DOES SUPER-DEPARTMENT REFORM IMPROVE PUBLIC

SERVICE PERFORMANCE IN CHINA?

Liang Ma

Nanyang Centre for Public Administration

Nanyang Technological University

Block S3.2 Level B4, 50 Nanyang Avenue, Singapore 639798

Tel: (65) 6592-2417

Fax: (65) 6791-7180

Email: liangma@ntu.edu.sg

DOES SUPER-DEPARTMENT REFORM IMPROVE PUBLIC SERVICE PERFORMANCE IN CHINA?

Abstract

Echoing the global public management reform movement, China's authorities advocated "super-department" reform (SDR) to curb interdepartmental conflict and administrative inefficiency. However, the related performance consequences have not been empirically investigated. We test the reform's effects on citizen satisfaction with public services through a natural experiment involving 25 counties in Guangdong Province (2009-2012) and the difference-in-differences method. The results show that the reform has improved public service performance, but its effects are marginal and unsustainable. We discuss the theoretical contributions and policy implications of the findings, and identify future research avenues.

INTRODUCTION

The past three decades have witnessed the global movement of pubic management reform (Pollitt and Bouckeart, 2011). "New Public Management" (NPM) has been initiated to improve government performance and regain public legitimacy, and most of its components (e.g., privatization, competition and deregulation) have been adopted around the world (Kettl, 2005). New emerging models like "whole-of-government" initiatives (Christensen and Lægreid, 2007), Post-NPM (Dunleavy *et al.*, 2006) and New Public Governance (Osborne, 2010) have also been introduced to mitigate NPM pathology. Different from individual instruments and continuous improvement, these reforms are usually radical, comprehensive and systematic. According to Christensen (2013, p. 171), "Compound administrative reforms are multidimensional and represent 'mixed' orders and combinations of competing, inconsistent, and contradictory organizing principles and structures, but also cultures that co-exist and balance interests, values, and claims to power." Comprehensive reform packages like the National Performance Review launched by U.S. Vice-President Al Gore in the 1990s and the Public Service Improvement Strategy initiated by the U.K. Labor Government in the 2000s require governments to take a "big bang" approach in implementing a battery of reform agendas (Walker and Boyne, 2006).

As these reforms are costly in terms of the financial input, manpower, managerial attention and organizational turbulence involved, it is imperative to gauge their actual

consequences. Different from single reform strategies like privatization and downsizing, the outcomes and influences of comprehensive reform strategies are usually difficult to disengage and estimate. The actual consequences of these reforms may be unintended or unanticipated (Perri 6, 2010), and whether and how they affect public service performance (PSP) is not yet fully understood (Walker and Boyne, 2006; Aiely, 2011). As suggested by Moynihan (2013, p. 500), "[O]ne area that is less tractable to some of the improved application of research methods is whole-of-government performance reforms. We usually lack a randomized control or variation in the rates of adoption that might provide inferences about the effects of adoption of such reforms."

There have been many endeavors to evaluate the effect of public management reform, either by assessing the reform as a whole package (Boyne *et al.*, 2003; Walker and Boyne, 2006; Gormley, 1989) or evaluating specific elements such as performance management (Boyne *et al.*, 2010; Grosso and Van Ryzin, 2012; Andrews and Van de Walle, 2013). One literature review of 519 empirical studies in European countries arrived at mixed findings in terms of the effect of public management reform on performance, and contextual factors have been found to significantly confound the results (Pollitt and Dan, 2013). Most of the research has involved case studies or cross-sectional surveys, impeding any causal inference about the performance effects. Furthermore, evidence from contexts unrelated to Western countries is rare. We add to the ongoing research stream by evaluating the performance consequences of comprehensive reform in China, which has been

under-examined in the literature (Xue and Liou, 2012; Christensen *et al.*, 2012).

The reform agendas and processes in China and the West have different foundations (Urio, 2012), and it is meaningful to examine their potentially different influences. How to sustain governance in a "fragmented authoritarianism" has become a significant challenge facing Chinese governments at all tiers (Lieberthal and Oksenberg, 1988). The central government advocated "super-department" or "mega-ministry" reform as the key means of curbing interdepartmental conflict and administrative inefficiency in the two most recent waves of institutional restructuring (2008 and 2013) (Yeo, 2009; Dong *et al.*, 2010). Local governments have experimented with different modes of super-department reform (SDR), and it is pertinent to investigate their performance consequences. We collected multisource data from 120 counties in southern Guangdong from 2009 to 2012, and implemented a "difference-in-differences" (DID) approach to examine whether and how SDR contributes to PSP. Our findings show that SDR has improved PSP, but its effects are marginal and unsustainable. The evidence reported in this paper contributes to the literature on public management reform and public service improvement, and governments in other countries and regions may learn from China's experience with administrative reform to deliver quality public service to their constituents.

The remainder of the paper is structured as follows. We begin by introducing the context of the reform and its key components, and then offer theoretical arguments and hypotheses. We then present the data and methods used in our empirical analyses, and report our key findings. Finally, we discuss the theoretical contributions and

policy implications of the findings, and conclude the paper by identifying the study's limitations and making suggestions for future research.

CONTEXT

The evolution of governance architecture in China

China adopts a unitary system of five administrative tiers, including central, provincial, prefectural, county-level and township governments (Lieberthal, 2004). Below the central government (the State Council and its subordinate ministries) are 33 provincial units, including 22 provinces (e.g., Guangdong), 4 municipalities (e.g., Beijing), 5 autonomous regions (e.g., Tibet) and 2 special administrative regions (Hong Kong and Macau). The provinces and autonomous regions directly supervise more than 300 prefectures and prefecture-level cities, and the central government has specifically designated 15 of them (e.g., Shenzhen) with sub-provincial status with extra administrative authority. Almost 3,000 counties, county-level cities and districts are controlled by municipalities, prefectures and prefecture-level cities, and over 40,000 towns, townships and sub-districts are at the bottom of the hierarchy.

Authoritarian China is solely ruled by the Chinese Communist Party (CCP), and other democratic parties and nonpartisan politicians play symbolic roles under the Chinese People's Political Consultative Conference (CPPCC). Deputies are periodically elected to people's congresses indirectly, but are usually considered rubber stampers without authentic power. A party-state regime organizes the authority, and the party's committee branches are juxtaposed with government agencies at every

level. The CCP committees at every level are the *de facto* authorities, and the corresponding governments are their administrative machines. The heads of authorities at each level (e.g., the party secretaries and governors) are appointed by upper-tier party committees, and sector chiefs are usually appointed by party committees at the same level.

The unitary five-tier regime makes China's hierarchical complexity hard to manage, particularly for local authorities in developed regions. The top-down chain of command is characterized by inertia and inefficiency in terms of vertical policy communication and implementation, and duplicated and overlapped functions at every level hinder horizontal coordination and collaboration. For instance, in the current prefecture-managing-county system, prefecture-level cities usually grab fiscal transfers and economic benefits from counties, and county governments have the disproportional fiscal and administrative authority to engineer economic growth and social development (Wang and Ma, 2014). Province-managing-county (PMC) reform has been experimented as an effective means of fiscal decentralization in China since the 1990s (Wang *et al.*, 2012). Some provinces such as Zhejiang and Guangdong have delegated substantial fiscal and administrative authority to county governments.

A typical county government is composed of roughly 100 agencies, making the bureaucracy unmanageable and coordination painful. Some of their functions are duplicated (e.g., the CCP Discipline Inspection Commission and Supervision Bureau), and their functional overlapping hinders policy coordination. The agencies' entrenched interests impede policy coordination and functional integration, and

efficiency and effectiveness are also largely sacrificed. Administrative reform has been emphasized as a key means of activating regime resilience in transitional China over the past three decades (Xue and Liou, 2012). Two primary themes of the several waves of reform are structural reorganization and downsizing, and SDR was initiated by the central government in the fifth and sixth rounds of administrative reform in 2008 and 2010, respectively (Yeo, 2009; Dong *et al.*, 2010).

Pilot of super-department reform in Shunde District

Although SDR was implemented in a top-down manner, some innovative local governments had already experimented with various alternatives to trigger national reform. Due to preferential policies and economic openness, local governments in Guangdong have served as pioneers in policy experiment and administrative reform. Located in the abundant Pearl River Delta region and bounded by Hong Kong to the north, Guangdong is among the most developed regions in China. The provincial government governs 21 prefecture-level cities, with Guangzhou and Shenzhen serving as sub-provincial cities. There are 121 county authorities in Guangdong, including 44 counties, 54 districts and 23 county-level cities.

Due to its rapid economic development, Shunde District in Foshan has been a remarkable pioneer in policy experiment and innovation (Li, 2004). On August 24, 2009, the Guangdong authorities chose it to pilot SDR. Guided by experts from the Chinese Academy of Governance and the experiences of Hong Kong and Singapore, Shunde merged agencies with similar functions and streamlined inter-functional

coordination. It dramatically amalgamated its 41 party-state organs into 16 super-departments, and this radical initiation attracted much attention and reports from across the nation.

An innovation of the Shunde reform is the consolidation of party and state agencies to overcome organizational overlapping and redundancy. The past decades have witnessed the separation of party functions from those of the government, but Shunde recombined party and government functions to form new super-departments. Shunde restructured its organs into several super-departments according to three main functions or missions: administrative affairs, economic adjustment and market regulation, and social governance and public service delivery. All agencies were merged into multiple umbrella institutions such as mega-security, mega-regulation, mega-planning, mega-construction, mega-economy and so forth. For instance, a new Bureau for Market Safety Regulation replaced several previous agencies such as the Bureau of Agriculture, Industrial and Commercial Bureau, Food and Drug Administration, Bureau of Safety Regulation, Bureau of Quality Regulation and Bureau of Health and Bureau of Economic Trading. District chiefs and party secretaries rather than past agency heads were put in charge of these agencies, and their engagement in operations facilitated decision making and coordination.

With the strong support of the provincial authority, the reform in Shunde has proceeded successfully via learn-by-doing and trial-and-error practices. The reform has streamlined business processes, improved administrative efficiency and alleviated the administrative burdens of citizens and businesses. One survey showed that most of

the responding citizens and businesses in Shunde were satisfied with the new mode of administration (Huang and Chen, 2012). Guangdong championed Shunde as a model and extended its experience to 25 selected counties in late 2010 to trigger learning and replication, creating a natural experiment that allows us to assess its consequences.

The diffusion of super-department reform in Guangdong Province

Governments in China commonly conduct policy experiments as effective instruments to drive local innovations in national policies (Florini *et al.*, 2012; Wu *et al.*, 2013). Local government officials are keen in policy innovation since it is a strong performance signal to attract their superiors' attention and increase the possibility of career advancement. Upper-tier governments usually generalize local experiments across their jurisdictions to prompt best practices (Heilmann, 2008). Shunde initiated SDR in 2009, and this experiment accumulated valuable experience for its peers to emulate and learn. Guangdong deemed the Shunde experiment replicable and generalizable, and has used it to encourage government innovations in other counties. A province-wide conference was held on April 29, 2010 to generalize the Shunde experiment, and Foshan immediately released its reform plan the next day to push forward a replication of the experiment in its remaining districts. On November 25, 2010, Guangdong formally promulgated a guideline and selected 25 pilot counties to popularize the Shunde experiment (see Figure 1 for their geographical distribution). Four counties in Foshan (excluding Shunde) and Yangjiang, respectively; two counties in Jiangmen and Yunfu, respectively; and one county in other

prefecture-level cities, respectively (excluding Shenzhen, Zhuhai, Dongguan and Zhongshan) were chosen as the pilot counties. As special economic zones, Shenzhen and Zhuhai have the discretion to specify their reform strategies, and were not mandated to reform. Dongguan and Zhongshan were upgraded from county-level cities, and they directly manage towns, townships and sub-districts without subordinate counties to reform.

Insert Figure 1 about here.

The reform aimed to empower county-level authorities with more administrative discretion, to merge government departments and flatten public organizations, and to establish a new model of a high-performance county management system. Guangdong pinpointed 11 proposals to amalgamate party committees and government agencies with similar functions[1], and set the maximum amount of agencies for pilot counties.[2] Although Guangdong suggested several consolidation strategies, the pilot counties were encouraged to learn from the Shunde experiment and left with ample discretion to boldly innovate according to local conditions. The pilot counties were required to complete their reforms by the first quarter of 2011, and all counties conceived and implemented their reform schemes in short order. In July 2012, Guangdong asked all prefecture-level and above cities but Zhuhai, Foshan, Yangjiang, Dongguan, Zhongshan and Shenzhen to choose at least two extra counties to experiment SDR and aimed to popularize it across the whole province in 2013.

THEORY

SDR is a mixture of NPM and "whole-of-government" approach in that it introduces a package solution to administrative inefficiency and interdepartmental conflicts, and almost all of its components have the potential to alter the government management system and improve PSP. We argue that SDR may positively influence PSP in several ways.

Consolidation and PSP

One core component of SDR is government consolidation and amalgamation, which is the rhythm of almost every wave of public administration reform (Seidman, 1997; Davis *et al.*, 1999). Comprehensive administrative reorganization is "a characteristic feature of twentieth-century bureaucratic and political life" (March and Olsen, 1983, p. 281). The amalgamation of government agencies with similar functions may help to cut costs and internalize inter-functional conflicts, thereby enhancing work efficiency and ameliorating citizens' public service experience (Christensen and Lægreid, 2007). The consolidation of fragmented policy functions helps a government streamline internal processes and attenuate cross-functional conflicts. A reduction in the number of subordinate agencies enables county governments to coordinate interdepartmental relations and concentrate their attention on policy priorities.

Evidence from federal, state and local reorganizations in many countries, however, has produced little support for its success (March and Olsen, 1983). Analysis of the structural reorganization of U.S. state government agencies revealed that it rarely

decreased employment and expenditure (Meier, 1980). Ministers and bureaucrats in Australia were found to have distinct preferences for the benefits and costs of policy coordination, suggesting that it is more complicated than expected (Craswell and Davis, 1994).

If amalgamation has not worked in other circumstances, may it also fail in the context of Guangdong? SDR is a typical organizational innovation that includes the adoption of restructuring, coordination and collaboration practices. Innovation usually involves a risk of failure (Damanpour *et al.*, 2009), but Shunde's experience may help the pilot counties lower their innovation costs and risks while improving their innovation performance. SDR may reduce bureaucratic redundancy and remove red tape, which in turn facilitate public service delivery and improve PSP. SDR echoes two elements of market orientation: customer orientation and inter-functional coordination, which have been found to enhance British citizens' satisfaction with public services (Walker *et al.*, 2011). Functional transformation and service orientation drive government agencies to highlight the delivery of quality public service to citizens rather than solely implementing mandates, which may enhance PSP (Xue and Liou, 2012). Agency consolidation strengthens governments' capacity in strategic planning and management, which is found to boost PSP (Andrews and Van de Walle, 2013).

Decentralization and PSP[3]

SDR is not merely about horizontal consolidation but also involves vertical

empowerment and administrative decentralization. In other words, SDR is also a form of decentralization reform (Wang and Ma, 2014). Prefecture-level cities and provincial departments were required to support the reform and delegate administrative authority to the pilot counties (e.g., administrative approvals of fixed-asset investment and social security), enabling them to customize developmental strategies to local conditions. For instance, previously citizens and businesses usually have to wait for several weeks to receive specific approvals, since county-level agencies must obtain permissions from their prefecture-level superiors. Thanks to SDR, prefecture-level governments may grant more discretion to county-level governments to directly approve some services, which may substantially shorten citizens' waiting time and improve their satisfaction with PSP (Zhou and Yang, 2011).

The pilot counties were mandated to transform their functions to focus on public service delivery (e.g., public education, health care, employment and social security), set up public service platforms or one-stop service centers, and harness information technology (IT) to lower administrative costs and improve work efficiency. They were also encouraged to strengthen public accountability, performance measurement, information disclosure, citizen feedback and cross-departmental coordination. The top-down mandate and political support from top-tier leaders also provided strong incentives and a beneficial climate for backing county governments to implement the reform agenda and improve their PSP (Huang and Chen, 2012). Government agencies take much more responsibility in designing service plans and implementing service

delivery. They become much more in charge of public services, rather than passively respond to commands from their upper-tier superiors. SDR helps government agencies to realign their strategies around core service functions (Walker *et al.*, 2011), and pilot counties enjoy much more autonomy and discretion to design and deliver public services, which may improve PSP (Wang *et al.*, 2012). The consolidation of Party committees and government agencies may delegate more discretion to the new super departments, which in turn helps bureaucratic agencies to perform more efficiently.

Does super-department reform matter?

The preceding arguments suggest that SDR as one form of administrative reform and organizational innovation may contribute to PSP in pilot counties, and we expect SDR to be positively related to PSP. A survey in Shunde shows that SDR has improved citizen satisfaction with PSP (Huang and Chen, 2012). The performance assessment of SDR initiated by the Guangdong Commission Office for Public Sector Reform in late 2011 found that SDR had generally achieved its goals, but the counties varied in terms of the depth and width of their reform. Over 30 percent of government agencies were cut, and more than 90 percent of administrative approval items were cancelled or merged in pilot counties (Zhou and Yang, 2011).

SDR may affect different dimensions of PSP. PSP is a complicated and multidimensional construct, and SDR may have distinct effects on its various dimensions (e.g., efficiency, effectiveness, responsiveness and equity). For instance,

although NPM may increase efficiency and effectiveness, equity and accountability may backfire (Bozeman, 2000). Evidence from British local governments reveals that efficiency, responsiveness, equity and effectiveness were affected by different NPM practices (e.g., public-private relationships and entrepreneurial strategic orientation) (Andrews and Van de Walle, 2013). We expect that SDR will positively affect each performance dimension (e.g., efficiency and equity), but the magnitude of the effects may vary.

METHODS

The unit of analysis and data sources

We conduct an empirical analysis using data from the full population of 121 county governments in Guangdong. Shunde as a pilot county started its initiation in 2009, over one year ahead of the other 25 experiment counties, and we exclude Shunde from the sample to eliminate its first-mover advantage. We use the 2009-2012 data to investigate the performance impact of the 2010 reform. The number of observations used in the analysis is 468.[4]

The data were taken from multiple sources to mitigate common method bias (Meier and O'Toole, 2013). We use the formal government document related to the reform to identify the list of the pilot counties. We derived the PSP data from an independent research program at South China University of Technology (SCUT). We collected the control variable data from the 2010 Population Census (National Bureau of Statistics, 2012).

Model specifications

The DID method is the standard approach used in natural experiment studies (Wang *et al.*, 2012; Grosso and Van Ryzin, 2012). The following equation is used in DID estimates:

$$P_{i,t} = \alpha + \beta_1\, S_{i,t} + \beta_2\, E_{i,t} + \beta_3\, S_{i,t} \times E_{i,t} + \beta_4\, C_{i,t} + \varepsilon. \qquad \text{(Equation 1)}$$

In the equation, i and t refer to county i and year t, respectively. P denotes PSP measures. We create two dummies and use their interaction terms to address the effects of SDR. S refers to the pilot counties, and we use code 1 for pilot counties (experiment group) and 0 for other counties (control group). E refers to the years during and after the experiment, using code 1 for the post-experiment years (2011 and onward) and 0 for the pre-experiment years (2010 and previous years). C refers to the control variables, and ε is the error term. Slope β_1 estimates the difference in P between the pilot and control counties, and slope β_2 estimates the difference in P between the pre- and post-experiment periods. The regression coefficient of the interaction term, β_3, refers to the DID or net effect of the treatment (experiment group after the treatment).

The dataset used in this study requires panel data models (Zhu, 2013). We use a random-effects model to test the hypothesis, and report robust standard errors to mitigate heterogeneity and autocorrelation concerns. The random-effects model is preferable to the fixed-effects model in this analysis for two reasons. First, some of our independent variables are time invariant and will be automatically dropped from the fixed-effects model. Second, the Hausman test suggests that a random-effects

16

model is more efficient for our model estimates.

Different from a laboratory experiment design, policy reform as a quasi-experiment design is typically non-randomly distributed, and the manipulations are usually beyond the control of the researchers. The pilot counties were not randomly selected to experiment SDR, and our estimates of the reform effect may be prone to sample selection bias (Konisky and Reenock, 2013). Guangdong selected the pilot counties primarily based on their representativeness in each corresponding prefecture-level city, and the selection criteria were not explicitly disclosed. We compare PSP of the pilot counties with that of the control counties in 2010, and find no statistically significant differences. In other words, the pilot counties were not chosen based on their PSP, and the threat of selection bias is negligible. Nevertheless, to mitigate the potential selection bias, we control for the demographic and socioeconomic variables.

Dependent variables

PSP as a multidimensional construct is difficult to gauge, particularly because multiple stakeholders perceive it differently. As the target customers of public management reform, citizens are eligible core information sources of PSP, and citizen surveys have commonly been used in studies (Walker *et al.*, 2011; Charbonneau and Van Ryzin, 2012). We use a unique citizen survey dataset to measure PSP. Starting in 2007, the SCUT team has annually evaluated government performance of 21 prefectural-level cities and 121 counties in Guangdong (Zheng, 2013), and we used their citizen survey data to measure PSP.

The SCUT team used a stratified quota sampling method in citizen survey. The sampling quotas were distributed in each county according to population size and demographical representativeness (e.g., gender and age), and the total number of respondents ranged from 23,000 to 28,000.[5] In each county, the residents (aged 18-70) were randomly sampled and interviewed in person.[6] The sample was generally representative of the population in gender, age, occupation and income, but relatively higher educated than the overall population due to the higher percentage of respondents in urban areas. The surveys were commonly conducted after the Spring Festival (with the Lunar New Year usually falling in February) and used to gauge PSP of the prior year. We use the latest data collected at the beginning of 2010-2013 (correspondent to the 2009-2012 period) to investigate the performance effect of the 2010 reform.

The SCUT team surveyed six PSP items: public service benevolence, public service efficiency, integrity, public service equity, government transparency and overall government performance. The surveys used 11-point Likert-type scales, ranging from 0 (strongly dissatisfied) to 10 (strongly satisfied). The six measures are highly correlated with one another[7], and we combine them to form a comprehensive PSP index (Favero and Meier, 2013). The factor analysis results are reported in Table 1 and a single factor emerges, explaining 80.17 percent of the total variance. The six measures are averaged and their mean is used to gauge PSP. We also use each measure as dependent variables to elicit the distinct effects of SDR on each dimension of PSP.

Insert Table 1 about here.

Control variables

As many environmental and organizational attributes may affect PSP, we control the following variables in line with the literature and in consideration of the available data. First, organizational performance changes incrementally to a large extent, and it is imperative to control for past performance (Walker *et al.*, 2011). We include the previous year's performance, and the estimates can be interpreted as the effects of the reform on performance improvement. Second, PSP can be enhanced by scale economy, and we control for population size to account for its effects.

Social deprivation and demographic diversity are found to be negatively related to PSP (Andrews and Van de Walle, 2013). Populations with a high degree of urbanization are more likely to enjoy quality public service, and we control for urbanization rate. We control for the average years of education, as counties with richer human capital may be pressed to provide better quality public service. Counties with higher proportions of minorities (non-Han nationalities) are usually less developed, and the people in those counties may be less likely to enjoy good quality public service. Guangdong has been both driven and plagued by the flux of rural migrant workers from other provinces (e.g., Sichuan), and public services are commonly depressed by the disproportional increase in users (Ngok, 2012). We use

the portion of immigrants to gauge the migration effect, and expect it to be negatively related to PSP (Andrews *et al.*, 2009). We also include unemployment rate, as it reflects economic wealth and may negatively affect PSP. Furthermore, China already encountered the healthcare challenge of population aging, and we include the portion of the population aged 65 and above to take it into account. All of the control variables except for aging are not normally distributed, and we use their logarithmic terms.

Finally, counties, county-level cities and districts as distinct administrative units have different levels of urbanization and infrastructure, and we create two dummies for county-level cities and districts, using counties as the reference group. Mountainous counties lag behind their peers in the flatlands in socioeconomic development, and we create a dummy variable to consider its effect. Counties nested in the same prefecture-level cities may be prone to similar environmental shocks and policy mandates, and we cluster standard errors at prefecture level to control for such unobserved influence.[8]

RESULTS

The descriptive statistics of the key variables are reported in Table 2. The dependent variables, PSP index and its six measures vary substantially across the counties. The means of PSP measures range from 4.77 (integrity) to 5.44 (overall). In general, local residents are moderately satisfied with PSP. Some of the counties outperform their peers with scores of nearly 9.00, and others perform disappointingly with scores as

low as 3.00. The counties also differ notably in terms of their sizes and demographics.

Insert Table 2 about here.

The pre-post comparison between the control and experiment groups (see Table 3) suggests that there are remarkable differences in all PSP dimensions before and after SDR. Although both the control and pilot counties exhibit improved PSP throughout the study period (except for overall performance in the control group), the pilot counties received greater increases in all PSP dimensions than their counterparts, particularly in public service equity (0.261) and government transparency (0.159) (see the last column). The magnitude of the treat effects, however, is rather weak in an 11-point Likert-type scale.

Insert Table 3 about here.

We begin by predicting the PSP index using all of the independent variables (Model 1 in Table 4), and then estimate all of the PSP measures (Model 2-7). The average value of variance inflation factor (VIF) for all of the models is approximately 2.90, implying that there is no severe multicollinearity threat. The models are all significantly different from 0 (see Wald χ^2). The R^2 value ranges from 0.09 to 0.21, and we can say that our independent variables moderately explain the variance in PSP.

The results suggest that SDR has a positive albeit statistically insignificant effect on PSP index, as denoted by the coefficient of the interaction term ($\beta=0.136$, $p>0.10$). The pilot counties improved their PSP index 0.136 point (or 1.36 percent) higher than their non-reforming peers, but the effect is not large enough to pass a significance test. The results reported in Model 2-7 show that all of the regression coefficients of the interaction term are positive but not statistically significant at a confidence level of 90 percent. Most of the regression coefficients of the reform county dummy are negative and those of the reform year dummy are positive, implying that PSP in the pilot counties was relatively poorer before the reform but improved after the reform.

We run alternative models as robustness tests, and the primary results are substantially unchanged. Because the dependent variables are highly correlated, the residuals of the six models (Model 2-7) may be interdependent, violating the independence assumption of the ordinary least square (OLS) regression. We rerun the models by seemingly unrelated regression (SUR) to simultaneously estimate the performance effects of SDR (Martin and Smith, 2005).[9] All of the regression coefficients of the interaction term are positive, and only those of public service equity ($\beta=0.282$, $p<0.05$) and government transparency ($\beta=0.195$, $p<0.05$) are statistically significant. The simultaneous estimates of the reform effects suggest that we can reject the null hypothesis that the coefficient estimate on the interaction term jointly across six PSP measures is equal to zero at a 99% confidence level ($\chi^2=22.7$,

p<0.01). In other words, the performance consequences of SDR are positive and statistically significant. In sum, SDR has improved PSP but the effect is marginal and attenuate.

As expected, our control variables also have notable effects on PSP. The lagged dependent variables are positive and statistically significant in all but two of the models (Models 4 and 5). The coefficients of population size and urbanization rate show inconsistent results across the different models, while average education level and migration rate are positively linked with PSP. As expected, minority rate, unemployment rate and aging are negative in most of the models. County-level city and district dummies are not statistically significant, whereas mountainous counties perform better than their peers in the flatlands.

DISCUSSION

Although SDR and other public management reform packages have been advocated as key to improving PSP and retaining government trust, the empirical evidence for this proposition remains scarce. It is meaningful and interesting to investigate the performance implications of these reforms. We examine the performance consequences of public management reform in this paper, which focuses on SDR in China's Guangdong counties. We use a natural experiment research design to conduct one of the first empirical studies related to the performance effects of SDR. The results suggest that SDR positively affects PSP, but its effects are attenuate.

Guangdong granted county governments with ample autonomy to adopt appropriate

reform measures according to local conditions. Strong top-down support also played a crucial role in safeguarding the implementation of the radical reform agenda, particularly in such a typical hierarchical system. However, the actual implementation of SDR agenda reveals notable shortages and barriers, originating from either the inertia of the current regime or the reform agenda itself. The marginal effect of SDR on PSP can be partially attributed to the complexities of vertical and horizontal coordination involved in the reform (March and Olsen, 1983; Meier, 1980; Craswell and Davis, 1994).

The key obstacle to reform is determining how to reconcile the conflict between super-departments and corresponding upper-tier authorities. As provincial and municipal governments have not yet gone through correspondent reforms, a typical county super-department must usually report to several upper-tier agencies. On the other hand, upper-tier agencies cannot easily figure out which super-department is in charge. The dysfunction of vertical communication even encourages upper-tier agencies to avoid contacting pilot counties, largely shrinking their fiscal transfer and other resources and opportunities. Agencies were merged, but the positions of their former heads were still conserved to smooth the radical reform process, largely complicating the coordination within the new super-departments. Agencies directly controlled by provincial or prefecture-level governments (e.g., national tax bureaus) often receive different rules and remunerations from county government agencies (e.g., local tax bureau), and how to reconcile their conflicts of interest under the same super-department umbrella is a tough problem facing reformers.

The cost, either direct or indirect, involved in SDR should be taken into account. The process of restructuring and reshuffling takes time and is costly, and the risk of downsizing demoralizes the civil service (Seidman, 1997; Davis *et al.*, 1999). If the benefits of SDR could not offset these costs, then we should be cautious about its promise in performance improvement. Although the provincial authority encouraged service orientation, the establishment and operation of a one-stop service center usually takes a long time. The cultivation of a service-oriented culture requires a thorough transformation of the current institutional system, which is far beyond the reform agenda. The public usually stereotypes PSP, and this is not easy to change in the short term. These reasons may partially explain why PSP were not substantially improved in the pilot counties.

Public management reform works differently for various PSP dimensions (Bozeman, 2000; Andrews and Van de Walle, 2013). Some dimensions improve and others remain constant or even deteriorate, making PSP remains unchanged overall or improves subtly. A key merit of the Guangdong reform was its comprehensiveness, as neither of its components could individually achieve an increase in PSP. The results show that SDR has been more profound in some performance dimensions (e.g., public service equity and government transparency) than others. The national government information disclosure regulations were enacted in 2007 and took effect in 2008. However, their implementation lagged due to local governments' reluctance and incapacity to become more transparent. Government openness was highlighted in SDR and was largely prompted through the use of IT, increasing residents' positive

perception of government transparency. As for equity, reform may enhance citizens' access to basic public services, and they may in turn perceive higher levels of equity and respect.

The evidence reported in this paper contributes to the literature related to public management reform and public service improvement. First, our results suggest that comprehensive public management reform matters to PSP, but its magnitude is marginal. Previous evidence of the performance consequences of public management reform has primarily relied on surveys and case studies (Walker and Boyne, 2006; Boyne *et al.*, 2003). We conduct a natural experiment to identify the performance effect of public management reform, and the DID methods that we use can be replicated and extended to other reforms and contexts to determine their effects. Second, we provide new evidence for the literature related to the innovation-performance nexus. In line with previous studies involving other contexts (Damanpour *et al.*, 2009), our results show that public management innovation is positively associated with organizational performance. This finding further enriches the literature by demonstrating with evidence that public management innovation matters for organizational performance. Finally, our findings imply that whole-of-government reform through the super-department approach works weakly in the distinct Chinese context (Christensen and Lægreid, 2007). Our findings thus contribute to the literature by providing new evidence of public management reform in China, which has been relatively under-examined (Wu *et al.*, 2013).

Our results reveal that the reform has marginally improved PSP as perceived by

citizens, and governments should thus continue their endeavors to restructure bureaucratic architecture to foster citizen-centered business processes and organizational routines. Our findings suggest that local governments in other regions in China may learn from the experiences of Shunde and Guangdong to fine-tune their reform endeavors. However, the results do not indicate that other governments should necessarily follow the Shunde model. Although the Chinese context is relatively unique, governments in other nations and regions, particularly those in less-developed and transition economies, can nevertheless learn from its experience in advancing public management reform and improving PSP.

The findings of our natural experiment are robust and relevant to the literature on public management reform and public service improvement. We argue that several mechanisms may help SDR improve PSP, but we do not investigate which mechanisms actually do so. DID permits us only to investigate whether the reform has worked, and we cannot identify the relative effects of each reform element. Future research could examine which reform components are most crucial in improving PSP (Walker and Boyne, 2006). Two key elements of package solution reform are comprehensiveness and radicalness, and the compatibility and complementarity of each reform component merits future research (Christensen and Lægreid, 2007).

Second, we explore the direct effect of SDR on PSP, and how the relationship conditions on other environmental and organizational attributes should be investigated further, if such investigation is deserved. Context is important in determining the influence of public management reform on performance (Pollitt and Dan, 2013). SDR

challenged the power balance and entrenched interest of upper-tier authorities, and how it is maintained during political conflict and pressure deserves future exploration (Wang, 2010). Leadership will and support are key predictors of reform effectiveness. Wang Yang, then Guangdong Party Secretary, is the core advocator of SDR, and his career advancement may be either a positive or a negative signal for its future, depending on his advocacies and the propensities of his successors.[10] It would be interesting to analyze the moderating effects of leadership support in such a radical reform process.

Finally, due to data limitations, we analyze only the short-term effect of the reform, and future studies could examine its long-term effects to determine whether they are sustainable. Administrative reform in China blows hot and cold. The reform may downsize government agencies and workforces, but several years later it may return to the status quo and require another round of reform. Whether the performance consequences of SDR are sustainable deserves further exploration. As a multidimensional construct perceived by various stakeholders, another research avenue involves eliciting the distinct effects of SDR on the perceptions of other stakeholders (e.g., businesses and administrators) (Walker *et al.*, 2011). As public organizations pursue multiple and often conflicting goals, the trade-off dilemma of public management reform also deserves further exploration (Moynihan, 2013).

ACKNOWLEDGEMENTS

Asian Group for Public Administration (AGPA), "Local and Regional Cooperation and Public Governance," in Singapore on September 26-28, 2013. I would like to thank the conference participants for their helpful comments and suggestions. The informative comments from the anonymous reviewers also helped a great deal. I am grateful to Huping Shang, Ruoxian Sun and Fanghui Zheng for the data collection work. Any remaining errors are my own.

NOTES

[1] For instance, a new Bureau of Urban Construction and Water Affairs was recommended to replace three former agencies (the Bureau of Urban Construction, Bureau of Water Supply and Bureau of Urban Management) to enforce laws related to urban construction, real estate administration, public utilities, the water supply and urban management. Former bureaus like the Bureau of Work Safety, Supervision Bureau and Bureau of Culture and Tourism Administration were advised to form a new Bureau of Market and Safety Inspection in charge of market regulation and law enforcement.

[2] County-level cities were instructed to merge their organizations to include less than 30 agencies. Large, medium and small counties had to keep fewer than 28, 24 and 20 agencies, respectively. Districts and districts created out of county-level cities could maintain 22 and 28 agencies at most, respectively.

[3] We thank one anonymous reviewer for suggesting us to discuss SDR's decentralization effects.

[4] The number of observations should be 480 (120×4), but the data on PSP of 12 counties in 2012 are missed. We impute the missing data by using the aggregated prefecture-level values minus the mean values of other counties in the same prefecture-level city weighted by their sample sizes. The model estimates of the imputed dataset are substantially similar with that of the original dataset (N=468), and we report the results of the original dataset.

[5] The samples included 100 respondents for populations under 300,000; 150 for populations of 300,000-500,000; 200 for populations of 500,000-800,000; 250 for populations of 800,000-1 million; 300 for populations of 1-1.5 million; 400 for populations of 1.5-2 million and 500 for populations higher than 3 million.

[6] About 10 percent of the respondents living in remote regions were randomly sampled and interviewed by telephone. The test showed no significant differences between the responses collected in person versus on the telephone (Zheng, 2013).

[7] The simple correlation coefficients are all larger than 0.70.

[8] We thank one anonymous reviewer for suggesting us to do so.

[9] The Breusch-Pagan test of independence (χ^2=4143.985, p<0.01) suggests that SUR is appropriate to estimate the models. We use the STATA command XTSUR to estimate the models in unbalanced panel data with random effect. The results are excluded to save space, but are available from the author upon request.

[10] Wang left Guangdong in December 2012 and was promoted to be Deputy Prime Minister of China in 2013.

REFERENCES

6, P. (2010) 'When Forethought and Outturn Part: Types of Unanticipated and Unintended Consequences' in H. Margetts, P. 6 and C. Hood (Eds.), *Paradoxes of Modernization: Unintended Consequences of Public Policy Reform.* Oxford: Oxford University Press, 44-60.

Aiely, G. (2011) Why People (Dis) Like the Public Service: Citizen Perception of the Public Service and the Npm Doctrine. *Politics & Policy*, 39:6, pp. 997-1019.

Andrews, R., *et al.* (2009) EU Accession and Public Service Performance. *Policy & Politics*, 37, pp. 19-37.

Andrews, R. and Van de Walle, S. (2013) New Public Management and Citizens' Perceptions of Local Service Efficiency, Responsiveness, Equity and Effectiveness. *Public Management Review*, 15:5, pp. 762-83.

Boyne, G., *et al.* (2010) 'What if Public Management Reform Actually Works? The Paradoxical Success of Performance Management in English Local Government' in H. Margetts, P. 6 and C. Hood (Eds.), *Paradoxes of Modernization: Unintended Consequences of Public Policy Reform.* Oxford: Oxford University Press, 203-20.

Boyne, G. A., *et al.* (2003) *Evaluating Public Management Reforms: Principles and Practice*. Buckingham: Open University Press.

Bozeman, B. (2000) *Bureaucracy and Red Tape*. Upper Saddle River, NJ: Prentice-Hall.

Charbonneau, É. and Van Ryzin, G. G. (2012) Performance Measures and Parental Satisfaction with New York City Schools. *American Review of Public Administration*, 42:1, pp. 54-65.

Christensen, T. (2013) 'New Public Management and Beyond: The Hybridization of Public Sector Reforms' in G. S. Drori, M. A. Höllerer and P. Walgenbach (Eds.), *Organizations and International Management.* New York: Routledge, 161-74.

Christensen, T., *et al.* (2012) Imitating the West? Evidence on Administrative Reform from the Upper Echelons of Chinese Provincial Government. *Public Administration Review*, 72:6, pp. 798-806.

Christensen, T. and Lægreid, P. (2007) The Whole-of-Government Approach to Public Sector Reform. *Public Administration Review*, 67:6, pp. 1059-66.

Craswell, E. and Davis, G. (1994) The Search for Policy Coordination: Ministerial and Bureaucratic Perceptions of Agency Amalgamations in a Federal Parliamentary System. *Policy Studies Journal*, 22:1, pp. 59-73.

Damanpour, F., Walker, R. M. and Avellaneda, C. N. (2009) Combinative Effects of Innovation Types and Organizational Performance: A Longitudinal Study of Service Organizations. *Journal of Management Studies*, 46:4, pp. 650-75.

Davis, G., *et al.* (1999) What Drives Machinery of Government Change? Australia, Canada and the United Kingdom, 1950–1997. *Public Administration*, 77:1, pp. 7-50.

Dong, L., Christensen, T. and Painter, M. (2010) A Case Study of China's Administrative Reform: The Importation of the Super-Department. *The American Review of Public Administration*, 40:2, pp. 170-88.

Dunleavy, P., *et al.* (2006) New Public Management Is Dead—Long Live Digital-Era Governance. *Journal of Public Administration Research and Theory*, 16:3, pp. 467-94.

Favero, N. and Meier, K. J. (2013) Evaluating Urban Public Schools: Parents, Teachers, and State Assessments. *Public Administration Review*, 73:3, pp. 401-12.

Florini, A., Lai, H. and Tan, Y. (2012) *China Experiments: From Local Innovations to National Reform*. Washington, DC: Brookings Institution Press.

Gormley, W. T., Jr. (1989) *Taming the Bureaucracy: Muscles, Prayers, and Other Strategies*, Princeton, NJ: Princeton University Press.

Grosso, A. L. and Van Ryzin, G. G. (2012) Public Management Reform and Citizen Perceptions of the UK Health System. *International Review of Administrative Sciences*, 78:3, pp. 494-513.

Heilmann, S. (2008) From Local Experiments to National Policy: The Origins of China's Distinctive Policy Process. *China Journal*, 59, pp. 1-30.

Huang, D. and Chen, C. (2012) A Follow-up Investigation on the Effectiveness of Local Giant Department Reform: The Experience of Shunde District in Foshan City, Guangdong Province. *Journal of Public Administration*, 5:6, pp. 24-47.

Kettl, D. F. (2005) *The Global Public Management Revolution: A Report on the Transformation of Governance*. Washington: Brookings Institution Press.

Konisky, D. M. and Reenock, C. (2013) Case Selection in Public Management Research: Problems and Solutions. *Journal of Public Administration Research and Theory*, 23:2, pp. 361-93.

Li, L. C. (2004) The Prelude to Government Reform in China? The Big Sale in Shunde. *China Information*, 18:1, pp. 29-65.

Lieberthal, K. (2004) *Governing China: From Revolution through Reform*, 2nd ed. New York: W. W. Norton.

Lieberthal, K. and Oksenberg, M. (1988) *Policy Making in China: Leaders, Structures, and Processes*. Princeton, NJ: Princeton University Press.

March, J. G. and Olsen, J. P. (1983) Organizing Political Life: What Administrative

Reorganization Tells Us about Government. *The American Political Science Review*, 77:2, pp. 281-96.

Martin, S. and Smith, P. C. (2005) Multiple Public Service Performance Indicators: Toward an Integrated Statistical Approach. *Journal of Public Administration Research and Theory*, 15:4, pp. 599-613.

Meier, K. J. (1980) Executive Reorganization of Government: Impact on Employment and Expenditures. *American Journal of Political Science*, 24:3, pp. 396-412.

Meier, K. J. and O'Toole, L. J. (2013) Subjective Organizational Performance and Measurement Error: Common Source Bias and Spurious Relationships. *Journal of Public Administration Research and Theory*, 23:2, pp. 429-56.

Moynihan, D. P. (2013) Advancing the Empirical Study of Performance Management: What We Learned from the Program Assessment Rating Tool. *The American Review of Public Administration*, 43:5, pp. 499-517.

National Bureau of Statistics (2012) *Tabulation on the 2010 Population Census of the People's Republic of China by County*, Beijing: China Statistics Press.

Ngok, K. (2012) Serving Migrant Workers: A Challenging Public Service Issue in China. *Australian Journal of Public Administration*, 71:2, pp. 178-90.

Osborne, S. P. (Ed.) (2010) *The New Public Governance? Emerging Perspectives on the Theory and Practice of Public Governance*. London: Routledge.

Pollitt, C. and Bouckeart, G. (2011) *Public Management Reform: A Comparative Analysis - New Public Management, Governance, and the Neo-Weberian State*, 3rd ed. Oxford: Oxford University Press.

Pollitt, C. and Dan, S. (2013) Searching for Impacts in Performance-Oriented Management Reform: A Review of the European Literature. *Public Performance & Management Review,* 37:1 pp7-32.

Seidman, H. (1997) *Politics, Position, and Power: The Dynamics of Federal Organization*, 5th ed. New York: Oxford University Press.

Urio, P. (2012) *China, the West and the Myth of New Public Management: Neoliberalism and its Discontents*. New York: Routledge.

Walker, R. M. and Boyne, G. A. (2006) Public Management Reform and Organizational Performance: An Empirical Assessment of the UK Labour Government's Public Service Improvement Strategy. *Journal of Policy Analysis and Management*, 25:2, pp. 371-93.

Walker, R. M., *et al.* (2011) Market Orientation and Public Service Performance: New Public Management Gone Mad? *Public Administration Review*, 71:5, pp. 707-17.

Wang, Q. (2010) Administrative Reform in China: Past, Present, and Future. *Southeast Review of Asian Studies*, 32:1, pp. 100-19.

Wang, W., Zheng, X. and Zhao, Z. (2012) Fiscal Reform and Public Education

Spending: A Quasi-Natural Experiment of Fiscal Decentralization in China. *Publius: The Journal of Federalism*, 42:2, pp. 334-56.

Wang, Z. and Ma, L. (2014) Fiscal Decentralization in China: A Literature Review. *Annals of Economics and Finance*, 15:2, pp. 751-70.

Wu, J., Ma, L. and Yang, Y. (2013) Innovation in the Chinese Public Sector: Typology and Distribution. *Public Administration*, 91:2, pp. 347-65.

Xue, L. and Liou, K. T. (2012) Government Reform in China: Concepts and Reform Cases. *Review of Public Personnel Administration*, 32:2, pp. 115-33.

Yeo, Y. (2009) Remaking the Chinese State and the Nature of Economic Governance? The Early Appraisal of the 2008 'Super-Ministry' Reform. *Journal of Contemporary China*, 18:62, pp. 729-43.

Zheng, F. (2013) *2012 Red Book on Government Performance Evaluation in China*. Beijing: Xinhua Publisher.

Zhou, Z., and Yang, R. (2011) People are Satisfied with City/County Government Agency Reform. *South Daily*, 09-05, pp. A05.

Zhu, L. (2013) Panel Data Analysis in Public Administration: Substantive and Statistical Considerations. *Journal of Public Administration Research and Theory*, 23:2, pp. 395-428.

TABLES AND FIGURES

Figure 1: Pilot counties included in the Guangdong super-department reform

Table 1: Factor-analytic results for the public service satisfaction survey items

Dimension	Survey item	Factor loading
Public service benevolence	Are you satisfied with the service attitude of local government agencies over the past year?	0.932
Public service efficiency	Are you satisfied with the service efficiency of local government agencies over the past year?	0.924
Public service integrity	Are you satisfied with the honesty and cleanness of local government employees over the past year?	0.810
Public service equity	Are you satisfied with the justice and law enforcement of local government employees over the past year?	0.815
Government transparency	Are you satisfied with the openness of local government affairs over the past year?	0.906
Overall government performance	Are you satisfied with the local government's overall performance over the past year?	0.872
Eigenvalue		4.810

Notes: N = 468.The method is principle-component factor analysis without rotation. LR test: χ^2= 3550.75, p=0.0000. Cronbach's α=0.946.

Table 2: Descriptive statistics for the key variables

Variable	Mean	SD	Min	Max
PSP index	5.058	0.608	3.01	6.862
Public service benevolence	5.056	0.625	3.04	6.77
Public service efficiency	4.950	0.639	2.89	6.76
Public service integrity	4.772	0.818	2.74	8.79
Public service equity	5.057	0.740	3.04	8.95
Government transparency	5.074	0.632	3.12	7.03
Overall government performance	5.439	0.634	3.09	7.4
Reform county dummy	0.208	0.407	0	1
Reform year dummy	0.500	0.501	0	1
Interaction term (reform county × reform year)	0.104	0.306	0	1
Population size (10,000 residents)	75.429	59.443	6.03	449.928
Urbanization (%)	58.075	26.094	19.869	100
Education (years)	9.278	1.070	7.59	12.81
Minority (%)	2.150	7.412	0.001	59.92
Migration (%)	17.604	21.108	0.364	89.888
Unemployment (%)	2.533	1.233	0.76	8.002
Aging (%)	8.284	2.460	1.02	12.57
County-level city dummy	0.192	0.395	0	1
District dummy	0.442	0.499	0	1
Mountainous county dummy	0.417	0.495	0	1

Notes: N = 468.

Table 3: Pre-post comparison of PSP

Variable	Control group (N=95)		Experiment group (N=25)		Performance change		Experiment effect (DID)
	Pre-SDR	Post-SDR	Pre-SDR	Post-SDR	Control group	Experiment group	
PSP index	4.960	5.157	4.899	5.235	0.207	0.312	0.104
	(0.579)	(0.624)	(0.630)	(0.542)	(0.547)	(0.644)	
Public service benevolence	4.932	5.181	4.914	5.224	0.262	0.289	0.028
	(0.615)	(0.623)	(0.635)	(0.541)	(0.538)	(0.525)	
Public service efficiency	4.800	5.107	4.778	5.129	0.321	0.327	0.007
	(0.590)	(0.648)	(0.644)	(0.609)	(0.548)	(0.589)	
Public service integrity	4.583	4.968	4.535	5.028	0.390	0.458	0.068
	(0.791)	(0.770)	(0.977)	(0.679)	(0.736)	(0.948)	
Public service equity	5.008	5.131	4.808	5.229	0.133	0.393	0.261
	(0.767)	(0.689)	(0.892)	(0.560)	(0.668)	(0.831)	
Government transparency	4.986	5.165	4.895	5.264	0.189	0.348	0.159
	(0.606)	(0.652)	(0.647)	(0.552)	(0.587)	(0.627)	
Overall government performance	5.453	5.393	5.462	5.533	-0.051	0.054	0.104
	(0.642)	(0.629)	(0.727)	(0.509)	(0.574)	(0.750)	

Notes: The pre-SDR period is 2009-2010 and the post-SDR period covers 2011-2012. The standard deviations are in parentheses.

Table 4: Random-effects regression models of PSP

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
	PSP index	Benevolence	Efficiency	Integrity	Equity	Transparency	Overall
Reform county	-0.0429 (0.0746)	0.00981 (0.0563)	0.00179 (0.0663)	-0.0434 (0.138)	-0.190 (0.122)	-0.0623 (0.0801)	0.0273 (0.101)
Reform year	0.187*** (0.0385)	0.267*** (0.0416)	0.311*** (0.0406)	0.364*** (0.0679)	0.121** (0.0596)	0.168*** (0.0531)	-0.0555 (0.0397)
Interaction term	0.136 (0.127)	0.0577 (0.0904)	0.0369 (0.114)	0.113 (0.204)	0.298 (0.182)	0.192 (0.124)	0.102 (0.147)
Lagged dependent variable	0.139*** (0.0500)	0.167*** (0.0523)	0.154*** (0.0493)	0.0748 (0.0455)	0.0325 (0.0428)	0.182*** (0.0460)	0.120*** (0.0439)
Population size†	-0.00944 (0.0310)	0.0274 (0.0263)	0.00956 (0.0229)	-0.0420 (0.0537)	-0.0335 (0.0485)	-0.0163 (0.0289)	0.00183 (0.0370)
Urbanization†	-0.0766 (0.121)	-0.134 (0.0992)	-0.117 (0.126)	0.0266 (0.179)	-0.0439 (0.165)	-0.0952 (0.111)	-0.0685 (0.117)
Education†	1.256*** (0.408)	1.559*** (0.415)	1.432*** (0.475)	0.844 (0.557)	0.845* (0.506)	1.294*** (0.415)	1.588*** (0.381)
Minority†	-0.0169 (0.0203)	-0.00319 (0.0199)	0.00301 (0.0194)	-0.0257 (0.0307)	-0.0544*** (0.0206)	-0.00651 (0.0218)	-0.0216 (0.0250)
Migration†	0.0552 (0.0402)	0.0758* (0.0431)	0.0709 (0.0446)	0.0145 (0.0544)	0.0389 (0.0522)	0.0553 (0.0397)	0.0728* (0.0399)
Unemployment†	-0.0555 (0.0872)	-0.0651 (0.0822)	-0.0746 (0.0886)	-0.0394 (0.114)	-0.0408 (0.116)	-0.0170 (0.0825)	-0.114 (0.0951)
Aging	-0.0116 (0.0130)	-0.00785 (0.0116)	-0.0159 (0.0122)	-0.0162 (0.0229)	-0.0143 (0.0208)	-0.0129 (0.0111)	-0.00250 (0.0133)
County-level city	-0.0358 (0.0596)	-0.0335 (0.0697)	-0.00221 (0.0736)	-0.00628 (0.0659)	-0.0536 (0.0547)	-0.0329 (0.0606)	-0.102 (0.0678)
District	0.0740 (0.0984)	0.0112 (0.0996)	0.0534 (0.110)	0.161 (0.106)	0.217** (0.110)	-6.27e-05 (0.0891)	0.0289 (0.126)
Mountainous county	0.123* (0.0631)	0.122* (0.0665)	0.161** (0.0700)	0.136 (0.0860)	0.100 (0.0862)	0.145** (0.0639)	0.0644 (0.0798)
Constant	1.774** (0.729)	0.963 (0.674)	1.256* (0.735)	2.442** (1.056)	3.203*** (1.006)	1.602** (0.736)	1.480** (0.719)
R^2	0.165	0.209	0.214	0.110	0.0904	0.159	0.165
Wald χ^2	600.1***	725.5***	1087***	394.8***	274.8***	400.7***	878.1***

Notes: $N = 468$. The Huber-White standard errors clustered at prefecture level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. † denotes that the variable is logged.