Management Journal* 23(3) v-ix. doi: 10.1080/10967494.2020.1784646

Wu, H., Li, Y., Hao, Y., Ren, S. and Zhang, P. 2020. Environmental decentralization, local government competition, and regional green development: evidence from China. *Science of The Total Environment* 708(15), 135805. doi: 10.1016/j.scitotenv.2019.135085

Wu, X., Ramesh, M. and Yu, J. 2017. Autonomy and performance: decentralization reforms in Zhejiang Province, China. *Public Administration and Development* 37, pp. 94-109. doi: 10.1002/pad.1786

Xiao, H. et al. 2021. The impact of the COVID-19 pandemic on health services utilization in China: time-series analyses for 2016–2020. *The Lancet Regional Health - Western Pacific* 9(100122).

Xiao, Y., Husain, L. and Bloom, G. 2018. Evaluation and learning in complex, rapidly changing health systems: China's management of health sector reform. *Globalization and Health* 14(1). doi: 10.1186/s12992-018-0429-7.

Xie, Y. and He, B. 2010. Autonomy, local decentralization, and local protection. *Socialism Study* 2, pp. 94–100

Xing, C. and Zhang, R. 2021. COVID-19 in China: responses, challenges and implications for the health system. *Healthcare* 9 (82).

Xu, H., He, J., Zheng, K. and Kwan, H. 2010. The Chinese healthcare system. In: Johnson, J. and Stoskopf, C. eds. *Comparative Health Systems: Global Perspective*. Boston, MA: Jones & Bartlett.

Xu, W. and Lin, J. 2022. Fiscal decentralization, public health expenditure and public health–evidence from China. *Frontiers in Public Health* 10. doi: 10.3389/fpubh.2022.773728.

Yamamoto, K. 2006. Performance of semi-autonomous public bodies: linkage between autonomy and performance in Japanese agencies. *Public Administration and Development* 26(1), pp. 35–44. https://doi.org/10.1002/pad.369

Ye, L. and Ni, X. 2013. Performance management in China: in pursuit of accountability and legitimacy. *Proceedings of the International Conference on Collaboration among Government, Market, and Society: Forging Partnerships and Encouraging Competition*. Shanghai 25-26 May 2013.

Yip, P. and Tsang, E. 2007. Interpreting dummy variables and their interaction effects in strategy research. *Strategic Organization* 5(1), pp. 13-30. doi: 10.1177/1476127006073512.

Yip, W. et al. 2019. 10 years of health-care reform in China: progress and gaps in Universal Health Coverage. *Lancet* 394, pp. 1192-1206. doi: 10.1016/S0140-6736(19)32136-1

Yip, W. and Hsiao, W. 2015. What drove the cycles of Chinese health system reforms? *Health Systems & Reform* 1(1), pp. 52-61. doi: 10.4161/23288604.2014.995005.

Yip, W. and Hsian, W. 2009. China's health care reform: a tentative assessment. *China Economic Review* 20, pp. 613-619. doi: 10.1016/j.chieco.2009.08.003

Yip, W., Hsiao, W., Chen W., Hu, S. and Ma, J. 2012. Early appraisal of China's huge and complex health-care reforms. *Lancet* 379, pp. 833–842.

Young, T. 2005. *Population health: concepts and methods.* 2nd ed. Oxford, England: Oxford University Press.

Yu, W. 2008. Constructing a harmony society: China's health care system reform. *Proceedings of the conference on China's Policy Reforms: Social Services, Regulation, and Finance.* Stanford's Center for International Development, Stanford October 22 – 24 2008.

Yu, W. and Ma, L. 2015. External government performance evaluation in China: evaluating the evaluations. *Public Performance & Management Review* 39(1), pp. 144-171. doi: 10.1080/15309576.2016.1071171.

Yu, Y., Zheng, X. and Zhang, L. 2016. Yardstick competition and the formation of enterprise zones in China. *Emerging Markets Finance and Trade* 53(9), pp. 1961-1972. doi: 10.1080/1540496x.2016.1184143.

Yuan, S. 2022. Zero COVID in China: what next? *Lancet* 399(10338), pp. 1856-1857. doi: 10.1016/S0140-6736(22)00873-X.

Yuesti, A., Novitasari, L.G. and Rustiarini, N.W. 2016. Accountability of non-government organization from the perspective of stakeholder theory. *International Journal of Accounting and Taxation* 4(2), pp. 98-119. doi: 10.15640/ijat.v4n2a7.

Zeng, J. 2014. The debate on regime legitimacy in China: bridging the wide gulf between Western and Chinese scholarship. *Journal of Contemporary China* 23(88), pp. 612–635. doi: 10.1080/10670564.2013.861141.

Zhang, H. 2012. The laws on the Ethnic Minority Autonomous Regions in China: legal norms and practices. *Loyola University Chicago International Law Review* 9(2), pp. 249-264.

Zhang, J. et al. 2019. Effects of FDI on the efficiency of government expenditure on environmental protection under fiscal decentralization: a spatial econometric analysis for China. *International Journal of Environmental Research and Public Health* 16(14), pp. 1–19. doi: 10.3390/ijerph16142496.

Zhang, L., Lu, K. and Chen, M. 2017. The effect of local fiscal expenditures on population migration. *Second International Conference on Information Technology and Management Engineering*. Beijing, 15-16 January 2017. Lancaster, Pennsylvania: DEStech Publications, Inc.

Zhang, Y. 2020. "Promotion tournament 2.0": why local cadres expand health-care provision in China. *Governance* 33(4), pp. 897-914. doi: 10.1111/gove.12429.

Zhao, D. 2009. The Mandate of Heaven and performance legitimation in historical and contemporary China. *American Behavioral Scientist* 53(3), pp. 416-433. doi: 10.1177/0002764209338800.

Zhao, L., Liu, S. and Zhang, W. 2018. New trends in internal migration in China: profiles of the new-generation migrants. *China and World Economy* 26, pp. 18–41. doi: 10.1111/cwe.12227

Zhao, X. and Li, X. 2021. The influence of internal migration on regional innovation in China. *Economic Research-Ekonomska Istraživanja* 34(1), pp. 498-520. doi: 10.1080/1331677x.2020.1792325.

Zheng, X., Wang, J., Li, X. and Yu, Y. 2015. On the supply of China's healthcare resources in a decentralized healthcare system. *The Social Science Journal* 52(4), pp. 449–58. doi: 10.1016/j.soscij.2014.05.003

Zheng, Y., Wang, W., Liu, W. and Mingers, J. 2019. A performance management framework for the public sector: the balanced stakeholder model. *Journal of the Operational Research Society* 70(4), pp. 568-580. doi: 10.1080/01605682.2018.1448247.

Zheng, Y.N. 1999. *Discovering Nationalism in China: Modernization, Identity and International Relations.* Cambridge: Cambridge University Press

Zheng, Y.N. 2006. Explaining the sources of de facto Federalism in reform China: intergovernmental decentralization, globalization, and central-local relations. *Japanese Journal of Political Science* 7(2), pp. 101–126. doi: 10.1017/S1468109906002222.

Zheng, Y.N. 2007. *De facto federalism in China reforms and dynamics of central-local relations*. Hackensack: World Scientific.

Zhou, L.A. 2010. *Incentives and governance: China's local governments.* Singapore: Cengage Learning Asia Pte. Ltd.

Zhou, L.L., Xu, X., Antwi, H.A., and Wang, L. 2017. Towards an equitable healthcare in China: evaluating the productive efficiency of community health centers in Jiangsu. *International Journal for Equity in Health* 16(89). doi: 10.1186/s12939-017-0586-y

Zhou, L., Ding, D. and Sun, Q. 2022. Under the epidemic, how to make the road to hemodialysis still "accessible" (in Chinese). Xinhua Daily Telegraph 28 April. Available at: http://www.news.cn/mrdx/2022-04/28/c_1310578597.htm [Accessed: 04 December 2022].

Zhou, X. et al. 2021. The party–government relationship in the Chinese bureaucracy: evidence from patterns of personnel flow. *Chinese Journal of Sociology* 7(3), pp. 315–346. doi: 10.1177/2057150X211031055.

Zhou, Z. 2009. Study of government performance management in China: a historical review and critical assessment. *Journal of Public Administration* 2(1), pp. 34-57.

Zhu, J. 2014. DEAFrontier software. In: Zhu, J. ed. *Quantitative Models for Performance Evaluation and Benchmarking*. New York: Springer New York, pp. 399-407.

Zhu, X. 2017. Inter-regional diffusion of policy innovation in China: a comparative case study. *Asian Journal of Political Science* 25(3), pp. 266-286. doi: 10.1080/02185377.2017.1339619.

Zhu, Y. 2011. "Performance legitimacy" and China's political adaptation strategy. *Journal of Chinese Political Science* 16(2), pp. 123-140. doi: 10.1007/s11366-011-9140-8.

Zhu, Y., Li, Y., Wu, M. and Fu, H. 2022. How do Chinese people perceive their healthcare system? Trends and determinants of public satisfaction and perceived fairness, 2006–2019. *BMC Health Services Research* 22(22). doi: 10.1186/s12913-021-07413-0

APPENDICES

Appendix 1. Summary statistics for HED by province (2006 to 2017)

Provinces	Max	Min	Mean	Median	SD
Anhui	0.971	0.829	0.929	0.954	0.046
Fujian	0.939	0.785	0.888	0.889	0.042
Gansu	0.938	0.805	0.891	0.909	0.042
Guangdong	0.971	0.869	0.934	0.937	0.030
Guangxi	0.921	0.750	0.877	0.904	0.052
Guizhou	0.944	0.793	0.892	0.922	0.057
Hainan	0.829	0.708	0.783	0.795	0.039
Hebei	0.965	0.774	0.914	0.933	0.052
Heilongjiang	0.914	0.824	0.880	0.892	0.031
Henan	0.956	0.814	0.913	0.928	0.043
Hubei	0.982	0.794	0.912	0.916	0.060
Hunan	0.965	0.870	0.929	0.939	0.031
Inner Mongolia	0.946	0.779	0.894	0.912	0.048
Jiangsu	0.948	0.863	0.912	0.917	0.026
Jiangxi	0.968	0.896	0.948	0.955	0.020
Jilin	0.903	0.753	0.862	0.881	0.047
Liaoning	0.935	0.842	0.912	0.922	0.027
Ningxia	0.861	0.719	0.805	0.829	0.051
Qinghai	0.814	0.603	0.726	0.743	0.068
Shaanxi	0.958	0.712	0.865	0.881	0.067
Shandong	0.964	0.855	0.936	0.954	0.031
Shanxi	0.940	0.811	0.881	0.874	0.050
Sichuan	0.967	0.859	0.936	0.947	0.033
Xinjiang	0.940	0.812	0.884	0.879	0.039
Yunnan	0.967	0.785	0.887	0.907	0.053
Zhejiang	0.929	0.814	0.880	0.886	0.038

Appendix 2. Summary statistics for TED by province (2006 to 2017)

Provinces	Max	Min	Mean	Median	SD
Anhui	0.883	0.764	0.827	0.835	0.043
Fujian	0.897	0.795	0.845	0.835	0.029
Gansu	0.820	0.679	0.766	0.775	0.038
Guangdong	0.933	0.859	0.894	0.890	0.019
Guangxi	0.835	0.734	0.789	0.796	0.034
Guizhou	0.825	0.704	0.764	0.771	0.035
Hainan	0.752	0.636	0.710	0.713	0.033
Hebei	0.866	0.518	0.804	0.834	0.090
Heilongjiang	0.789	0.667	0.739	0.753	0.042
Henan	0.882	0.812	0.845	0.841	0.026
Hubei	0.889	0.716	0.837	0.871	0.065
Hunan	0.912	0.768	0.836	0.816	0.050
Inner Mongolia	0.860	0.800	0.835	0.837	0.016
Jiangsu	0.912	0.831	0.873	0.876	0.025
Jiangxi	0.875	0.807	0.843	0.842	0.023
Jilin	0.826	0.717	0.762	0.767	0.031
Liaoning	0.876	0.808	0.848	0.845	0.020
Ningxia	0.763	0.579	0.701	0.722	0.058
Qinghai	0.690	0.530	0.625	0.630	0.054
Shaanxi	0.818	0.582	0.725	0.748	0.073
Shandong	0.904	0.848	0.881	0.878	0.018
Shanxi	0.804	0.711	0.757	0.746	0.031
Sichuan	0.876	0.828	0.851	0.855	0.016
Xinjiang	0.785	0.558	0.695	0.712	0.062
Yunnan	0.835	0.736	0.792	0.800	0.027
Zhejiang	0.935	0.876	0.896	0.893	0.018

Appendix 3. Summary statistics for RD by province (2006 to 2017)

Provinces	Max	Min	Mean	Median	SD
Anhui	0.906	0.802	0.866	0.883	0.037
Fujian	0.915	0.888	0.902	0.902	0.008
Gansu	0.722	0.552	0.661	0.659	0.048
Guangdong	0.809	0.752	0.780	0.779	0.015
Guangxi	0.891	0.734	0.769	0.761	0.042
Guizhou	0.825	0.753	0.796	0.807	0.027
Hainan	0.674	0.643	0.660	0.658	0.009
Hebei	0.820	0.619	0.769	0.784	0.054
Heilongjiang	0.783	0.597	0.733	0.753	0.057
Henan	0.952	0.912	0.934	0.939	0.014
Hubei	0.948	0.724	0.859	0.902	0.089
Hunan	0.849	0.796	0.823	0.823	0.020
Inner Mongolia	0.851	0.714	0.797	0.798	0.036
Jiangsu	0.938	0.890	0.915	0.915	0.015
Jiangxi	0.938	0.829	0.894	0.909	0.038
Jilin	0.785	0.728	0.753	0.748	0.019
Liaoning	0.974	0.799	0.917	0.966	0.073
Ningxia	0.752	0.656	0.727	0.735	0.027
Qinghai	0.743	0.661	0.695	0.686	0.026
Shaanxi	0.755	0.569	0.689	0.678	0.047
Shandong	0.969	0.861	0.913	0.889	0.041
Shanxi	0.743	0.648	0.707	0.732	0.035
Sichuan	0.790	0.690	0.733	0.731	0.029
Xinjiang	0.893	0.804	0.854	0.862	0.031
Yunnan	0.819	0.771	0.789	0.783	0.014
Zhejiang	0.944	0.889	0.915	0.920	0.018

Appendix 4. Summary statistics for BCC-DEA efficiency scores by province (2006 to 2017)

Provinces	Max	Min	Mean	Median	SD
Anhui	0.928	0.647	0.802	0.828	0.107
Fujian	1.000	0.942	0.979	0.976	0.019
Gansu	0.980	0.733	0.848	0.850	0.080
Guangdong	1.000	0.978	0.995	1.000	0.008
Guangxi	1.000	0.977	0.997	1.000	0.007
Guizhou	1.000	0.825	0.911	0.903	0.063
Hainan	0.988	0.883	0.940	0.941	0.030
Hebei	1.000	0.767	0.914	0.933	0.083
Heilongjiang	0.911	0.497	0.736	0.766	0.140
Henan	1.000	0.691	0.901	0.941	0.106
Hubei	1.000	0.708	0.873	0.910	0.116
Hunan	1.000	0.747	0.896	0.931	0.095
Inner Mongolia	0.754	0.578	0.675	0.692	0.068
Jiangsu	0.890	0.657	0.797	0.800	0.070
Jiangxi	0.952	0.890	0.914	0.913	0.017
Jilin	0.904	0.546	0.744	0.784	0.135
Liaoning	0.960	0.545	0.788	0.832	0.159
Ningxia	1.000	0.989	0.998	1.000	0.004
Qinghai	1.000	0.822	0.956	0.984	0.055
Shaanxi	0.898	0.556	0.771	0.831	0.133
Shandong	1.000	0.758	0.857	0.845	0.081
Shanxi	1.000	0.416	0.613	0.626	0.158
Sichuan	0.938	0.704	0.827	0.854	0.090
Xinjiang	1.000	0.711	0.910	0.964	0.102
Yunnan	1.000	0.885	0.967	0.983	0.040
Zhejiang	1.000	0.873	0.953	0.968	0.044

Appendix 5. Summary statistics for Super-efficiency DEA scores by province (2006 to 2017)

Provinces	Max	Min	Mean	Median	SD
Anhui	0.632	0.144	0.479	0.503	0.136
Fujian	1.019	0.742	0.865	0.840	0.092
Gansu	0.836	0.334	0.557	0.529	0.142
Guangdong	1.056	0.914	0.989	1.008	0.044
Guangxi	1.019	0.771	0.983	1.005	0.066
Guizhou	1.003	0.477	0.639	0.554	0.184
Hainan	0.721	0.173	0.635	0.677	0.145
Hebei	1.003	0.175	0.652	0.620	0.239
Heilongjiang	0.526	0.051	0.376	0.411	0.123
Henan	1.012	0.178	0.627	0.622	0.242
Hubei	1.016	0.313	0.669	0.668	0.227
Hunan	1.004	0.327	0.625	0.601	0.218
Inner Mongolia	0.398	0.070	0.303	0.326	0.077
Jiangsu	0.622	0.049	0.533	0.597	0.152
Jiangxi	0.652	0.488	0.585	0.578	0.044
Jilin	0.535	0.092	0.374	0.436	0.138
Liaoning	0.678	0.267	0.463	0.436	0.144
Ningxia	1.068	0.910	0.991	1.004	0.047
Qinghai	1.033	0.502	0.764	0.815	0.197
Shaanxi	0.495	0.053	0.399	0.463	0.127
Shandong	1.008	0.274	0.549	0.512	0.171
Shanxi	1.159	0.039	0.385	0.375	0.259
Sichuan	0.543	0.310	0.464	0.489	0.068
Xinjiang	1.009	0.173	0.729	0.846	0.276
Yunnan	1.017	0.625	0.850	0.844	0.139
Zhejiang	1.007	0.137	0.703	0.632	0.249

Appendix 6. Summary statistics for PSR by province (2006 to 2017)

Provinces	Max	Min	Mean	Median	SD
Anhui	99.621	99.125	99.42083	99.474	0.16
Fujian	99.567	98.993	99.35783	99.385	0.168
Gansu	99.266	98.245	98.94942	99.097	0.338
Guangdong	99.56	99.135	99.39675	99.41	0.13
Guangxi	99.364	98.871	99.149	99.183	0.147
Guizhou	99.514	98.644	99.20633	99.459	0.351
Hainan	99.603	99.025	99.33908	99.371	0.177
Hebei	99.676	99.127	99.4095	99.447	0.184
Heilongjiang	99.448	98.949	99.22483	99.241	0.134
Henan	99.632	99.086	99.466	99.589	0.19
Hubei	99.587	99.260	99.461	99.495	0.1
Hunan	99.563	99.052	99.346	99.376	0.163
Inner Mongolia	99.459	98.670	99.180	99.249	0.249
Jiangsu	99.689	99.371	99.574	99.616	0.101
Jiangxi	99.719	99.209	99.547	99.594	0.159
Jilin	99.407	99.015	99.158	99.117	0.133
Liaoning	99.392	98.810	99.115	99.148	0.197
Ningxia	99.187	98.450	98.896	98.87	0.232
Qinghai	99.304	98.369	98.997	99.085	0.314
Shaanxi	99.629	98.851	99.348	99.411	0.236
Shandong	99.608	99.256	99.450	99.519	0.119
Shanxi	99.375	98.916	99.147	99.163	0.144
Sichuan	99.647	98.998	99.407	99.449	0.192
Xinjiang	98.706	98.023	98.411	98.446	0.21
Yunnan	99.393	98.659	99.073	99.094	0.23
Zhejiang	99.626	99.285	99.465	99.475	0.117

Appendix 7. VIF values from efficiency and effectiveness models controlling for TED and RD

Appendix 7.1 VIF values from efficiency models (DEA) controlling for TED and RD

Dep: DEA	VIF	Dep: DEA	VIF
TED	3.81	RD	2.37
Lngdppc	9.16	Lngdppc	9.00
Lnpop	4.76	Lnpop	4.02
Popden	4.33	Popden	4.43
Urban	4.81	Urban	4.83
Edu	4.11	Edu	4.16
Elderly	2.14	Elderly	2.12
Unemp	1.74	Unemp	1.82
Lnhexppc	4.66	Lnhexppc	4.89
Solvency	4.55	Solvency	5.47
Fragmentation	3.11	Fragmentation	3.10
Minority	1.47	Minority	1.50
Mean VIF	4.05	Mean VIF	3.98

Appendix 7.2 VIF values from effectiveness models (PSR) controlling for TED and RD

Dep: PSR	VIF	Dep: PSR	VIF
TED	3.95	RD	2.30
Lngdppc	9.13	Lngdppc	8.88
Lnpop	5.95	Lnpop	4.64
Popden	4.82	Popden	4.86
Urban	6.06	Urban	6.06
Edu	5.29	Edu	5.35
Birthrate	2.32	Birthrate	2.32
Hinstitutions	1.80	Hinstitutions	1.8
Beds	5.39	Beds	5.28
lnhexppc	6.31	lnhexppc	6.06
Solvency	4.87	Solvency	5.62
Fragmentation	3.55	Fragmentation	3.51
Minority	1.54	Minority	1.56
Mean VIF	4.69	Mean VIF	4.48

Appendix 8. Non-linear effects of FD on healthcare efficiency and effectiveness

Appendix 8.1 Non-linear effects of FD on healthcare efficiency

	BCC-DEA Fixed effects			BCC-DEA Tobit regression			Super-efficiency DEA Fixed effects		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
FD^2	0.261**	1.753	0.057	0.253	2.564***	-0.141	0.334	3.045	0.112
	(0.122)	(1.520)	(0.673)	(0.156)	(0.732)	(0.612)	(0.197)	(2.150)	(0.621)
FD (HED)	0.052			0.115			-0.061		
	(0.238)			(0.312)			(0.332)		
FD (TED)		-2.210			-3.342***			-4.162	
		(2.318)			(1.079)			(3.234)	
FD (RD)			0.477			0.790			0.426
			(1.071)			(0.963)			(1.011)
Lngdppc	-0.111	-0.073	-0.029	-0.018	0.011	0.046	-0.104	-0.053	-0.020
	(0.092)	(0.091)	(0.087)	(0.060)	(0.059)	(0.058)	(0.105)	(0.105)	(0.099)
Lnpop	-1.403	-1.225	-0.953	-0.090*	-0.092*	-0.082	-1.839	-1.418	-1.353
	(0.866)	(0.812)	(0.819)	(0.050)	(0.047)	(0.051)	(1.175)	(1.040)	(1.137)
Popden	-0.001	-0.001	-0.001	0.000	-0.000	-0.000	-0.000	-0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)	(0.003)	(0.002)	(0.002)
Urban	-0.000	-0.002	-0.001	-0.001	-0.003	-0.001	0.001	-0.001	0.000
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.008)
Edu	0.005	0.005	0.006	0.009***	0.008**	0.009***	0.005	0.005	0.006
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)
Elderly	-0.013	-0.011	-0.010	0.007	0.004	0.005	-0.028*	-0.025*	-0.024*
	(0.009)	(0.009)	(0.009)	(0.006)	(0.006)	(0.006)	(0.014)	(0.015)	(0.014)
Unemp	0.009	0.000	0.012	-0.012	-0.022	-0.016	0.012	0.002	0.016
	(0.020)	(0.020)	(0.020)	(0.017)	(0.017)	(0.017)	(0.025)	(0.026)	(0.025)
Lnhexppc	0.058	0.070	0.035	0.078***	0.084***	0.054**	0.036	0.055	0.013
	(0.042)	(0.044)	(0.037)	(0.027)	(0.027)	(0.028)	(0.043)	(0.048)	(0.037)
Solvency	0.003***	0.003*	0.001	0.002***	0.002**	0.001	0.003***	0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Fragmentation	-3.128	-3.556	-2.868	-0.015	-0.706	-0.652	-2.346	-2.787	-2.059
	(4.950)	(4.824)	(5.016)	(1.200)	(1.162)	(1.218)	(7.236)	(6.734)	(7.314)
Minority regions				0.017	0.013	-0.012			
				(0.066)	(0.062)	(0.069)			
Time	0.029**	0.024*	0.024**				0.040**	0.032*	0.035**
	(0.011)	(0.013)	(0.011)				(0.015)	(0.017)	(0.015)
Constant	25.916*	23.483	17.421	1.616*	2.912***	0.917	33.378	27.465	24.210
	(14.938)	(13.987)	(13.995)	(0.924)	(0.981)	(0.981)	(20.252)	(17.998)	(19.453)
R-sq within	0.70	0.70	0.71				0.68	0.69	0.69
Pseudo R-sq				-1.19	-1.22	-1.22			
F-test	47.19	43.56	38.22				44.45	35.43	60.60
Prob > F	0.000	0.000	0.000				0.000	0.000	0.000
Wald chi2				527.66	539.59	551.56			
Prob > chi2				0.000	0.000	0.000			
LR test				203.46	197.31	205.21			
Prob≥chi2				0.000	0.000	0.000			

N = 312. Standard errors in parentheses: *** p < 0.01, ** p < 0.05, * p < 0.1. For Tobit regressions: uncensored observations: 276, left-censored: 0, righ-censored: 36

Appendix 8.2 Non-linear effects of FD on healthcare effectiveness (PSR)

Dep: PSR	(1)	(2)	(3)	(4)	(5)	(6)
		Fixed effects			Random effects	
FD^2	-0.139	-2.685***	-1.526	-0.212	-3.119***	-1.903**
	(0.155)	(0.921)	(1.402)	(0.185)	(0.890)	(0.744)
FD (HED)	0.288			0.513		
	(0.393)			(0.368)		
FD (TED)		3.485**			4.261***	
		(1.371)			(1.307)	
FD (RD)			1.983			2.621**
			(2.122)			(1.168)
Lngdppc	0.214***	0.251***	0.215***	0.234***	0.265***	0.242***
	(0.057)	(0.061)	(0.054)	(0.074)	(0.073)	(0.070)
Lnpop	0.409	0.111	0.009	-0.017	0.022	0.000
	(0.754)	(0.678)	(0.786)	(0.068)	(0.058)	(0.067)
Popden	-0.003**	-0.003**	-0.003*	0.000	0.000**	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Urban	0.005	0.007	0.006	-0.002	-0.001	-0.001
	(0.006)	(0.006)	(0.006)	(0.003)	(0.003)	(0.003)
Edu	0.007*	0.008*	0.007*	0.005	0.007*	0.006*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Birthrate	-0.031***	-0.026***	0.003	-0.024***	-0.013**	0.034**
	(0.010)	(0.007)	(0.031)	(0.006)	(0.006)	(0.016)
Hinstitutions	2.598***	2.413***	2.688***	2.081***	1.671***	2.067***
	(0.742)	(0.658)	(0.696)	(0.516)	(0.505)	(0.508)
Beds	0.011	0.008	-0.027***	0.037**	0.035**	-0.019***
	(0.035)	(0.030)	(0.008)	(0.016)	(0.016)	(0.006)
lnhexppc	0.032	0.015	0.044	0.049	0.047	0.062*
	(0.040)	(0.040)	(0.036)	(0.032)	(0.032)	(0.032)
Solvency	-0.079	-0.125	0.049	-0.103	-0.125	0.025
	(0.115)	(0.138)	(0.114)	(0.099)	(0.109)	(0.105)
Fragmentation	-1.849	-1.679	-2.694	-4.526***	-4.173***	-4.679***
	(2.309)	(1.744)	(2.432)	(1.548)	(1.381)	(1.522)
Minority				-0.271***	-0.277***	-0.278***
				(0.094)	(0.077)	(0.094)
time	0.009	0.011	0.011			
	(0.016)	(0.014)	(0.016)			
Constant	90.294***	93.990***	96.450***	96.746**	94.478***	95.564***
	(12.973)	(11.520)	(13.204)	(1.314)	(1.277)	(1.347)
F-test	132.93	92.87	94.34			
Wald chi2 test				1514.06	1501.60	1576.39
Prob	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-sq within	0.866	0.874	0.871			
R-sq overall				0.594	0.619	0.597

N = 312. Standard errors in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.