Decentralisation and Public Health Services Provision: Evidence from Administrative Decentralisation Reform in China

January 9, 2019

Abstract

This paper adopts the Difference-in-Difference method and combines it with propensity score matching to identify the causal effect of administrative decentralisation reform in 167 counties, on the local public health services provision from 2000 to 2012, using the quantity of hospital beds per 10,000 residents as an indicator. The results show that there is a significant negative effect of decentralisation reform on the other.

1 Introduction

The central government of China has started endorsing and promoting administrative decentralization reform since 1990s. Specifically, socio-economic affairs previously administrated and audited by prefecture-level cities which then submitted to the provincial level for approval, shall be directly reported by counties to the provincial-level government for further decision. And discretionary power over some local issues and fiscal management will be delegated to county governments[1]. In this way, the administrative efficiency and effectiveness are considered to be enhanced because of administrative hierarchy reduction. This devolution of administrative authority is also conducive to the autonomy of county-level economic development and intergovernmental competition (Bardhan, 2002), which corresponded with central government's intention. Futhermore, according to the *Decentralisation Theorem* by Oates (1999), local governments will be in a better position to provide public outputs that meet the demands of their local residents. But as some people have pointed out that because

[1]China adopts a five-level administrative framework: central government-province-prefecture-level city-county-town.

of the strong linkage between the promotion of local officials and economic development performance in China, local officials are very likely to be committed to accomplishing their economic development missions and allocate more resources to it while overlooking the investment and provision of local public goods and services. Therefore, being given more discretionary power as the reform goes on, public services provision condition may worsen and undermine local citizens' well-being condition and the overall balanced development. Thus it requires further study on decentralisation reform and its potential impact on the provision of local public services.

This reform did not rollout in all provinces and counties at one time. Considering the great disparity among provinces, central government leave it for provinces themselves to make decisions. Limited number of provincial government carried out the administrative decentralisation reform in all the counties in its jurisdiction at one time while others selected a first pilot wave of counties to reform and then decide whether and when to include other counties gradually, which consequently allocated counties into control and treatment groups, though not on a random basis, and thus created a quasi-experiment situation for causal inference.

This paper mainly adopts the Difference-in-Difference method and combines it with propensity score matching, as a way to better eliminate selection bias and potential unobservable variables to identify the causal effect of administrative decentralisation reform on public health services provision, focusing on medical facility provision. The results show that there is a significant negative effect of decentralisation reform on the other. But unfortunately, the robustness of this identification was not able to be tested rightly.

2 Identification Strategy and Data

2.1 Data

Data for this study are collected from China Statistical Yearbook (county-level) and Statistical Yearbook by each province from 2000. Given that central government has certain policy inclination in less developed provinces, if more incomparable provinces and counties are included, though it seems to cover a larger area, the causal effect will be miscellaneous, which lacks precision. Therefore, according to *the Reform plan for the division of fiscal and expenditure responsibilities in the basic public services,* I restricted the study area to 3 provinces because of their same ranking in the plan: Zhejiang, Jiangsu and Fujian.[2] These provinces also share economic and geographical similarities, which hopefully will make the control and

[2] In the reform plan, provinces(cities) are divided into 5 grades. Central government will take on difference proportions of fiscal expenditure on public services for local government in different grades: 1st grade 80%, 2nd grade 60%, 3rd grade 50%, 4th grade 30%, 5th grade 10%(Only Shanghai and Beijing in this grade). Provinces covered in this paper are in the 4th grade, which means that they have better and similar basis in basic public services and they have considerable economic capacity to pay for the corresponding expenditure.

treated counties more overall identical. I also excluded counties that experienced changes in their administrative divisions. And Jiangsu does not have hospital beds number data after 2012 so the time period ends in 2012. Finally, I have data for 119 reformed counties and 48 unreformed counties on 7 years (every other year from 2000 to 2012). Reform time for counties are different: 75 counties reformed on 2003; 41 counties on 2010; 3 counties on 2012.

Because of the non-random assignment setting of the reform and 3 provinces are included, I have to acknowledge that selection bias should be one of the main concerns and Propensity Score Matching will be used to try to overcome it. And because of multiple reform times in the treatment group, I also tried to do a Granger test as a robustness test.

2.2 Dependent variable and Treatment Assignments

Public health services provision is the variable that this paper interested in. Given that the county-level data available to the public is very scanty, this paper chooses "the number of hospital beds owned by per 10,000 people"(at county level) as the dependent variable. As to treatment assignments, in the provinces where the reform is incrementally promoted, if a county is assigned as a pilot county by its provincial government in the policy document, then it will be divided into treatment group, otherwise control group. The assignment criteria normally will be mentioned in the policy document as well.[3]For those provinces that rollout the reform throughout its jurisdiction at once, all of its counties will be viewed as treatment receivers. As to the timing of treatment assignment, namely the actual reform start time, it is decided according to the date that the policy document was enacted: if the date was in the fist half of the year, then the start point should be current year, otherwise the next year. It is for the reason that the data this paper use (as noted in the section 2.1) is collected at the end of each year, thus it might be reasonable to make this distinction so as to obtain more precise causal effect estimations.

2.3 Difference-in-Difference

The aim of this paper is to estimate the causal effect of governance decentralisation reform on the provision of medical facilities. Ideally, authentic causal effect can be calculated by comparing both of the outcomes of a county being treated and not treated. However, the *counterfactual* result cannot be directly observed, that is to say, we cannot observe the outcome of a county would have produced had it not been introduced into reform, as the *Fundamental Problem of Causal Inference* states. To overcome this problem, this paper adopts the Difference-in-Difference method to perform an estimation of the *average treatment effect on the treated*(ATT). If the average treatment effect is significant, then we might be

[3]See2.4. It also to some extent indicates that the dependent variable of this paper is not one of the variables accounting for selection bias so we can lessen our suspicion on estimating an inverse causal relationship.

convinced that the governance decentralisation has a substantial impact on its provision of public health services.

There are two more specific reasons on why a DiD method is the most appropriate one in this case. Firstly, reforming on the government administrative system probably will take years for different levels of governments to fit in with each other. Compared with matching method, we will be able to include observations on several years in DiD and see the trend of its impact and control on some time-variant observable variables which may also influence on the provision of public health services. Secondly, another concern is the selection bias in this causal inference, i.e., whether counties being introduced into reform or not were not randomly assigned. Reformed counties and unreformed counties may have already differenced in some way before the reform happened. Given that a convincing instrument variable is difficult to find, DiD can help us to estimate on the treatment effect while fixed on some time invariant differences in pre-treatment and post-treatment periods[4].

2.3.1 DiD Setup

The units of this study are counties, denoted as i. The indicator for treatment is D_i,

$$D_i = \begin{cases} 1 & \text{if county was assigned to reform,} \\ 0 & \text{if county was not assigned to reform.} \end{cases} \tag{1}$$

The indicator for treatment period is T,

$$T_i = \begin{cases} 1 & \text{Post-treatment period,} \\ 0 & \text{Pre-treatment period.} \end{cases} \tag{2}$$

The potential outcome can be defined as $\mathrm{Y}_{di}(t)$,

$$Y_{di}(t) = \begin{cases} \mathrm{Y}_{1i}(t) & \text{Potential outcome for county i in period t when treated,} \\ \mathrm{Y}_{0i}(t) & \text{Potential outcome for county i in period t when controled.} \end{cases} \tag{3}$$

Then the causal effect for county i at time t can be defined as,

$$\tau_i(t) = Y_{1i}(t) - Y_{0i}(t) \tag{4}$$

The observed outcome for a given county i is,

$$Y_i(t) = Y_{1i}(t)D_i(t) + Y_{0i}(t)(1 - D_i(t)) \tag{5}$$

[4]Assumptions required for estimation are explained in section 2.3.1.

Because of the *fundamental problem of causal inference*, we can only observe the outcome for each county i either under treatment or control, but by using the DiD method, we can turn to calculate the *average treatment effect on the treated*(ATT),

$$\tau_{ATT} = E[Y_{1i}(1) - Y_{0i}(1)|D_i = 1] \tag{6}$$

To make further identification, two assumptions must be made. The first one is *parallel trends assumption*, which in this paper means that the trends of "the quantity of hospital beds" would be the same in both reformed and unreformed counties in the absence of the administrative decentralisation reform. We can herefrom identify $\tau_{ATT} as$,

$$\tau_{ATT} = \{E[Y_i(1)|D_i = 1] - E[Y_i(1)|D_i = 0]\} - \{E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0]\} \tag{7}$$

The plausibility of parallel trends assumption will be investigated in 3.1 below. And the second assumption is that there is no other shocks on the dependent variables during the period of reform. I have tried to lessen this concern by restricting on study area.

2.3.2 Estimation

Based on Angrist and Pischke(2009, 299), the basic DiD fixed-effect regression equation in this study can be put as:

$$Y_{it} = \gamma_i + \lambda_t + X'_i\beta + \delta D_{it} + \delta + \varepsilon_{it} \tag{8}$$

The left-hand side of this equation is the dependent variable, i.e."the quantity of hospital beds owned by per 10,000 people", for county i at time t. γ_i is the county fixed effect, which captures unobservable confounders that are individually varying but time invariant; λ_t is the year fixed effect, which can capture unobservable and additive confounders that are individually invariant but time varing. X'_i is the vector of observable time-varying covariates for each county, where this paper chose per capita GDP, population density, fiscal expenditure and fixed investments(West and Wang, 1995;Kristiansen and Santoso, 2006). β is the vector of coefficients for covariates. D_{it} is the indicator for the reform of county i in time t, and δ is the causal effect of introducing administrative decentralisation reform on the provision of hospital bess. ε_{it} is the error term, where E($\varepsilon_{it}|i,t$)=0. To also control on specific time trend across difference counties, γ_{1i} can be included in the equation as (Angrist and Pischke 2009: 299),

$$Y_{it} = \gamma_{0i} + \gamma_{1i}t + \lambda_t + X'_i\beta + \delta D_{it} + \delta + \varepsilon_{it} \tag{9}$$

OLS regression will be adopted to make the DiD estimation. In addition, as Angrist and Pischke(2009) noted, regional shocks are very likely to be highly serially correlated. Thus standard error clustered by *county* instead of *county*$\times$*year* will be reported in the regression results.

2.4 Matching Before Difference-in-Difference

As previously noted, selection bias is a great concern, thus I chose Propensity Score Matching (PSM) to try to eliminate it. By calculating the propensity score of each county entering the reform, it can reduce multiple dimensions of covariates into one (Rosenbaum and Rubin, 1983). Thus, it helps to overcome the *curse of dimensionality*, that is to avoid a substantial number of counties being eliminated because of being not able to match on a matchable county when increasing covariates' quantity. This paper uses a "straightforward" approach of combining matching and difference in difference as in the paper by Ladd and Lenz (2009), which is matching the treated and control group on observable covariates and drop units that are unmatched, in order to make two groups more similar on these characteristics. Base on the policy document of Zhejiang province and data availability, I chose the proportion of rural population to account for urbanisation level, population density for economic growth potential and per capita GDP for regional productivity, as covariates to estimate the probability of a county being introduced into the reform.[5]

3 Results

3.1 Plausibility of Parallel Trends

As noted in 2.3.2, a prerequisite for DiD is the parallel trends of dependent variable, which in this paper, is the parallel trends of hospital beds quantities in the reformed and unreformed counties before the start of decentralisation reform (in 2003). I used the full group of counties and post-matching counties to plot. From Figure 1 and 2, we can see that, hospital beds quantities basically keep increasing over time. The gap between the outcomes for both groups narrowed after the reform and then widened before reaching stable again. And before Matching, the trends during 2000-2002 seems to be parallel, but the trends between 2002 and 2004 intersect each other, which is around the reform start point, but we do not know whether this is only caused by the reform or not. While after matching, the parallel trends for reformed and unreformed counties before 2002 still holds. For 2002-2004 period, there is no intersection between the two groups and the trends become somewhat more parallel, which indeed helped to correct the selection bias among them. Given that there were no observations on the reform start year, the parallel trends assumption for groups after matching is reasonably plausible, which satisfied the prerequisite for using DiD.

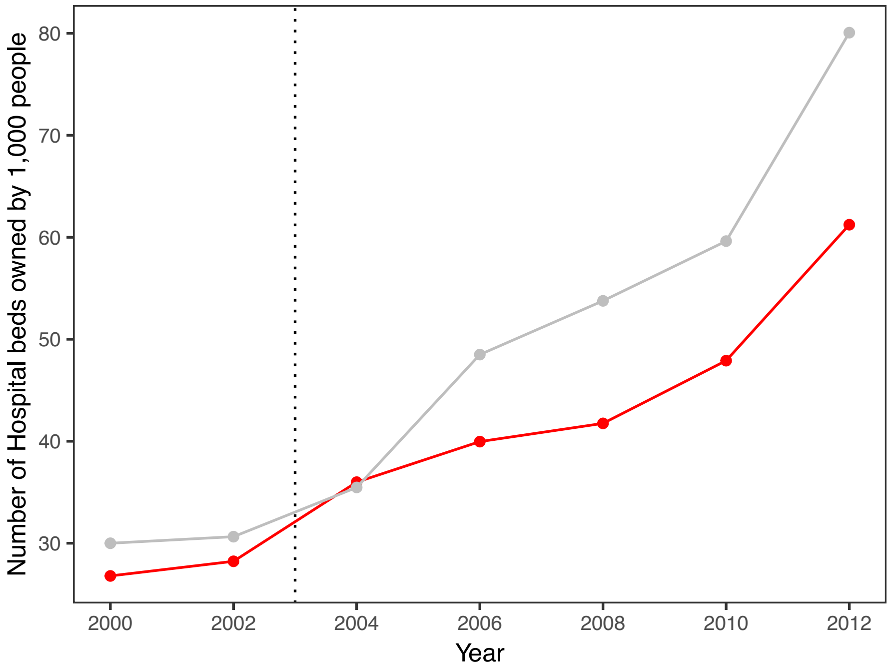

Figure 1: Trends in the quantity of hospital beds owned by per 10,000 people using all counties from 2000 to 2012

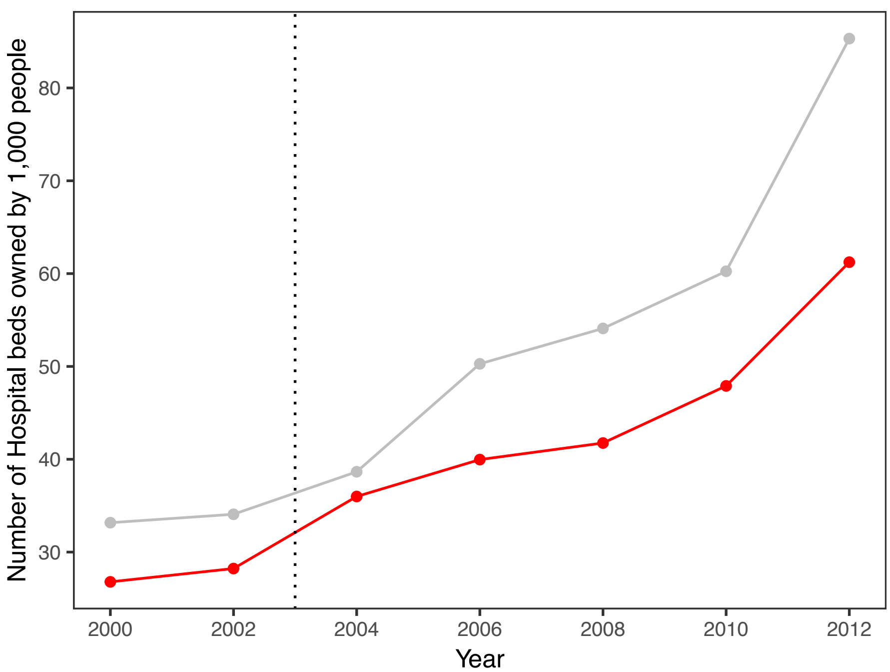

Figure 2: Trends in the quantity of hospital beds owned by per 10,000 people using matched counties from 2000 to 2012

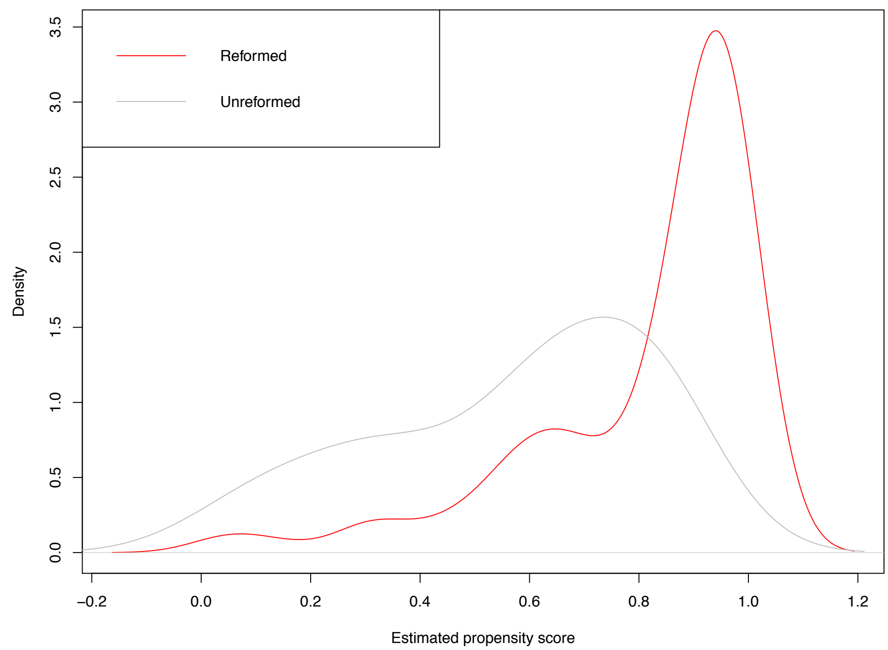

Figure 3: Propensity Score Distribution for matched counties by treatment status

Figure 4: Standardized Mean Differences for Propensity Score and Covariates Before and After Matching

Table S1: Regression Results After Matching

	Hospital Beds per 10,000 people			
	1	2	3	4
Decentralisation Reform	17.45	-6.36	-4.30	-2.47
Clustered SE	1.33	1.47	1.55	1.31
P-Value	$6.30e^{-36}$	$1.68e^{-5}$	$5.6e^{-3}$	0.06
Observations	1001	1001	1001	1001
Counties	143	143	143	143
County FEs	Yes	Yes	Yes	Yes
Year FEs	No	Yes	Yes	Yes
Covariates	No	No	Yes	Yes
County Specific Trends	No	No	No	Yes
Effect Size	64.44%	-12.50 %	-8.80 %	-5.25%
Ub Effect Size	81.96%	-7.25 %	-2.76 %	0.24%
Lb Effect Size	49.99%	-17.19%	-14.14%	-10.17%

Note: The mean value for hospital beds quantity by per 10,000 people is 44.54.

3.2 Matching

This paper used Logit model to estimate propensity score for each county and then adopted the genetic matching to match on reformed and unreformed counties for it helps to "obtain better levels of balance without requiring the analyst to correctly specify the propensity score" (Diamond and Sekhon, 2013). All of the unreformed counties are matched and 24 reformed counties are unmatched and dropped, leaving 143 matched counties in total. From Figure 3 we can see that the ranges of propensity score for two groups are almost the same, which satisfies the *common support* assumption for Matching. And reformed counties assemble around higher propensity score range than unreformed counties. But because the number of reformed counties are more than the unreformed, the distribution pattern for two groups are not very identical as in a ideal situation.

In Figure 4, all of the standardized mean differences for covariates as well as the propensity score are closer to 0 after matching, which proves the effectiveness of the matching process and relieves selection bias to some degree.

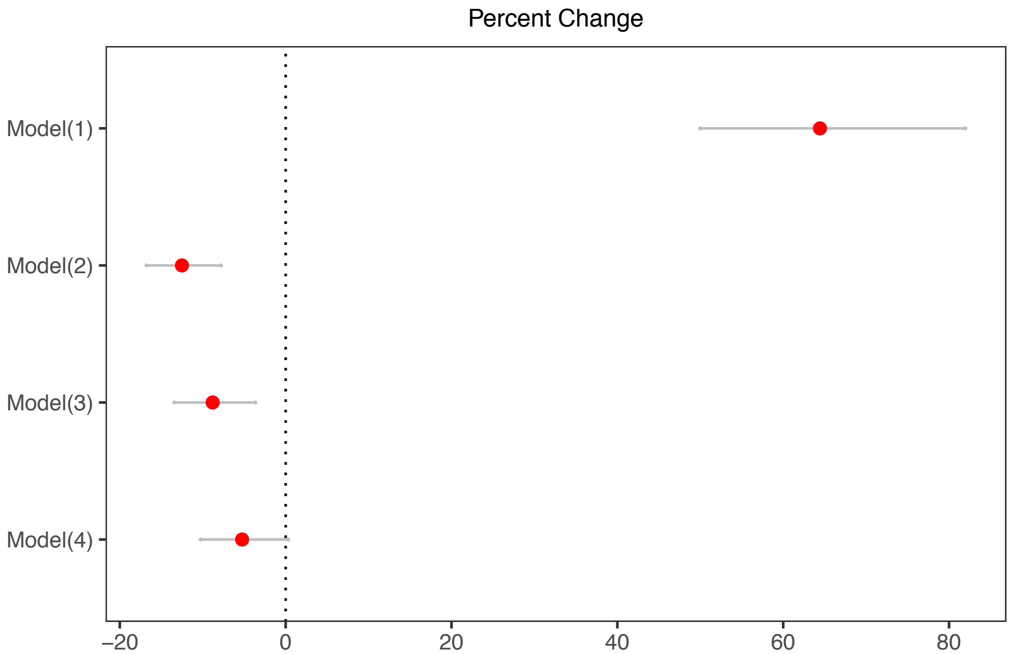

Figure 5: The effect sizes of decentralisation reform for models using matched counties

3.3 DiD Regression

In Table S1, four models are reported, which are oneway county fixed effect, twoway county-year fixed effect, twoway fixed effect with covariates and county specific trends included respectively. Table S2 in the Appendix shows regression results using full groups. The first row in the results part shows the coefficients of the treatment, i.e., administrative decentralisation reform. And Figure 5 presents the effect size of treatment in percentage. When only county fixed effects were included in the model, the reform had a very significant and strong positive effect on counties' hospital beds quantity. While after controlling on year fixed effects, the effect decreased dramatically to -6.36, which suggests that there are some strong individual invariant but time varying variables affecting negatively on hospital beds provision during the decentralisation reform period. After further controlling on covariates and county specific trends, the negative effect of the reform decreases slightly to -2.47 when fully controlled, meaning that after county governments engaging in the decentralisation reform, there will be an average decrease of 2.47 hospital beds (per 10,000 people) in these reformed counties ($p=0.06<0.1$). The county clustered standard errors do not change much across the models. But same as the coefficients, the p-values of them are also very sensitive and increase significantly as more fixed effects and variables adding in, but still reach a significance level of 0.1. The results of Regression using full groups of counties are similar to

[5] According to the Zhejiang Provincial Government (2002), "status in regional production, economic size, urbanisation level, and economic growth potential" were mentioned.

only including matched counties, but the effect of decentralisation reform on hospital beds provision is not significant after fully controled (p=0.11>0.1).

4 Robustness Test

As there are multiple reform times across counties in the treatment group, the average reform effect will be more robust if it can stand the Granger test, which is testing on whether, conditional on county and time fixed effect , past D_{it} predict Y_{ist}, while future D_{it} do not. To perform this test, I can modify equation9 to (Angrist and Pischke,2009: 237),

$$Y_{it} = \gamma_i + \lambda_t + \sum_{\tau=0}^{m} \delta_{-\tau} D_{i,t-\tau} + \sum_{\tau=1}^{q} \delta_{+\tau} D_{i,t+\tau} + X_{it}^{'}\beta + \varepsilon_{it}$$

which allows for m posttreatment effects (lags), and q anticipatory effects (leads) of reform. If the anticipatory effects is close to 0 and posttreatment significantly not equal, then I do not need to worry about the causal interpretation of my previous identification. But as I run the codes for leads and lags, the P value of reform effects on all of the post and pre-treatment time points are equal to exact 0, which is very problematic, thus I do not report it here.

5 Conclusions

This paper tried to identify the causal effect of the administrative decentralisation reform on the public health services provision in 3 provinces in China from 2000 to 2012, using the quantity of hospital beds per 10,000 residents as an indicator. After using propensity score matching and difference in difference method, I found that the reform has some significantly negative effect on the public health services provision, which is contrary to some views that administrative decentralisation will allow local governments to be more efficient and meet with residents' demand. And in the DiD regression, there seems to be a strong time varying trend that is co-influencing the public health service provision. As to the external validity of this paper, because of variant policy environment and original background of different provinces, the results of this paper cannot be generalised. But the identification strategy can be replicated on other object provinces with some modifications. However, the causal effect has not been substantially assessed under the expected robustness test, Granger test, which remains some suspicion on the its internal credibility.

A Appendix

Table S2: Regression Results

	Hospital Beds per 10,000 people			
	1	2	3	4
Decentralisation Reform	17.45	-7.45	-5.19	-2.02
Clustered SE	1.33	1.43	1.50	1.26
P-Value	$2.3e^{-36}$	$2.3e^{-7}$	$5.7e^{-4}$	0.11
Observations	1169	1169	1169	1169
Counties	167	167	167	167
County FEs	Yes	Yes	Yes	Yes
Year FEs	No	Yes	Yes	Yes
Covariates	No	No	Yes	Yes
County Specific Trends	No	No	No	Yes
Effect Size	64.44%	-14.33%	-10.44%	-4.34%
Ub Effect Size	81.96%	-9.45%	-4.81%	1.02%
Lb Effect Size	49.99%	-18.71%	-15.45%	-9.16%

Note: The mean value for hospital beds quantity per 10,000 people is 44.54.

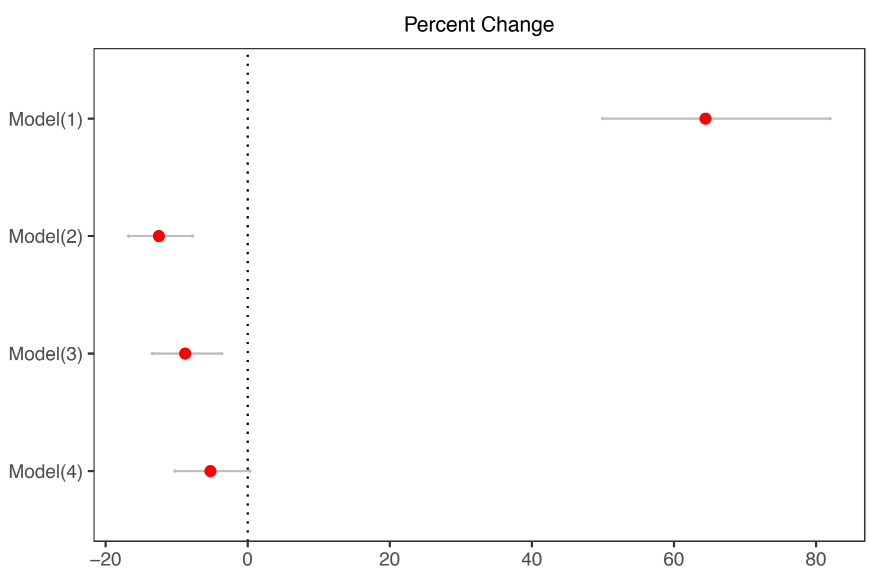

Figure 6: The effect sizes of decentralisation reform for models using full groups of counties

Bibliography

Akin, J., Hutchinson, P., & Strumpf, K. (2005). Decentralisation and government provision of public goods: The public health sector in Uganda. The Journal of Development Studies, 41(8), 1417-1443.

Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Bardhan, P. (2002). Decentralization of governance and development. Journal of Economic perspectives, 16(4), 185-205.

Diamond, A., & Sekhon, J. S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. Review of Economics and Statistics, 95(3), 932-945.

Gerber, A. S., & Green, D. P. (2012). Field experiments: Design, analysis, and interpretation. WW Norton.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. Journal of Statistical Software, 42(8), 1-28.

Kristiansen, S., & Santoso, P. (2006). Surviving decentralisation?: Impacts of regional autonomy on health service provision in Indonesia. Health Policy, 77(3), 247-259.

Oates, W. E. (1999). An essay on fiscal federalism. Journal of economic literature, 37(3), 1120-1149.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41-55.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. Statistical science: a review journal of the Institute of Mathematical Statistics, 25(1), 1.

Vlahov, D., & Galea, S. (2002). Urbanization, urbanicity, and health. Journal of Urban Health, 79(1), S1-S12.

West, L. A., & Wong, C. P. (1995). Fiscal decentralization and growing regional disparities in rural China: Some evidence in the provision of social services. Oxford Review of Economic Policy, 11(4), 70-84.