Jiang *et al. Health Economics Review* (2025) 15:16
https://doi.org/10.1186/s13561-025-00600-3

Health Economics Review

RESEARCH

Open Access

Impact of pilot public hospital reform on efficiencies: a DEA analysis of county hospitals in East China, 2009–2015

Wei Jiang[1], Xuyan Lou[2], Qiulin Chen[3], Lina Song[4] and Zhuo Chen[2,5*]

Abstract

Background China started a pilot public hospital reform in 2012 to improve governance and efficiency in healthcare services delivery among county-level hospitals. This study aims to investigate the impact of the pilot reform on hospital efficiency and productivity by using a unique dataset of county hospitals in East China during 2009–2015.

Methods A three-stage approach is used. First, this study uses the output-oriented data envelopment analysis (DEA) to estimate hospital efficiency with variable returns to scale. Second, propensity score matching is used to address potential biases associated with the selection of counties for the pilot program. In the third stage, we assess the impact of the pilot reform on efficiency by using a Tobit Difference-in-Differences approach.

Results The average level of hospital efficiency for the whole sample experienced a rapid drop in 2013, then returned to a peak in 2014. Except in the reform year (2012), the overall hospital efficiency for the post-reform period is higher than that for the pre-reform period. The baseline model results show that the pilot reform is associated with a 3% decline in pure technical efficiency and a 2.3% increase in hospital scale efficiency, respectively. Our findings are robust when we apply bootstrapped DEA efficiency scores and use different specifications.

Conclusion The findings of this study suggest no improvements in overall hospital efficiency associated with the pilot reform, possibly due to the combined effects of inefficient governance and hospital scale expansion. This study suggests that further efforts are needed to increase county hospital performance by strengthening management and optimizing resource utilization.

Keywords Public hospital reform, Hospital efficiency, Data envelopment analysis, Propensity score matching, China

*Correspondence:
Zhuo Chen
Zhuo.Chen@nottingham.edu.cn; zchen1@uga.edu
[1] College of Accounting, Ningbo University of Finance & Economics, Ningbo, China
[2] School of Economics, Faculty of Humanities and Social Sciences, University of Nottingham Ningbo China, Ningbo, China
[3] Faculty of Applied Economics, Universityof Chinese Academy of Social Sciences, Beijing, China
[4] Department of Industrial Economics, Nottingham University Business School, University of Nottingham, Nottingham, UK
[5] Department of Health Policy and Management, College of Public Health, University of Georgia, Athens, USA

Introduction

For developing and developed countries alike, assessing health system performance is critical to implementing appropriate strategies to control costs and efficiently use scarce resources [1, 2]. China's health system has undergone tremendous transformation over the last couple of decades [3, 4]. The 2009 national health reform aimed to increase access to care, expand insurance coverage, and strengthen primary care capacity. However, the early accomplishments of the reform were accompanied by soaring health expenditures [4, 5]. The public hospital reform in 2012 attempted to rein in the increases in healthcare expenditure and improve the performance of China's health system through the zero-markup drug policy and hospital governance. Previous studies show that the health reform has promoted health technology adoption and reduced inpatient spending but led to overtreatment in some settings [6–8]. However, few empirical studies investigate the policy effect on hospital performance and operational management, i.e., how public hospital utilizes limited healthcare resources, an important policy objective of China's health reform.

This paper focuses on the change in hospital performance from the perspective of efficiency measures. Early research used efficiency as a comprehensive and encompassing indicator for evaluating the performance of health systems. Efficiency was defined as the attainable level of a health system to what it could achieve [9]. The farther from the ideal health attainment, the lower the efficiency ratio. Under this context, the World Health Organization (WHO) analyzed the comparative efficiency of health system performance in 191 member states [10]. As for methods of efficiency measures, parametric and non-parametric methods are the main approaches to efficiency measurement in health care [11, 12]. Giuffrida and Gravelle used two parametric estimators, corrected ordinary least squares (COLS) and stochastic frontier approach (SFA), to evaluate the performance of 90 family health service authorities in the UK [13]. Hollingsworth and Wildman used Data Envelopment Analysis (DEA) for efficiency measurement with data from the WHO study [10, 14]. Hunt and Link conducted a bootstrapped DEA approach to measure the technical efficiency of local hospitals in the US [15]. In this study, we use DEA to measure the efficiency of county-level hospitals in East China.

There is an extensive research body using DEA to measure efficiency and analyze determinants of efficiency in China's health system. Jiang et al. analyzed the efficiency of a large sample of county hospitals in China from 2008 to 2012 [16]. They found decreasing hospital efficiency after the 2009 health reform and geographic disparities in technical efficiency between the eastern and western regions. Current studies adopted the Malmquist DEA efficiency index to analyze the efficiency change in provincial and county-level hospitals [17, 18]. Through the decomposition of productivity, the researchers attributed the improvement of hospital efficiency to technological progress after the 2009 health reform. Social health insurance and personal health expenditures may have unintended effects on local healthcare systems. A recent study reported an inverse relationship between personal health expenditure and the performance of provincial health systems [19].

Few studies explore the change in hospital efficiency when specific policies are implemented in the health system. Taking hospital privatization in Germany as an example, hospitals in Germany experienced a set of reforms over the last decades. Many public hospitals in Germany converted to private hospitals and started to implement the Diagnosis-related Group (DRG) financing scheme between 1997 and 2007. Tiemann and Schreyogg found that hospitals converted from public to private performed better in hospital operations and obtained higher quality of care [20]. Specifically, privatized hospitals were more likely to improve quality under the DRGs scheme [21]. Another study in Japan showed that hospital efficiency was not significantly affected by the change in regulation and improvement in financial conditions [22]. During hospital reform, local public hospitals in Japan recruited more healthcare staff while the number of beds was constrained, which caused a misallocation of medical resources on health services provision and quality of care. Besides, the regulatory costs imposed on the pharmaceutical and medical device industries slowed down the technological progress of healthcare system [22]. However, limited research detected the association between policies' impacts and hospital efficiency changes in China, especially concerning the public hospital reform in county hospitals.

Using county-level hospital data across six provinces and one metropolitan area in East China, this study investigates the impacts of pilot reform on hospital efficiency. Hospital efficiency scores are measured by using DEA. The result shows that all county hospitals experienced a short-term decline in hospital efficiency in 2013. We use propensity score matching to address the selection biases of counties associated with the inclusion in the pilot reform. Then, a difference-in-differences approach is applied to estimate the impact of the pilot reform on hospital efficiency. The result indicates that public hospital reform is associated with a reduction of 3% in pure technical efficiency scores. However, hospitals in the pilot counties are more efficient than those that are not. For robustness analysis, we use semi-parametric bootstrapping to address the potential bias in efficiency estimation. The estimates from different methods are robust in terms of magnitude.

Our research has three key contributions to existing literature. First, this study is the first to explore the impact of pilot public hospital reform on hospital efficiency. Few studies conducted policy evaluation studies on a large sample of county hospitals and assessed the reform's impacts on healthcare services delivery [16, 17]. Our findings reveal that the difference in efficiency between the pilot and non-pilot hospitals has unexpectedly narrowed after the pilot public hospital reform. Second, this research extends current studies on the potential mechanism behind the efficiency changes by decomposing efficiency. The insignificance of the pilot reform is driven by the mixed impact of the decline in pure technical efficiency and hospital scale expansion. Third, this study fills the gap in policy heterogeneity in terms of hospital types. Our findings provide evidence that the pilot reform has a limited effect on the efficiency decline of traditional Chinese medicine hospitals.

Literature review

Healthcare system and county hospitals in China

In expanding healthcare accessibility to the population, the main concerns are the heavy burden of healthcare expenditures and inefficient utilization of medical resources [23–25]. China's health reform has achieved universal health insurance coverage by the end of 2012. More than 90% of Chinese citizens are covered by one of the three major health insurance schemes[1], but with expanding healthcare costs over the past decades [26]. Government health expenditure reached 1.74 trillion Chinese Yuan in 2019, around three times that of 2009 (National Health Commission, 2020). The annual growth rate of total health expenditure was 14.98% from 2009 to 2019, which was 3.77% points higher than the annual growth rate of GDP. However, the health system in rural area still faces challenges from limited resource allocation, the lack of healthcare workforce, and diagnostic inaccuracies [26–28]. The dramatic growth of public spending and severe urban-rural disparity in healthcare utilization alarmed policymakers, who prioritized controlling the growth of healthcare costs and improving the capacity building of healthcare systems.

In China, public hospitals follow a three-tier healthcare services delivery system: tertiary hospitals, secondary hospitals, and primary healthcare facilities. Most tertiary and secondary hospitals are located in metropolitan areas and prefecture cities with adequate or better resources. As part of the public healthcare system, county-level hospitals delivered services to 900 million people, covering 70% of county residents' medical service needs [29]. However, most health resources were concentrated among tertiary hospitals, while resources available to the secondary and primary care facilities were around 40% [4]. The imbalanced resource distribution intensified the inequality of healthcare services and quality of care provided by county-level hospitals. Outpatient visits in tertiary hospitals experienced an 8% annual increase from 2009 to 2015 [27]. In contrast, a research based on the China Family Panel Studies survey pointed out that less than 60% of rural residents selected county hospitals or primary care facilities for health services [30]. To achieve a better quality of care and reduce inequality, public county hospitals need to improve operational efficiency through resource optimization and professional management. Hence, our study focuses on the change in hospital performance, and try to investigate the impact of pilot reform on county-level hospital efficiency.

The pilot public hospitals reform

China launched a systematic health reform in 2009 with the overarching goal of expanding insurance coverage and ensuring accessible and affordable services for the population. The reform has three planned phases [25]. The first phase was from 2009 to 2012. The goals of the health reform in this period were to meet the priorities of infrastructure building and strengthen the capacity of primary healthcare facilities. The second phase of the health reform was from 2012 to 2015. Recognizing the ballooning healthcare costs and inequalities in the delivery system, China has focused on improving the governance of public hospitals. The last phase of health reform is from 2017 to 2020, aiming to establish an integrated health system for both medical and public health services.

As an important component of the health reform, a multi-stage pilot reform in county-level hospitals was implemented from 2012 to 2015. The main elements of the reform included strategic planning, financing policy reform, department reorganization, and performance assessment. In June 2012, China's central government kickstarted the pilot reform in 311 counties. The pilot reform is planned to cover the entire counties in mainland China at the end of 2015. To keep most medical services provision within the county, the county governments integrated resources to operate at least one general or traditional Chinese medicine hospital for residents. The pilot design of county hospital reform consists of three components: first, the zero-markup drug policy (ZMDP) to remove the 15% markup on drug prices sold at public hospitals; second, restructuring hospital governance and managerial compensation; third, promote cooperation and partnership between major hospitals and community health centers, i.e., tertiary hospitals support county hospitals through medical training and paired assistance in labor resources. According to the National Health Commission guidelines[2], hospitals in pilot counties will receive government subsidies and have

[1] Three health insurance schemes: Urban Employee Basic Medical Insurance (UEBMI). Urban Resident Basic Medical Insurance (URBMI). New Cooperative Medical Scheme (NCMS).

[2] Weblink: http://www.mof.gov.cn/gp/xxgkml/shbzs/201211/t20121125_2500611.htm.

more autonomy on hospital expansion, human resources, and payment incentives.

Methods

Data sources

Our study combines three separate datasets. The first dataset is the National Health Statistical Information Report System, which provides the financial statements and information on facilities and staff of hospitals from 2009 to 2015. We extract hospital data from this dataset to measure the efficiency of county hospitals. This study only includes general and traditional Chinese medicine hospitals to ensure the comparability of healthcare services delivery. Mental health hospitals, long-term care facilities, and dental clinics are excluded. We also reviewed the data with specific hospital grades and fiscal budgets to ensure the hospital received financial support and administration from the corresponding county government.

The second dataset is the official documents issued by the National Health Commission[3], which contains information on public hospital reform and the list of pilot counties. Due to the data availability, we restrict our sample to 205 hospitals in the treated group from 93 pilot counties, which represents 29% of all pilot counties in China during the first batch county-level public hospital reform. As for the regional distribution, our study focuses on the healthcare system in Eastern China. County public hospitals from Shanghai, Anhui, Fujian, Jiangsu, Jiangxi, Shandong, and Zhejiang are included in our sample.

This study specifies two inputs and two outputs in the DEA process for measuring hospital efficiency scores. The inputs considered serve as proxies for delivering daily medical services to patients. Following previous studies, we use the number of full-time equivalency health care professionals (physicians, nurses, medical technicians, and other health professionals) and the number of hospital beds [20, 22, 31, 32]. In China, providing fundamental medical services to the residents is one of the key objectives of county hospitals in the new healthcare reform phase. Following previous studies, we use emergency visits and inpatient days as the output indicators [16, 21]. Outpatient and emergency visits represent the number of patients from outpatient and emergency visits per year. Inpatient days is the total number of occupied bed days for patients.

Aside from efficiency measurement, our study also uses the average length of stay and bed use rate as the indicator of healthcare utilization. Table 1 presents summary statistics for the hospital variables. Most county hospitals in this study are secondary hospitals (88%). As for hospital type, approximately 60% of hospitals are general hospitals, while the rest of them are traditional Chinese Medicine hospitals.

The main covariates used in PSM specification and DID analysis are demographical and socioeconomic information collected from the third dataset, the China Statistical Yearbook (County-level) and the National Economic and Social Development Statistical Bulletin (County-level). Our study matches individual hospitals with the socioeconomic information of the counties where the hospitals are located. Those variables include socioeconomic status indexes, demographic indicators, governmental expenditure, and local healthcare infrastructure. We also provide a comparative summary of the counties' characteristics of pilot and non-pilot hospitals in Table 1. The average population of pilot and non-pilot counties is about 643,700 persons. Gross Domestic Product per capita (GDPPC) is calculated as the value of county-level GDP divided by the population of this county. The number is close to 40,000 RMB in our study.

The original dataset contains 749 county-level hospitals from 366 counties. Observations with missing values and outliers are excluded from the requirement of efficiency measures. We construct a 7-year balanced panel data to assess the policy impact of the first-batch public hospital reform before and after 2012, the year of the pilot reform. Hospitals are divided into pilot and non-pilot groups based on the list of the first-batch pilot counties in public hospital reform. After PSM matching, 190 pilot and 175 non-pilot hospitals from 172 counties (85 pilot and 87 non-pilot counties) are selected as the final sample.

We use a three-stage strategy to estimate the impact of the pilot reform on hospital efficiency. First, hospital yearly efficiency scores are analyzed using the output-oriented DEA approach. Second, we use propensity score matching to pair pilot counties with those counties with similar group characteristics. Third, the impact of the pilot reform is assessed by using a difference-in-differences (DID) specification of the Tobit regression model. Efficiency scores calculated in the first stage were used as the dependent variable in the DID analysis.

Data envelopment analysis

Data envelopment analysis is a classical non-parametric method for measuring the efficiency and productivity of a decision-making unit (DMU) [33]. Over the past decades, DEA has been widely applied to assess efficiency and performance in healthcare, education, operational research, and transport industries [34]. We use the DEA approach to measure county hospitals' efficiency for the following reasons. First, DEA approach has its advantages in handling multiple

[3] http://www.nhc.gov.cn/tigs/s3581/201307/582017786d09475491c405c4fec5ad91.shtml.

Table 1 Summary of Hospital and County characteristics (before Propensity score matching specification)

Variable (Observations = 5243)	Mean	S.D.	Median	Min	Max
Input and Output in Efficiency Measurement					
The number of health care professionals	434.48	332.27	340	10	2895
The number of beds	375.87	304.59	290	0	2600
The number of outpatient visits	275.06	268.21	195	0.33	2876.23
The number of inpatient days	125.47	112.49	90.37	0	948.87
Health care utilization					
Average length of stay	8.16	2.64	7.88	0	51.18
Bed use rate	0.88	0.23	0.93	0	2.91
Hospital characteristics					
Type (Dummy, general hospital or not) (%)	0.58				
Tertiary hospital (%)	0.05				
Secondary hospital (%)	0.88				
Primary healthcare facilities (%)	0.07				
County characteristics					
GDP	264.61	294.55	173.79	10.91	3080.02
Population	64.37	37.46	57.42	4.4	198.64
GDP per capita	3.93	2.57	3.31	0.45	18.57
Government revenue	18.8	23.01	11.04	0	284.76
Government expenditure	30.16	23.39	23.88	2.71	255.36
Services industry proportion (%)	0.34	0.07	0.34	0.12	0.61
The number of beds in healthcare institution	31.7	12.89	29.37	6.91	130.53
Share of Education Enrollment in Population (‱)	8.03	5.08	6.71	0.28	32.15

Notes: 1) Input and Output in Efficiency Measurement: The number of health care professionals stands for the proxy of full-time equivalents, including physicians, nurses, medical technicians, and other health care professionals. The number of Outpatient visits is the number of emergency and outpatient visits (1,000 people). The number of inpatient days is the total number of days during which patients receive medical services at the hospital.2) County characteristics include GDP per capita (0.1 billion RMB units); Population, regional permanent population (10,000 people); GDPPC, GDP per capita (10,000 RMB); Government revenue and expenditure (0.1 billion RMB); The output in the services industry as a share of GDP; The number of beds in healthcare institutions in population (10, 000); The share of Secondary school and primary school enrollment in regional permanent population (10, 000)

outputs compared with other parametric efficiency measures. Second, DEA is more flexible in model setting than other parametric methods (such as SFA or COLS) because no assumptions are needed on the functional form of the production frontier [35].

Through linear programming, DEA generates a ratio that reflects the distance of each DMU to the efficient frontier. The efficient frontier is based on the best estimated input-output relationship under a certain assumption and framework [34]. The DEA approach could be fitted with different frameworks in its production frontier. In the output-oriented DEA model, the efficient frontier is derived from those units that achieved maximum outputs under given inputs. The efficient frontier in the input-oriented model is the best practice for accomplishing the minimum inputs in production for given outputs [36]. In this study, we concentrate on the efficiency of county-level hospitals in China and assess the impact of public hospital reform on efficiency. We use the output-oriented DEA model for analysis because it fits the context that county hospitals exploited given inputs to get maximum outputs. On the one hand, limited medical resources are allocated to county hospitals in China. County Hospital leaders had no control over inputs (like human resources) in the short run [37–39]. On the other hand, the main target for county hospitals over the observation period was enhancing the capacity building of primary healthcare facilities and providing most medical services to rural residents [4].

Farrell is among the first to apply the basic DEA approach with one input and one output for the DMU [40]. Charnes et al. extended Farrell's work and proposed the DEA approach with a constant return to scale assumption, CRS Model [41]. Banker et al. extended it into a variable return to scale (VRS) model [42]. The main difference between CRS and VRS model is whether a unit increase (decrease) in the input will contribute to a proportional increase (decrease) in the output. The constant return to scale model is more appropriate when the DMUs operate at the optimal scale, reflecting the perfect competition in the market. When DMUs faces incomplete competition, the VRS Model is used to alter the efficiency scores from CRS model by considering the scale of each DMU. In this study, public hospitals in

China operate under budget constraints, regulations, and imperfect competition [43–45]. Thus, we adopt a variable return-to-scale model to calculate the efficiency scores of county-level hospitals [14]. The model is denoted as follows:

$$max\ \theta_0(w,\ \mu) = \frac{\sum_{i=1}^{r} \mu * y_r}{\sum_{i=1}^{k} w * X_k} \quad (1)$$

$$s.t.\ \sum_{i=1}^{r} \mu * Y_r - \sum_{i=1}^{k} w * X_k \leq 0,\ because\ \theta \leq 1$$

$$\sum_{i=1}^{N} w_N = 1$$

$$w_1,\ w_2,\ \ldots,\ w_k \geq 0$$

$$\mu_1,\ \mu_2,\ \ldots,\ \mu_r \geq 0$$

Efficiency score θ_0 is the estimated efficiency of DMU_0 for N decision DMU with K inputs and R outputs. Where $Y(r \times n)$ and $X(k \times n)$ are the matrices for hospitals with k inputs and r outputs, respectively. Hospitals maximized the efficiency ratio under the constraints of given inputs. w and μ are vectors for the weight of inputs and outputs. The weights of the input-output combination are obtained by dual programming. The efficiency scores calculated by DEA model denote the overall efficiency (hospital efficiency in this study), which measures how health resources are transformed into the production process of each county hospital in healthcare services delivery. Overall technical efficiency could be decomposed into pure technical efficiency and scale efficiency. Pure technical efficiency denotes the efficiency under a variable return of scale related to management and technology innovations. Scale efficiency is the gap between the actual and optimal scales, measuring whether a hospital operates at the optimal scale.

Propensity score matching

We aim to investigate the policy impact on hospital efficiency among county-level hospitals. However, the counties selected for the pilot reform are more likely to be distinct in terms of access to healthcare and economic resources. We use county-level covariates in a propensity score matching to address the potential selection bias to construct a counterfactual. We include county-level covariates because the pilot reform was implemented at the county level rather than for specific hospitals. Counties were enrolled in the pilot program, and hospitals in pilot counties were included in the pilot reform. This study used propensity score matching to balance the treatment and control groups and to match pilot counties with other counties that are similar to the pilot counties but not enrolled in the pilot program.

Propensity score matching results in a balanced comparison to create a "quasi-random experiment." This approach assigns treatment and control groups in terms of the conditional probability given by the observed covariates [46]. To adjust multivariate sampling to the control group, we use a logit regression model to estimate the propensity scores [47]. We set 2011 (one year before the pilot reform) as the base year. To find appropriate covariates that determine the pilot grouping, we select socioeconomics and health system covariates based on previous studies [20, 43]. The county-level covariates in the logit model include GDP per capita, government revenue, government expenditure, the number of beds in healthcare institutions, the output in the services industry as a share of GDP, and the share of education enrollment in the population. We minimize the distance of propensity scores of the nearest neighbor to reduce the disparity between pilot and non-pilot counties. We used one-to-one matching without replacement to match each county in the treatment group to at least one non-pilot county in the control group. After propensity score matching, counties under common support remain in the final sample. We examine standardized differences of the covariates to evaluate the balance of the two sub-samples. The result of the balanced test of covariates is shown in Table 8 in Appendix 1.

Difference-in-differences

This study uses a standard difference-in-differences specification to assess the impact of pilot reform on hospital efficiency. The dependent variables are the estimated scores of hospital efficiency, pure technical efficiency, and scale efficiency of hospitals by DEA measures. Our study applies the Tobit regression model because of two-side truncated efficiency scores computed by DEA as dependent variables. Specifically, we conduct a 7-year panel Tobit model with hospital individual fixed and year fixed effects as our main DID specification. The reasons for using two-way fixed effects DID model are as follows. First, adding individual fixed effects could deal with the unobserved factors and time-invariant characteristics of hospitals, affecting hospital efficiency. Second, we use year effects to control the common policy impacts on all hospitals after the nationwide healthcare reform launched in 2009. Aside from the pilot hospital reform, a set of healthcare policies and guidelines were released from 2009 to 2015. The year-fixed effects address the time-specific effects on all county hospitals. The DID model can be specified as follows:

$$\theta_{it} = \lambda\,(Post_{it} * Pilot_i) + \beta\, Z_{ct} + a_i + \gamma_t + \epsilon_{ict} \quad (2)$$

Where θ_{it} denotes the efficiency scores of hospital i in year t. λ is the coefficient of the interaction term, capturing the impacts of the pilot reform on county hospitals in different groups. $Post_{it}$ is the time dummy, which equals 1 in the years after pilot reform was implemented for hospital i in 2012, and 0 otherwise. As for the treated dummy, $Pilot_i$, equals 1 if the hospital i was in the pilot reform counties from 2009 to 2015 and zero otherwise. a_i and γ_t represent hospital individual fixed effects and year fixed effects, respectively. ϵ_{ict} denotes the error term.

We control the county-level characteristics that have potential confounding effects on hospital efficiency. Z_{ct} represents a set of control variables of socioeconomic status, health system, and demographic information on county c where hospital i located. Specifically, we included GDP per capita, local government revenue, local government expenditure, the proportion of services industry to GDP, the number of beds in healthcare institutions, and the share of education enrollment[4] in the regression model [17, 21, 24]. GDP per capita, government revenue, government expenditure, and the number of beds in healthcare institutions were log-transformed following prior studies and for easy interpretation.

This study does not use hospital-level characteristics as the control variables. The reasons are as follows. As mentioned in the data sector, our study only includes general and traditional Chinese medicine hospitals associated with local governmental financial support. Those inclusion criteria control the variation of hospital type and subsidy. Besides, the efficiency indicator in our study represents how county hospitals utilize medical resources to deliver healthcare services. Compared with the hospital characteristics, this indicator is more relevant to local demographic information and health infrastructure.

The key assumption in DID identification strategy is the change in dependent variables for the treatment and control group should share identical trends before the policy implementation. In this study, we use a more flexible form of DID specification to test the parallel trend assumption (PTA) and capture the dynamic effects of the pilot reform. The event study method is as follows:

$$\theta_{it} = \sum_{k=-3}^{3} \lambda_k * D_{ki} + \beta Z_{ct} + a_i + \gamma_t + \epsilon_{ict} \tag{3}$$

Where D_{ki} denotes the interactive term of the treatment dummy and period dummy. Our study contains 7-year observation periods, which could be divided into 3-year pre-intervention period, policy implementation time, and 3-year post-intervention period. K is the footprint of periods. We drop one year before the pilot reform ($k = -1$) to avoid multicollinearity. λ_k is the coefficient indicating dynamic policy effects.

Robustness check

This study transforms hospital efficiency scores from conventional DEA estimates to Bias-corrected estimates. Conventional DEA estimates would be biased because of the unknown efficient frontier and exogenous affecting efficiency measures. First, the efficiency score generated by DEA is the deviation from the estimated efficient frontier (a set of observed units) to each DMUs, rather than the distance of true full-efficiency units to other DMUs. However, the true possible full-efficiency frontier could not be illustrated due to the finite sample in DEA framework [44]. Second, the inefficiency of hospitals could not be entirely explained by resource misallocation or scale inefficiency. Hospital efficiency is also affected by external factors and organizational characteristics [20, 48]. Conventional DEA may overestimate the inefficiency of hospital governance. Simar and Wilson proposed a bias-corrected efficiency estimator based on the two-stage bootstrapped process to address the potential bias in efficiency measures [49]. The details of bootstrapped DEA procedure are in Appendix 2.

This study uses bias-corrected estimates in hospital technical efficiency measurement based on Simar and Wilson's approach. Specifically, we implemented 1000 times bootstraps to produce bias-corrected estimates and 500 bias-corrected bootstraps to construct confidence intervals of bootstrapped estimates [50]. The bias-corrected estimates are obtained using **simarwilson**, and implemented using Stata 18.0 (Stata Corp, College Station, TX).

Aside from the robustness check in replacing efficiency scores, we conduct a set of estimations of the policy impact. The original estimation of the panel data model with two-way fixed effects is Least Squares Dummy Variable (LSDV) or Within-Group estimation. However, those estimators would be consistent and unbiased under the assumption that the error terms are independent and asymptotically normally distributed. Due to the trimmed distribution of the error term in censored regression models (Tobit model), Least Squares Dummy Variable and other estimation methods will be inconsistent in this sense. Honore applied a trimmed least absolute deviations estimator to construct the moment conditions for one-sided censoring data with fixed effects [51]. By using this method, Honore maintained assumptions on consistency and asymptotically normal distribution. Alan et al. extended this method in two-sided censoring panel data regression [52]. In our study, we follow Alan et al.

[4] Education enrollment: Student enrollment in regular secondary and primary school per 10, 000 persons. This variable reflects the age structure for county-level population. The share of aging population in county is not applied due to data availability.

Table 2 Descriptive Statistics of Socioeconomics for Pilot and Non-pilot hospitals during 2009–2015 (after propensity score matching specification)

Variables	Pilot Hospitals (*N* = 190)		Non-Pilot Hospitals (*N* = 175)	
	Mean	S.D.	Mean	S.D.
GDP	356.0	246.1	268.4	344.2
Population	73.3	31.3	67.0	44.2
GDP per capita	4.9	2.7	3.9	2.8
Government revenue	26.7	20.4	18.7	28.0
Government expenditure	37.1	22.5	31.3	27.6
Services industry proportion (%)	35.0	7.8	35.5	6.9
The number of beds in healthcare institution	32.3	12.2	31.3	11.3
Share of Education Enrollment in Population (‱)	8.7	4.4	8.4	6.0

Notes: Pilot hospitals are defined as hospitals located in the counties where the first phase of pilot reform was implemented in 2012. Socioeconomic status variables include GDP per capita (0.1 billion RMB units); Population, regional permanent population (10, 000 people); GDPPC, GDP per capita (10, 000 RMB units); Government revenue and expenditure (0.1 billion RMB units); The output in the services industry as a share of GDP; The number of beds in healthcare institutions in population (10, 000); The share of Secondary school and primary school enrollment in regional permanent population (10, 000)

method and use **two_side** command to conduct the main estimation. In contrast, we also applied Tobit panel data model with LSDV, and Tobit model with random effects as the robustness check.

Results

Descriptive statistics

Table 8 in Appendix 1 reports the result of propensity score matching specification. Our study uses data from 2011 for the matching. Column (3) of Table 8 in Appendix 1 shows the difference in most county-level characteristics between pilot and non-pilot counties before matching. Only the proportion of the services industry in GDP demonstrates insignificant differences before matching. Column (6) of Table 8 in Appendix 1 shows the matching result, which suggests no significant difference in county characteristics between the treatment and control groups after a one-to-one matching. Therefore, the matching helped to reduce the bias associated with the selection into the pilot program.

Table 2 presents the descriptive statistics of control variables after matching. Those variables included socioeconomic status indexes, demographic indicators, and information on healthcare conditions. Hospitals located in pilot counties are usually with better economic infrastructure. The pilot counties' GDP per capita, government revenue, and government expenditure were higher than those of the non-pilot counties over the research period. There is no significant difference between pilot and non-pilot counties regarding average education.

The summary statistics of the inputs and outputs used in hospital efficiency measures are shown in Table 3. The average number of health care professionals and hospital beds in pilot hospitals is larger than in non-pilot hospitals. The input variables in DEA estimates experienced steady growth from 2009 to 2015 in both pilot and non-pilot hospitals. The number of health care professionals in pilot hospitals increased annually by 6.4%, from 390 in 2009 to 567 in 2015. At the same time, the number of healthcare professionals in non-pilot hospitals shows a 6.7% annual increase. As for the number of hospital beds, the difference between pilot and non-pilot hospitals has been narrowed from 38 in 2009 to 29 in 2015.

As for the outputs, the number of outpatient visits in pilot hospitals increased from 273.7 thousand visits per year in 2009 to 430.6 thousand visits per year in 2015, representing a 7.8% annual growth. Non-pilot hospitals demonstrate a similar trend, with a 7.9% annual growth rate. The average number of inpatient days in pilot hospitals increased slightly lower than that of non-pilot hospitals from 2009 to 2015. In conclusion, pilot hospitals and non-pilot hospitals experience similar increases in trends in outputs and inputs for efficiency measures. The disparity in healthcare services delivery between pilot and non-pilot hospitals widened after the pilot reform.

Hospital efficiency measurement

Table 4 illustrates the efficiency scores of both pilot and non-pilot hospitals during 2009–2015. We estimate hospital efficiency, pure technical efficiency, and scale efficiency by using DEA in an output-oriented setting. The average level of hospital efficiency for the whole sample experienced a rapid drop in 2013, then returned to the peak in 2014. Except in the reform year (2012), the overall hospital efficiency for the post-reform period is higher than that for the pre-reform period. In addition, the efficiency gap between pilot and non-pilot hospitals has narrowed from 0.083 in 2010 to 0.061 in 2015.

Table 3 Summary statistics of the inputs and outputs in Efficiency Measurement (after propensity score matching specification)

Variables	Health care professionals		Beds		Outpatient visits		Inpatient days	
	Pilot (*N* = 190)	Non-pilot (*N* = 175)	Pilot (*N* = 190)	Non-pilot (*N* = 175)	Pilot (*N* = 190)	Non-pilot (*N* = 175)	Pilot (*N* = 190)	Non-pilot (*N* = 175)
2009								
Mean	390	347	317	279	273.7	198.2	105.6	89.4
S.D.	(271)	(243)	(246)	(224)	(226)	(178.6)	(90.9)	(80.6)
2010								
Mean	421	370	346	301	294	209.4	121	97
S.D.	(304)	(269)	(279)	(246)	(233.9)	(191)	(109.3)	(88.8)
2011								
Mean	456	392	369	333	343.5	241.2	132.4	111.9
S.D.	(334)	(306)	(289)	(267)	(283.3)	(220.3)	(114.1)	(99.3)
2012								
Mean	485	429	435	377	379.4	273.4	150.6	125.3
S.D.	(343)	(331)	(357)	(303)	(308.9)	(241.8)	(132.7)	(109.5)
2013								
Mean	514	461	460	430	396.4	291.6	155.9	139.8
S.D.	(374)	(348)	(350)	(348)	(331.7)	(254.1)	(128.5)	(121.7)
2014								
Mean	536	480	477	456	424.9	316.1	162.3	147.9
S.D.	(392)	(353)	(346)	(358)	(352)	(280.1)	(125)	(124.4)
2015								
Mean	567	511	500	471	430.6	313.4	165	146.3
S.D.	(407)	(372)	(359)	(380)	(355.7)	(283.1)	(128)	(124.5)

Notes: Pilot hospitals are defined as hospitals located in the counties that were included in the list of the first phase of pilot reform in 2012. Outpatient visits is the number of emergency and outpatient visits (1, 000 people). The number of inpatient days is the total number of days during which patients received medical services at the hospital

Table 4 Efficiency of pilot and non-pilot hospitals in 2009–2015

Year	Hospital efficiency			Pure technical efficiency			Scale efficiency		
	Pilot (*N* = 190)	Non-pilot (*N* = 175)	Difference	Pilot (*N* = 190)	Non-pilot (*N* = 175)	Difference	Pilot (*N* = 190)	Non-pilot (*N* = 175)	Difference
2009	0.671	0.620	0.051	0.695	0.639	0.056	0.968	0.973	−0.005
2010	0.613	0.530	0.083	0.662	0.567	0.095	0.932	0.943	−0.011
2011	0.640	0.579	0.061	0.684	0.614	0.07	0.942	0.950	−0.008
2012	0.676	0.627	0.049	0.708	0.659	0.049	0.957	0.955	0.002
2013	0.455	0.411	0.044	0.556	0.506	0.05	0.844	0.845	−0.001
2014	0.742	0.690	0.052	0.758	0.710	0.048	0.980	0.976	0.004
2015	0.727	0.666	0.061	0.752	0.695	0.057	0.968	0.963	0.005

Notes: Hospital efficiency, the efficiency of the hospital in producing outputs for given inputs in healthcare services delivery. Pure technical efficiency, the efficiency of the hospital in producing outputs for given inputs in healthcare services delivery under a variable return of scale. Scale efficiency, the ratio of the efficiency under constant return of scale to the efficiency under variable return of scale, evaluates the gap between the actual scale to the optimal scale. Difference, the gap of efficiency between pilot and non-pilot hospitals

The change in pure technical efficiency shows the same trend as hospital efficiency. Both pilot and non-pilot hospitals have increased efficiency scores after the pilot reform. Pure technical efficiency reached to peak in 2014. The average pure technical efficiency of pilot hospitals and non-pilot are 0.758 and 0.710, respectively. The difference in the pure technical efficiency is slightly smaller than the hospital efficiency. Given these trends, the results suggest county hospitals have improved efficiency after pilot reform. The gap between pilot and non-pilot hospitals has been gradually narrowing since 2013.

Although most county hospitals are not fully scale efficient (scale efficiency = 1), the average scale efficiency achieves a high interval over the observation period. Hospitals have experienced a steady improvement in scale efficiency since 2009. The average scale efficiencies of pilot and non-pilot hospitals were 0.968 and 0.963 in 2015, respectively. Scale efficiency peaked in 2012 and slightly declined from 2013 to 2015. The inverse U-shaped curve of scale efficiency indicates that county hospitals fail to allocate resources efficiently when resources are beyond the optimal scale. Besides, non-pilot hospitals' scale efficiency is better than pilot hospitals, with a relatively small gap for the most time.

Baseline results

Table 5 shows the results of DID specification with two-way fixed effects after the propensity score matching process. We use hospital efficiency, pure technical efficiency, and scale efficiency as the dependent variables in the regression model. The interaction term between the pilot reform dummy and post-reform indicator measures the differential policy effects on hospitals. As shown in Column (1) of Table 5, the pilot reform has a negative effect on hospital efficiency when we control the year-fixed effects and hospital-fixed effects. In Column (2), the results indicate that pilot hospitals have a decrease of 3% ((p-value < 0.05) in pure technical efficiency from the pilot reform. Compared with overall hospital efficiency, the result of pure technical efficiency shares a similar magnitude but is highly significant. In contrast, the result in Column (3) suggests that the pilot reform leads to a significant 2.3% improvement in scale efficiency.

Table 5 Impact of pilot reform on hospital efficiency

Variables	(1) Hospital Efficiency	(2) Pure Technical Efficiency	(3) Scale Efficiency
Post*Pilot Reform	−0.019	−0.030**	0.023**
	(0.012)	(0.012)	(0.010)
Log GDP per capita	−0.050	−0.068	0.019
	(0.054)	(0.055)	(0.034)
Log Government Revenue	0.008	−0.003	0.007
	(0.019)	(0.018)	(0.021)
Log Government Expenditure	−0.050*	−0.037	−0.015
	(0.028)	(0.025)	(0.037)
Services Industry Proportion	0.232	0.300*	−0.061
	(0.156)	(0.161)	(0.140)
Log Healthcare Institution Bed	−0.034	−0.022	−0.040***
	(0.021)	(0.021)	(0.015)
Share of Education Enrollment	0.002	−0.000	0.004
	(0.004)	(0.004)	(0.003)
Year Effects	Yes	Yes	Yes
Individual Effects	Yes	Yes	Yes
N	2555	2555	2555

Note: The numbers in parentheses are standard errors. Significant levels are denoted as *p < 0.1, **p < 0.05, ***p < 0.01. Control variables include GDP per capita, government revenue per capita, government expenditure per capita, the output in the services industry as a share of GDP, the number of beds in healthcare institutions in population (10,000), and the share of education enrollment in the regional permanent population

Regarding socioeconomic factors, the GDP per capita of the county is negatively related to hospital efficiency and pure technical efficiency of the local county hospitals. In contrast, hospitals in better-off counties are more likely to be scale-efficient. We fail to find statistical evidence that government revenue affects any kind of efficiency. As for the industrial structure and health system infrastructure, we also find that hospitals in counties with a high proportion of services industry obtained higher hospital efficiency and pure technical efficiency. In contrast, the services industry proportion has a negative coefficient to scale efficiency. It is worth noting that the number of beds in local healthcare institutions is negatively related to scale efficiency. This negative impact on scale efficiency might be attributed to the substitution effect between local healthcare facilities and county hospitals.

Results of the robustness check

We apply a bias-corrected efficiency estimate as a robustness check for hospital efficiency. The new efficiency scores are generated using 1000 replications in Simar-wilson's bootstrapping procedure. The result of bias-corrected efficiency estimates is shown in Column (4) of Table 9 in Appendix 1. The pilot reform is negatively associated with bias-corrected efficiency. The coefficient of the bias-corrected efficiency score is identical to the original hospital efficiency score. The difference in significance level might be attributed to the different settings in standard error calculation.

The results of the event study are in Fig. 1 in Appendix 1. This study estimates the dynamic policy effects on hospital efficiency, pure technical efficiency, scale efficiency, and bias-corrected efficiency. The omitted period is one year before the implementation of pilot reform. We report both coefficients and 95% confidence intervals for policy impact by year. From the trends of the pre-stage, we fail to reject that the treatment group and the control group had the same trend of efficiency changes before the intervention of the pilot reform. Notably, pure technical efficiency significantly increased two years before the policy implementation. Hospital efficiency, pure technical efficiency, and bias-corrected efficiency have identical time trends over the whole observation period, which presents a negative impact after the pilot reform. As for scale efficiency, we find a perfect parallel time trend before the treatment. The impact of pilot reform materializes quickly in the change of scale efficiency. The magnitude of the coefficient demonstrates a positive effect immediately when the pilot reform was launched.

This study conducts different estimation methods as the robustness check to the baseline results. Tobit model with Least Squares Dummy Variable (LSDV) and random effects settings are introduced to estimate policy impact. Table 10 in Appendix 1 reports the results of robustness in hospital efficiency. As for the policy impact, the magnitude of the LSDV method estimate is identical to the baseline estimates. The results of control variables are also robust compared with the baseline model. However, the estimates of policy impact in the random effects model demonstrate a slightly lower magnitude, comparing the estimates of the baseline model and LSDV model. The difference in estimation could be attributed to the misspecification without addressing year and individual fixed effects. Table 11 in Appendix 1 and Table 12 in Appendix 1 clearly demonstrate similar patterns as Table 10 in Appendix 1 in robustness checking. Overall, the magnitude and significance of estimates in pure technical efficiency and scale efficiency are robust between the baseline model and Tobit model with LSDV setting.

Heterogeneous treatment effect

We also investigate the impact of the pilot reform on healthcare utilization and the heterogenous treatment effects in terms of hospital types. As mentioned in the background section, the zero-markup drug policy shares the overlapping pilot list with county hospital pilot reform. Previous studies on zero-markup drug policy found that the average length of stay in primary care facilities expanded due to the compensation of profit loss [53]. Fu et al. found a negative and insignificant impact on average stay length and inpatient admissions [24]. In our study, pilot hospital reform is negatively associated with county hospitals' average length of stay and bed use rate (Table 6). The average length of stay in the pilot hospitals reduces to around 0.31 days after the policy implementation. This result could be explained by the constraints from the demand side. County hospitals experienced scale expansion in labor and capital resources (supply side). However, the demand side of health seeking may not consistently increase as the supply side.

Furthermore, the impact of the pilot reform on efficiency is heterogeneous in terms of hospital type. We conducted a subgroup analysis by separating the sample into general hospitals and traditional Chinese medicine hospitals (TCM hospitals). Overall, general hospitals experience a larger decline in efficiency after the policy implementation. Table 7 shows that general hospitals have a 2.8% decrease in overall hospital efficiency and a 3.7% decrease in pure technical efficiency. In contrast, the coefficients for TCM hospitals suggest that the pilot reform has a lower and insignificant effect on efficiency for TCM hospitals.

As shown in Table 4, we observed a drop in average hospital efficiency in 2013. One of the potential mechanisms is that the shift of an efficient frontier will widen the gap between the most efficient hospitals and other hospitals. The change of the efficient frontier may result in relatively lower efficiency scores for hospitals that experienced the adjustment from the pilot reform. Hence, we provide further analysis of the distribution of hospitals in terms of different efficiency categories. As Table 13 in Appendix 1 presented, we observe a rapid decline in the number of second-tier efficient hospitals in 2013. Additionally, the trend of efficiency decline reversed in 2014, consistent with the hospital average efficiency scores findings.

Table 6 Impact of pilot reform on health care utilization

Variables	(1) Average Length of Stay	(2) Bed Use Rate
Post*Pilot Reform	−0.314***	-0.030**
	(0.109)	(0.012)
Log GDP per capita	−0.804	−0.043
	(0.820)	(0.047)
Log Government Revenue	0.993***	0.007
	(0.208)	(0.020)
Log Government Expenditure	−0.328	−0.041
	(0.259)	(0.032)
Services Industry Proportion	0.698	0.135
	(2.074)	(0.165)
Log Healthcare Institution Bed	0.095	−0.096***
	(0.207)	(0.028)
Share of Education Enrollment	−0.011	0.011**
	(0.031)	(0.005)
Year Effects	Yes	Yes
Individual Effects	Yes	Yes
N	2555	2555
Adj-R^2	0.814	0.574

Note: The numbers in parentheses are standard errors. Significant levels are denoted as $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Control variables include GDP per capita, government revenue per capita, government expenditure per capita, the output in the services industry as a share of GDP, the number of beds in healthcare institutions in population (10,000), and the share of education enrollment in the regional permanent population

Table 7 Heterogenous treatment effects: impact of pilot reform on hospital efficiency (by hospital types)

Variables	**General Hospitals**			**Traditional Chinese Medicine Hospitals**		
	Hospital Efficiency	**Pure Technical Efficiency**	**Scale Efficiency**	**Hospital Efficiency**	**Pure Technical Efficiency**	**Scale Efficiency**
Post*Pilot Reform	−0.028*	−0.037**	0.015	−0.007	−0.023	0.042
	(0.014)	(0.015)	(0.010)	(0.020)	(0.019)	(0.030)
Log GDP per capita	−0.064	−0.062	−0.032	−0.016	−0.065	0.205
	(0.068)	(0.069)	(0.028)	(0.079)	(0.079)	(0.173)
Log Government Revenue	−0.013	−0.010	−0.021	0.030	−0.003	0.043
	(0.020)	(0.022)	(0.017)	(0.038)	(0.030)	(0.041)
Log Government Expenditure	−0.046*	−0.050*	0.009	−0.052	−0.009	−0.068
	(0.028)	(0.030)	(0.022)	(0.056)	(0.037)	(0.121)
Services Industry Proportion	0.329*	0.349*	0.057	0.114	0.233	−0.316
	(0.186)	(0.196)	(0.172)	(0.282)	(0.284)	(0.324)
Log Healthcare Institution Bed	−0.021	−0.002	−0.038***	−0.051	−0.050	−0.049*
	(0.024)	(0.025)	(0.013)	(0.039)	(0.038)	(0.028)
Share of Education Enrollment	−0.003	−0.005	0.003	0.008	0.007	0.008
	(0.005)	(0.006)	(0.004)	(0.006)	(0.006)	(0.005)
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	1538	1538	1538	1017	1017	1017

Note: The numbers in parentheses are standard errors. Significant levels are denoted as $^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$. Control variables include GDP per capita, government revenue per capita, government expenditure per capita, the output in the services industry as a share of GDP, the number of beds in healthcare institutions in the population (10, 000), the share of education enrollment in population

Discussion

This study aims to measure the efficiency changes of county-level hospitals and investigate the impact of pilot reform on hospital efficiency. Using data envelopment analysis, we conducted the efficiency measurement of county-level hospitals in East China from 2009 to 2015. Our results suggest that hospitals in pilot counties are 6.1% (percentage point) more efficient than those in non-pilot counties. County hospitals' overall efficiency and pure technical efficiency declined one year after the policy was implemented, then experienced a dramatic increase in 2014. Results of DID specification point to a mixed effect on hospital efficiency. On the one hand, the pilot reform is negatively associated with the pure technical efficiency of county hospitals. On the other hand, the scale efficiency of hospitals in pilot counties increases by 2.3% after policy implementation.

Although surprising at first glance, our findings on efficiency estimates are consistent with recent studies on China's health system and county hospital efficiency measures. While some researchers reported that most county hospitals performed better in hospital efficiency after public hospital reform [16, 17]. Technological diffusion theory explain the decline in hospital efficiency and pure technical efficiency [54]. After the national healthcare reform, hospitals were encouraged to introduce modern technologies to increase revenue. However, technological diffusion is driven by hospitals and healthcare staff. It takes time for county hospitals to achieve technical progress after systematic reform. We also need to be cautious about the disparity of technology diffusion in different regions. Chai et al. observed heterogeneities in health system productivity in different provinces [23]. They found some less well-off counties experienced a shock of declining productivity after 2009. Policy-makers need to pay more attention to the regional heterogeneity of healthcare infrastructure, which may lead to the negative impact of the pilot reform and exacerbate the inequality of healthcare services.

Another possible reason for the inconclusive association between the pilot reform and hospital efficiency is the outward movement of the efficiency frontier. We examined this hypothesis by illustrating the distribution of hospitals in different efficiency categories. It is noted that many hospitals experienced downshifts in efficiency in 2013. In contrast, the number of most efficient hospitals remains stable over the observation period. During the pilot reform, few highly efficient hospitals quickly adjusted to the change and pushed the frontier up, eventually resulting in relatively lower efficiency measures for other hospitals.

Our study complements existing research on hospital efficiency measurement from a policy evaluation perspective. The results of DID approach showed that most pilot county hospitals experience a decline in hospital efficiency and pure technical efficiency, indicating inefficient resource utilization during the first phase of pilot reform. This phenomenon might be attributed to inadequate hospital management [35]. County hospitals in China are under the auspices of the local government and health bureau. The principal managers are often appointed by government bodies. Most county hospital managers are medical specialists but lack managerial experience running a hospital. Once the pilot hospitals receive government support, hospital managers invest the resources in scale expansion (such as adding more beds or hiring staff). The findings of a positive association in scale efficiency partially verified the hypothesis of scale expansion in pilot hospitals. Previous studies on government subsidies also revealed that hospitals with more local government support are more likely to face difficulties in resource utilization [16, 17]. Challenges in hospital leadership and management practice call for further reforms to improve hospital governance, enhance regulation, and promote a comprehensive assessments of hospital performance [55].

Aside from hospital governance, balancing the multiple objectives of a hospital may also be relevant to the reduction in hospital efficiency associated with the pilot reform. Findings in Japan's health system show that hospitals would experience a short-term efficiency decline when faced with the trade-off between medical service provision and quality of care [22]. The public hospital reform in China has multiple mandates for county hospitals, including reducing the reliance on medical sales, controlling the cost of medical treatment, establishing the role of gatekeeper of family physicians for healthcare utilization, etc. The main objectives of the pilot reform at that period were to maintain social welfare, enhance healthcare capacity, and provide accessible healthcare services for the population [56]. Devoting resources to those objectives may lead to better accessibility and equity but with short-term reductions in technical efficiency among county hospitals.

The sensitivity analysis explores the policy impact heterogeneity in terms of hospital types. The results of heterogenous treatment effects suggest that the pilot reform has limited effects on Traditional Chinese Medicine hospitals to improve efficiency and achieve scale efficiency. The role in healthcare service delivery could be one possible explanation. Compared with general hospitals, TCM hospitals mainly provide herbal medicine and minority general medical services to county residents [57]. Besides, the difference in resource allocation strategies could result in inconsequential impacts among TCM hospitals. The Zero-markup Drug Policy was implemented in the pilot county hospitals after 2012, which strongly reduced the cost of medicine in medical treatment. However, the cost of Traditional Chinese Medicine is not controlled by the ZMDP. Even if they received governmental subsidies and medical resources, TCM hospitals might be less motivated to expand scale or purchase high-value devices to compensate for reduced drug sales than general hospitals.

Conclusion

This study examines the association between the pilot public hospital reform and county hospital efficiency. We find an unintended consequence of the reform: no improvement in hospital efficiency in healthcare services delivery was found after the pilot reform in county hospitals. The findings imply inefficient utilization of medical resources and shed light on scientific management, which helps to address the potential conflict on a multitude of policy objectives for public hospitals.

Our research also extends current studies in hospital efficiency by investigating the impact of the pilot reform and provides valuable policy implications to county hospitals and the health system. County-level hospital leaders must improve their internal management and hospital operations competency. Then, county hospitals would be more likely to achieve optimal scale efficiency. Policymakers could regard efficiency scores as one of the monitoring indices in assessing the performance of hospitals in certain periods and guiding decision-making.

Our study has at least three limitations. First, the sample of county hospitals is selected from six provinces in East China. The impacts of the pilot reform may differ in the central and western areas; thus, we need to use caution in extrapolating our results to other regions of China. Second, we analyze the policy effects shortly after the reform. It may take longer for the pilot hospitals to materialize the efficiency gains. Last, we could not assess the change in the quality of care due to the data limitations. Further studies are needed to investigate the quality performance by adding health outcomes indicators.

Appendix 1

Table 8 Covariates before and after propensity score matching specification (Match year: 2011)

Covariates	Before matching			After matching		
	(1)	(2)	(3)	(4)	(5)	(6)
	Pilot (*N* = 206)	Non-pilot (*N* = 543)	difference	Pilot (*N* = 190)	Non-pilot (*N* = 175)	difference
GDP per capita	4.483	3.194	1.289***	4.293	4.493	−0.2
Government Revenue	21.815	12.132	9.683***	19.933	23.246	−3.313
Government Expenditure	31.553	22.301	9.252***	29.750	34.760	−5.01
Healthcare Institution Bed	30.731	27.079	3.652***	29.201	29.501	−0.3
Services Industry Proportion	0.329	0.325	0.004	0.327	0.319	0.008
Education Enrollment	8.615	7.472	1.143*	8.347	9.956	−1.609

Note: Statistical significance: ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$. One-to-one matching without replacement and repeated counties. Covariates include GDP per capita, government revenue per capita, government expenditure per capita, the output in the services industry as a share of GDP, the number of beds in healthcare institutions in the population (10, 000), the share of education enrollment in population

Table 9 Impact of Pilot Reform on Hospital Efficiency (with Bias-corrected efficiency)

Variables	(1) Hospital Efficiency	(2) Pure Technical Efficiency	(3) Scale Efficiency	(4) Bias-Corrected Efficiency
Post*Pilot Reform	−0.019	−0.030**	0.023**	−0.019*
	(0.012)	(0.012)	(0.010)	(0.010)
Log GDP per capita	−0.050	−0.068	0.019	−0.052
	(0.054)	(0.055)	(0.034)	(0.044)
Log Government Revenue	0.008	−0.003	0.007	0.007
	(0.019)	(0.018)	(0.021)	(0.023)
Log Government Expenditure	−0.050*	−0.037	−0.015	−0.016
	(0.028)	(0.025)	(0.037)	(0.025)
Services Industry Proportion	0.232	0.300*	−0.061	0.299**
	(0.156)	(0.161)	(0.140)	(0.143)
Log Healthcare Institution Bed	−0.034	−0.022	−0.040***	−0.027*
	(0.021)	(0.021)	(0.015)	(0.016)
Share of Education Enrollment	0.002	−0.000	0.004	−0.003
	(0.004)	(0.004)	(0.003)	(0.003)
Year Effects	Yes	Yes	Yes	Yes
Individual Effects	Yes	Yes	Yes	Yes
N	2555	2555	2555	2555

Note: The numbers in parentheses are standard errors. Significant levels are denoted as *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Bias-corrected Efficiency is the DEA estimates after 1000 times Simarwilson's bootstrapped-DEA approach. Control variables include GDP per capita, government revenue per capita, government expenditure per capita, the output in the services industry as a share of GDP, the number of beds in healthcare institutions in the population (10, 000), the share of education enrollment in population

Table 10 Robustness: impact of Pilot Reform on Hospital Efficiency (different estimation methods)

Variables (Y = Hospital efficiency)	(1) Baseline Model	(2) Tobit Model with LSDV	(3) Tobit Model with Random Effects
Post*Pilot Reform	−0.019	−0.019**	−0.015**
	(0.012)	(0.007)	(0.008)
Log GDP per capita	−0.050	−0.049*	0.030
	(0.054)	(0.030)	(0.019)
Log Government Revenue	0.008	0.008	0.046***
	(0.019)	(0.013)	(0.017)
Log Government Expenditure	−0.050*	−0.046**	−0.032*
	(0.028)	(0.020)	(0.018)
Services Industry Proportion	0.232	0.226**	0.555***
	(0.156)	(0.099)	(0.095)
Log Healthcare Institution Bed	−0.034	−0.029**	−0.041***
	(0.021)	(0.012)	(0.016)
Share of Education Enrollment	0.002	0.001	0.002
	(0.004)	(0.002)	(0.002)
Year Effects	Yes	Yes	NO
Individual Effects	Yes	Yes	NO
Rho			0.417
Log likelihood		2242.98	1091.47
N	2555	2555	2555

Note: The numbers in parentheses are standard errors. Significant levels are denoted as *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. Control variables include GDP per capita, government revenue per capita, government expenditure per capita, the output in the services industry as a share of GDP, the number of beds in healthcare institutions in the population (10, 000), the share of education enrollment in population

Table 11 Robustness: impact of pilot reform on pure technical efficiency (different estimation methods)

Variables (Y = Pure technical efficiency)	(1) Baseline Model	(2) Tobit Model with LSDV	(3) Tobit Model with Random Effects
Post*Pilot Reform	−0.030**	−0.029***	−0.024***
	(0.012)	(0.008)	(0.008)
Log GDP per capita	−0.068	−0.063**	0.054**
	(0.055)	(0.030)	(0.024)
Log Government Revenue	−0.003	−0.003	0.033**
	(0.018)	(0.014)	(0.017)
Log Government Expenditure	−0.037	−0.038*	−0.023
	(0.025)	(0.021)	(0.016)
Services Industry Proportion	0.300*	0.279***	0.635***
	(0.161)	(0.102)	(0.096)
Log Healthcare Institution Bed	−0.022	−0.017	−0.032*
	(0.021)	(0.013)	(0.018)
Share of Education Enrollment	−0.000	−0.001	0.004**
	(0.004)	(0.002)	(0.002)
Year Effects	Yes	Yes	NO
Individual Effects	Yes	Yes	NO
Rho			0.578
Log likelihood		2291.96	1193.35
N	2555	2555	2555

Note: The numbers in parentheses are standard errors. Significant levels are denoted as $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Control variables include GDP per capita, government revenue per capita, government expenditure per capita, the output in the services industry as a share of GDP, the number of beds in healthcare institutions in the population (10, 000), the share of education enrollment in population

Table 12 Robustness: impact of Pilot Reform on Scale Efficiency (different estimation methods)

Variables (Y = Scale efficiency)	(1) Baseline Model	(2) Tobit Model with LSDV	(3) Tobit Model with Random Effects
Post*Pilot Reform	0.023**	0.011**	0.003
	(0.010)	(0.005)	(0.004)
Log GDP per capita	0.019	0.019	−0.007
	(0.034)	(0.019)	(0.018)
Log Government Revenue	0.007	0.005	−0.003
	(0.021)	(0.008)	(0.017)
Log Government Expenditure	−0.015	−0.009	−0.005
	(0.037)	(0.012)	(0.012)
Services Industry Proportion	−0.061	−0.067	0.060*
	(0.140)	(0.063)	(0.034)
Log Healthcare Institution Bed	−0.040***	−0.026***	−0.028***
	(0.015)	(0.008)	(0.008)
Share of Education Enrollment	0.004	0.004**	−0.002**
	(0.003)	(0.001)	(0.001)
Year Effects	Yes	Yes	NO
Individual Effects	Yes	Yes	NO
Rho			0.326
Log likelihood		3395.14	2349.91
N	2555	2555	2555

Note: The numbers in parentheses are standard errors. Significant levels are denoted as $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Control variables include GDP per capita, government revenue per capita, government expenditure per capita, the output in the services industry as a share of GDP, the number of beds in healthcare institutions in the population (10, 000), the share of education enrollment in the population

Table 13 The distribution of hospitals by different efficiency categories

Efficiency Categories	2009	2010	2011	2012	2013	2014	2015
1	7	7	7	8	6	11	7
0.8–0.99	47	25	37	40	6	92	90
0.5–0.79	247	227	243	276	84	232	231
0.2–0.49	60	102	75	39	258	29	34
< 0.2	4	4	3	2	11	1	3

Notes: Table 13 in Appendix 1 presents the distribution of hospitals' numbers in different efficiency categories. Hospitals in this study were divided into five groups based on efficiency from highest to lowest. The number displayed in each cell is the number of hospitals within that efficiency interval. For example, there were 47 hospitals with efficiency values between 0.8 and 0.99

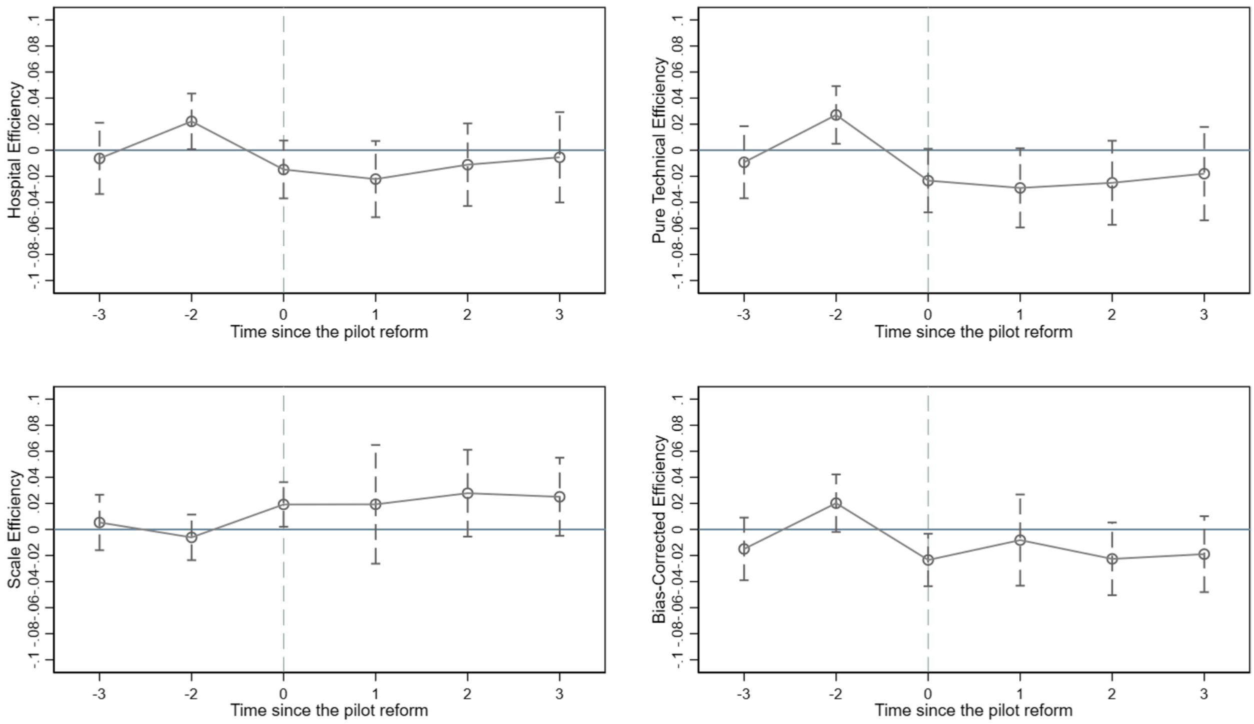

Fig. 1 The event study results in Hospital efficiency, Pure technical efficiency, Scale efficiency, and Bias-corrected efficiency

Appendix 2

The Procedure of Bias-corrected Efficiency Scores

Simar and Wilson first focused on the data-generating process of estimated efficiency. A truncated regression[5] model illustrated the relationship between efficiency and exogenous factors

$$\theta_i = Z_i\beta + \epsilon_i$$

In this data-generating process, θ_i is the technical efficiency and Z_i denotes the exogenous factors influencing hospital efficiency. ϵ_i is the error item following a truncated distribution with zero mean and constant variance.

The main steps of bias-corrected efficiency estimates are as follows. Step 1–4 generate the bias-corrected efficiency scores. Step 5–7 expand the confidence intervals of bias-corrected efficiency to all DMUs.

1. Estimate efficiency scores by using conventional DEA.

[5] Efficiency scores estimated by DEA are bounded to the (0, 1], or [1, ∞] interval based on the different specifications. Efficiency with the [1, ∞] interval is applied for one-sided truncated regression model. Any efficiency score > 1 represents inefficiency. Efficiency with the (0, 1] interval is the regular case in two-sided truncated model.

2. For $\theta_i > 1$, use maximum likelihood to obtain $\widehat{\beta}$ and $\widehat{\epsilon}$ in the truncated regression of θ_i.
3. Loop step 3.1 to 3.4 B_1 times and generate estimates of $\widehat{\theta}_i$ for DMU_i, i = 1-N.

   3.1 For each DMU, draw an artificial $\widetilde{\epsilon}_i$ from N $(0,\widehat{\epsilon}_i^2)$ at $1 - Z_i\widehat{\beta}$.
   3.2 Compute artificial efficiency scores based on the $\widehat{\beta}$ and $\widetilde{\epsilon}_i$.
   3.3 Generate artificial x and y of each DMU with $\widetilde{y_i} = \left(\widehat{\theta}_i / \widetilde{\theta}_i\right) y_i$.
   3.4 Replace DMU with artificial one generated in step 3.3, let them be the sample to draw the bias-corrected efficiency score for each DMU.

4. For DMU_i, i = 1-N, calculated the bias-corrected efficiency score $\widehat{\theta_i^{bc}}$.
5. Use maximum likelihood to estimate new $\widehat{\widehat{\beta}}$ and new $\widehat{\widehat{\epsilon}}$.
6. Loop step 6.1 to 6.3B_2 times to obtain $\widehat{\widehat{\theta}}_i$.

   6.1 For each DMU, draw an artificial $\widetilde{\widetilde{\epsilon}}_i$ from N $(0,\widehat{\epsilon}_i^2)$ at $1 - Z_i\widehat{\widehat{\beta}}$.
   6.2 Compute artificial efficiency scores based on the $\widehat{\beta}$ and $\widetilde{\epsilon}_i$ for each DMU.
   6.3 Use maximum likelihood to estimate $\widehat{\widehat{\theta}}_i$ on Z_i to obtain bootstrapped $\widehat{\widehat{\beta}}$ and $\widetilde{\widetilde{\epsilon}}_i$.

7. Use bootstrapped $\widehat{\widehat{\beta}}$ and $\widetilde{\widetilde{\epsilon}}_i$ to construct the confidence interval for β and ϵ_i.

Abbreviations

DEA	Data envelopment analysis
DID	Difference-in-differences
DMUs	Decision making units
LSDV	Least Squares Dummy Variable
PSM	Propensity score matching
PTA	Parallel trend assumption
TCM	Traditional Chinese medicine
ZMDP	The Zero-markup Drug Policy

Acknowledgements
This research was supported by the General Program of the National Natural Science Foundation of China (Grant#: 72174098). The funding source had no involvement in the research process. We are grateful to Professors Qing-Ping MA, Zhanchun Feng, Rufei Guo, and seminar participants at Wuhan University, The Chinese Economists Society (CES) 2021 annual conference, and the China Health Policy and Management Society (CHPAMS) 2021 symposium. The views in this article do not represent the official positions of authors' employers. All remaining errors are our own.

Authors' contributions
Conceptualization, Wei Jiang and Zhuo Chen; Data Curation, Qiulin Chen and Wei Jiang; Formal analysis, Wei Jiang; Writing-original draft, Wei Jiang; Software, Wei Jiang; Writing-review and editing, Xuyan Lou, Lina Song, and Zhuo Chen.

Data availability
Deidentified data will be available upon request from the corresponding author.

Declarations

Ethics approval and consent to participate
The study uses administrative data and does not contain any studies with human participants or animals performed. The authors received ethical approval for secondary data analysis from the University of Nottingham Ningbo China on Aug 6th, 2020.

Consent for publication
Not applicable. This article does not contain any studies with human participants performed by any of the authors.

Competing interests
The authors declare no competing interests.

Received: 22 July 2024 Accepted: 11 February 2025
Published online: 03 March 2025

References

1. Mitropoulos, Mitropoulos I, Sissouras A. Managing for efficiency in health care: the case of Greek public hospitals. Eur J Health Econ. 2013;14(6):929–38.
2. Savedoff WD. The health financing transition: a conceptual framework and empirical evidence. Soc Sci Med. 2014;105:112–21.
3. Yip W, William C, Hsiao. What drove the cycles of Chinese Health System reforms? Health Syst Reform. 2015;1(1):52–61.
4. Yip W, Fu H, Chen AT, Zhai T, Jian W, Xu R, Pan J, Hu M, Zhou Z, Chen Q, Mao W, Sun Q, Chen W. 10 years of healthcare reform in China: Progress and gaps in Universal Health Coverage. Lancet. 2019;394(10204):1192–204.
5. Zhang Y, Zhou Z, Si Y. When more is less: what explains the overuse of health care services in China? Soc Sci Med. 2019;232:17–24.
6. Chen C, Dong W, Shen JJ, Cochran C, Wang Y, Hao M. Is the prescribing behavior of Chinese physicians driven by financial incentives? Soc Sci Med. 2014;120:40–8.
7. Qian J, He AJ, Yin JD. C. The medical arms race and its impact in Chinese hospitals: implications for health regulation and planning. Health Policy Plann. 2019;34(1):37–46.
8. Zhang Y, Ma Q, Chen Y, Gao H. Effects of Public Hospital Reform on Inpatient expenditures in Rural China. Health Econ. 2017;26(4):421–30.
9. Murray CJL, Frenk J. A framework for assessing performance of health system. Bull World Health Organ. 2000;78(6):717–31.
10. World Health Report. 2000. Health Systems: Improving Performance. World Health Organization: Geneva, 2000.
11. Wagstaff A. Estimating efficiency in the Hospital Sector: a comparison of three statistical cost Frontier models. Appl Econ. 1989;21(5):659–72.
12. Hollingsworth B. Non-parametric and parametric applications measuring efficiency in health care. Health Care Manag Sci. 2003;6(4):203–18.
13. Giuffrida A, Gravelle H. Measuring performance in primary care: Econometric analysis and DEA. Appl Econ. 2001;33(2):163–75.
14. Hollingsworth B, Wildman J. The efficiency of Health Production: re-estimating the WHO Panel Data Using Parametric and non-parametric approaches to provide additional information. Health Econ. 2003;12(6):493–504.
15. Hunt DJ, Link CR. Better outcomes at lower costs? The effect of public health expenditures on hospital efficiency. Appl Econ. 2020;52(4):400–14.
16. Jiang S, Min R, Fang PQ. The impact of healthcare reform on the efficiency of public county hospitals in China. BMC Health Serv Res. 2017;17(1):1–8.
17. Cheng Z, Tao H, Cai M, Lin H, Lin X, Shu Q, Zhang RN. Technical efficiency and productivity of Chinese county hospitals: an exploratory study in Henan province, China. BMJ Open. 2015;5(9):1–10.

18. Zhao L, Wang L, Li S, Zhang Y. Evaluation and analysis of hospital efficiency in China based on macro- and micro-level analyses. J Public Health (Germany). 2020;28(2):191–7.
19. Gong G, Chen Y, Gao H, Su D, Chang J. Has the efficiency of China's healthcare system improved after healthcare reform? A network data envelopment analysis and tobit regression approach. Int J Environ Res Public Health. 2019;16(23);4847.
20. Tiemann O, Schreyögg J. Changes in hospital efficiency after privatization. Health Care Manag Sci. 2012;15(4):310–26.
21. Karmann A, Roesel F. Hospital Policy and Productivity – evidence from German States. Health Econ. 2017;26(12):1548–65.
22. Zhang X, Tone K, Lu Y. Impact of the Local Public Hospital Reform on the efficiency of medium-sized hospitals in Japan: an Improved slacks-based measure data Envelopment Analysis Approach. Health Serv Res. 2018;53(2):896–918.
23. Chai P, Zhang Y, Zhou M, Liu S, Kinfu Y. Health system productivity in China: a comparison of pre- and post-2009 healthcare reform. Health Policy Plann. 2020;35(3):257–66.
24. Fu H, Li L, Yip W. Intended and unintended impacts of price changes for drugs and medical services: evidence from China. Soc Sci Med. 2018;211(January):114–22.
25. Li L, Fu H. China's health care system reform: Progress and prospects. Int J Health Plann Manage. 2017;32(3):240–53.
26. Yip WCM, Hsiao WC, Chen W, Hu S, Ma J, Maynard A. Early appraisal of China's huge and complex healthcare reforms. Lancet. 2012;379(9818):833–42.
27. Li X, Lu J, Hu S, Cheng KK, De Maeseneer J, Meng Q, Mossialos E, Xu DR, Yip W, Zhang H, Krumholz HM, Jiang L, Hu S. The primary health-care system in China. Lancet. 2017;390(10112):2584–94.
28. Zheng D, Gong J, Zhang C. Efficiency of medical service systems in the rural areas of Mainland China: a comparative study from 2013 to 2017. Public Health. 2019;171:139–47.
29. Fang P, Min R, Zou X. Key points and pathway of County Public Hospital Reform in China. Chin Hosp Manage. 2014;34(1):4–8.
30. Zhou Z, Zhao Y, Shen C, Lai S, Nawaz R, Gao J. Evaluating the effect of hierarchical medical system on health seeking behavior: a difference-in-differences analysis in China. Soc Sci Med. 2021;268:113372.
31. Kao C, Pang RZ, Liu ST, Bai XJ. Optimal expansion paths for hospitals of different types: viewpoint of scope economies and evidence from Chinese hospitals. Eur J Oper Res. 2021;289(2):628–38.
32. Lindlbauer I, Schreyögg J, Winter V. Changes in technical efficiency after quality management certification: a DEA approach using difference-in-difference estimation with genetic matching in the hospital industry. Eur J Oper Res. 2016;250(3):1026–36.
33. Hollingsworth B. The measurement of efficiency and productivity of health care delivery. Health Econ. 2008;1131:1127–31.
34. Hussain Z, Huo C, Ahmad A, Shaheen WA. An assessment of economy- and transport-oriented health performance. Health Econ Rev. 2024;14(1):80.
35. Li H, Dong S, Liu T. Relative efficiency and productivity: a preliminary exploration of public hospitals in Beijing, China. BMC Health Serv Res. 2014;14:158.
36. Stefko R, Gavurova B, Kocisova K. Healthcare efficiency assessment using DEA analysis in the Slovak Republic. Health Econ Rev. 2018;8(1):6.
37. Chai P, Zhang Y, Zhou M, et al. Technical and scale efficiency of provincial health systems in China: a bootstrapping data envelopment analysis. BMJ Open. 2019;9:e027539.
38. Renner A, Kirigia JM, Zere EA, et al. Technical efficiency of peripheral health units in Pujehun district of Sierra Leone: a DEA application. BMC Health Serv Res. 2005;5:77.
39. Zhao L, Wang L, Li S, et al. Evaluation and analysis of hospital efficiency in China based on macro- and micro-level analyses. J Public Health (Berl). 2020;28:191–7.
40. Farrell MJ. The measurement of productive efficiency. Journal of the Royal Statistical Society, Series a (General). 1957;120(3);253–90.
41. Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision-making units. Eur J Oper Res. 1978;2(6):429–44.
42. Banker RD, Charnes A, Cooper WW. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Manage Sci. 1984;30(9):1078–92.
43. Jing R, Xu T, Lai X, Mahmoudi E, Fang H. Technical efficiency of public and private hospitals in Beijing, China: a comparative study. Int J Environ Res Public Health. 2020;17(1);82.
44. Chai P, Wan Q, Kinfu Y. Efficiency and productivity of health systems in prevention and control of non-communicable diseases in China, 2008–2015. Eur J Health Econ. 2021;22(2):267–79.
45. Cheng Z, Cai M, Tao H, He Z, Lin X, Lin H, Zuo Y. Efficiency and productivity measurement of rural township hospitals in China: a bootstrapping data envelopment analysis. BMJ Open. 2016;6(11):e011911.
46. Rosenbaum PR, Rubin DB. The Central Role of the Propensity score in Observational studies for Causal effects. Biometrika. 1983;70(1):41–55.
47. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am Stat. 1985;39(1):33–8.
48. Monchuk DC, Chen Z, Bonaparte Y. Explaining production inefficiency in China's agriculture using data envelopment analysis and semi-parametric bootstrapping. China Econ Rev. 2010;21(2):346–54.
49. Simar L, Wilson PW. Estimation and inference in two-stage, semi-parametric models of production processes. J Econ. 2007;136(1):31–64.
50. Badunenko O, Tauchmann H. Simar and Wilson two-stage efficiency analysis for Stata. Stata J. 2019;19(4):950–88.
51. Honore B. Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects. Econometrica. 1992;60(3):533–65.
52. Alan S, Honoré BE, Hu L, Leth-Petersen S. Estimation of Panel Data Regression models with two-sided censoring or truncation. J Econometric Methods. 2014;3(1):1–20.
53. Yi H, Miller G, Zhang L, Li S, Rozelle S. Intended and unintended consequences of China's zero markup drug policy. Health Aff. 2015;34(8):1391–8.
54. Skinner J, Staiger D. Technology diffusion and productivity growth in health care. Rev Econ Stat. 2015;97(5):951–64.
55. Liu X, Wang Q, Sara B, Yang W, Dong S, Li H. (2018). The Technical Efficiency of Community Health Service Centers in Wuhan, China: Estimation and Policy Implications. Inquiry. 2018;55:1–9.
56. Barber SL, Borowitz M, Bekedam H, Ma J. The hospital of the future in China: China's reform of public hospitals and trends from industrialized countries. Health Policy Plann. 2014;29(3):367–78.
57. Jiang X, He P, Zhu D, Shi X, Meng Q. Different impacts of the zero-markup drug policy on county general and traditional Chinese medicine hospitals: evidence from Shandong province, China. Int J Equity Health. 2020;19(1):219.

Publisher's Note