The Unintended Impact of Fiscal Hierarchy Reforms on Local Air Quality: Evidence from China*

Yongli Chen†

August 15, 2022

Abstract

This paper explores the causal relationship between the flattening of a fiscal hierarchy and environmental consequences using the Province-Managing-County (PMC) reform in China as a natural experiment. The reform enabled county governments to directly connect with provincial governments in fiscal affairs, bypassing the intermediary prefecture-level governments. Using a county-level panel data set from 2000 to 2012 and a Difference-in-Differences design, I find that local air quality measured by PM2.5 concentration was negatively affected by the PMC reform. Further analysis suggests that the deterioration in local air quality was not caused by better economic performance, since reformed counties' GDP per capita also witnessed a significant drop compared with non-reformed counties. These impacts of the PMC reform were more negative for counties in provinces with larger number of subordinate governments under direct supervision before the reform. Mechanism analysis shows that reformed counties spent less on environmental protection, which could weaken environmental monitoring and enforcement, so firms in reformed counties generated more air pollutants. The findings of this paper imply that flattening a fiscal hierarchy may lead to undesirable outcomes if the monitoring from the upper-level government becomes less effective due to the increased number of subordinates.

Keywords: Fiscal hierarchies, PMC reform, Air quality

JEL Codes: H11, H32, Q53

*I thank the Atmospheric Composition Analysis Group at Dalhousie University for publicly sharing the PM2.5 data on their website. I also thank Prof. Yang Xie, Prof. Richard Arnott, Prof. Kurt Schwabe, Prof. Ruoyao Shi and Prof. Ruixue Jia for their helpful comments and suggestions. All errors are my own.

†University of California, Riverside. Email address: ychen573@ucr.edu

1 Introduction

Air pollution is a problem for many developing countries around the world and is believed to kill more people worldwide than AIDS, malaria, breast cancer, or tuberculosis (Rohde and Muller, 2015). China, as the largest developing country, is experiencing such a problem after decades of economic growth at the expense of the environment. In China, air pollution was previously estimated to contribute to 1.2 to 2 million deaths annually (Rohde and Muller, 2015). Severe and persistent air pollution represented by excessive PM2.5 has become one of the main concerns of the Chinese public in recent years.

Air pollution control, or environmental protection in a broader sense, is essentially an issue of local public good provision. To promote efficient provision of local public goods, flattening fiscal hierarchy has been introduced in many countries to pass more revenue and spending authorities down to local governments (Hooghe et al., 2016; Ma and Mao, 2018). However, there is a key trade-off in such flattening reform that determines its actual effects. On the one hand, flattening government's fiscal hierarchy can reduce cumulative loss in fiscal funds across hierarchical levels, so environmental protection that were previously not a major concern of local officials could be allocated more funds. But on the other hand, flattening government's fiscal hierarchy may weaken the effectiveness of monitoring from the upper-level government, so the funds that should be allocated to environmental protection might be used for other purposes. Therefore, the net effect of fiscal hierarchy reforms on environmental protection is actually uncertain. To provide causal evidence on the relationship between fiscal hierarchy reforms and local environment, this paper explores a natural experiment in China, which is known as the Province-Managing-County (PMC) reform, and documents the unintended impact of this reform on local air quality.

The PMC reform started in 2003 and covered more than 1000 counties across the country. This reform enabled county governments to bypass the middle-level prefectural governments to receive funds from the provincial government and report expenditures to the provincial government directly. Therefore, this reform is a de facto flattening reform in terms of fiscal relations between different layers of government. The staggered rollout of this reform provides rich spatial and temporal variations, allowing me to use a difference-in-differences (DiD) method to explore its effects.

I construct a panel-data set which consists of 1807 counties in China from 2000 to 2012, accounting for 89.6% of all the counties across the country. One challenge in terms of the data that makes it hard to analyze the effect of the PMC reform on local environment is that there is no publicly available pollution data at county level for the sample period. To solve this challenge, I use annual average PM2.5 concentration retrieved from satellite data as the indicator for local air quality.

The analysis of this paper yields several findings. First, the adoption of the PMC reform increased a county's annual average PM2.5 concentration by 1.3%. This statistically significant effect was more pronounced in the long run. This finding indicates that the PMC reform had a negative impact on local air quality. Further analysis shows that the deterioration in local air quality was not caused by better economic performance, since reformed counties' GDP per capita also witnessed a significant drop by an average of 3.1% compared with non-reformed counties.

Second, since the PMC reform increased the number of subordinates of provinces, the impact of the reform could vary with the initial number of subordinate governments, which I call initial span of control, of provincial governments. Heterogeneity analysis shows that the PMC reform reduced local PM2.5 for counties in provinces with smaller initial span of control, but this effect is not significant. In provinces with greater initial span of control, the PMC reform had a more negative effect on local air quality. Similar heterogeneity pattern holds for economic performance.

Third, further analysis on county government spending suggests that the PMC reform led to more administrative expenditures but fewer aggregate expenditures and economic development expenditures. The reform was also associated with fewer expenditures on environmental protection. For counties in provinces with greater initial span of control, the impact on administrative expenditures was more positive, while the impacts on environmental protection and economic development expenditures were more negative. This could result from the problem that less effective monitoring became worse if a provincial government had greater initial span of control. Less environmental spending weakened environmental monitoring and enforcement, so firms in reformed counties used fewer exhaust treatment facilities and emitted more untreated smoke, which resulted in deteriorated local air quality.

This paper contributes to three strands of literature. First, a large number of existing literature evaluates the officially stated purposes of the PMC reform, namely, accelerating county-level economic growth (Li et al., 2016; Ma and Mao, 2018) and alleviating the financial difficulties of county governments (Li et al., 2016). These are the intended impacts of the PMC reform. However, the unintended impact of the PMC reform has not received enough attention. To evaluate a policy comprehensively, its unintended consequences should also be considered. This paper complements the existing literature by providing empirical evidence on the impact of the PMC reform on local air quality. The findings of this paper suggest that the adjustment of fiscal relations between different layers of government could generate undesirable outcomes when monitoring problems cannot be properly addressed.

This paper also fits into the literature on organization hierarchy. Classic works and some recent papers have discussed in theory the trade-off between deepening vertical layers and broadening the horizontal span of control (Simon, 2013; Calvo and Wellisz, 1978; Qian, 1994; Chen and Suen, 2019). In specific, it is argued that having fewer tiers within an organization can reduce cumulative loss across hierarchical levels, but the effectiveness of supervision to reduce moral hazard is weakened as a result of the increased span of control (Qian, 1994). However, there have not been many empirical explorations along this line. This paper shows empirically that a greater *initial* span of control highlights the negative side in the trade-off. Thus, flattening government hierarchies may not be a good policy choice if the initial number of subordinates of a supervising body is already large.

Finally, there has been a growing literature analyzing environmental issues in China from the perspective of political economy. For example, Jia (2017) explores the variation in provincial governors' connections with key officials in the center and finds that gaining connections increases pollution. He et al. (2020b) find that local government officials are incentivized to enforce tighter environmental standards on firms immediately upstream of a monitoring station. In addition, Kahn et al. (2015), Chen et al. (2018), and Bo (2021) investigate how promotion incentives affect local officials' performance on environmental protection. Different from existing literature which mainly focuses on the career concerns of local officials, this paper contributes to the understanding of the relationship between institutional reforms and environmental consequences. This paper finds that fiscal hierarchy reforms could result in a deterioration in local air quality through weak monitoring. This finding highlights the role of institutional reforms in understanding environmental issues in China and suggests the necessity to extend the research scope of this strand of literature from official-specific contexts towards a broader institutional side.

The rest of this paper is organized as follows. Section 2 introduces the institutional background of the PMC reform. Section 3 describes the empirical strategy and data. Section 4 presents main empirical results. Section 5 further discusses the impact of the PMC reform on economic performance. Section 6 shows the heterogeneous impacts of the PMC reform across the initial span of control of provincial governments. Section 7 analyzes the potential mechanisms from both government and firm sides. Section 8 concludes.

2 Institutional background

The hierarchical structure of governance in China consists of five layers of government. From the highest to the lowest, they are the central government, province or municipality-level government, prefecture or city-level government, county-level government, and township-level government. Between the 1980s and the early 2000s, this hierarchical system featured a highly centralized fiscal managing system. Governments at the prefecture level undertook a strict control of fiscal revenues and expenditures of the lower county-level governments. The risk of this institutional arrangement is that prefectural governments are inclined to withhold revenue sources allocated to counties and at the same time shift expense responsibilities to counties due to the priority of urban construction and development that had been pursued by most prefectural governments in China (Wang et al., 2012).

The central government, therefore, enacted a fiscal reform known as the PMC reform, aiming to lessen the fiscal stress of county governments. In contrast to the old system where county governments were strictly under the control of prefectural governments, this reform enabled county governments to directly receive funding from the provincial government and report to the provincial government about expenditures. This hierarchical change is illustrated using Figure 1, where County D is selected as the reformed county. As prefectural governments cannot withhold fiscal sources allocated to PMC counties and shift expenditure responsibilities to them, the PMC reform is expected to promote fiscal decentralization in China (Wang et al., 2012). In fact, the PMC system already existed in Zhejiang and Hainan provinces in the late 1980s. The nation-wide PMC reform started in 2003 in some provinces located in the middle and north parts of China, and then the program continued expanding to other provinces. Although the Ministry of Finance, in July 2009, released *the Opinions on Enhancing the Province-Managing-County Reform*, proposing to implement this reform across all provinces except the minority autonomous regions by 2012, it seems that the PMC reform was slowed down and even at a standstill since 2012. By the end of 2012, 1099 counties from 24 provinces, representing approximately 56% of all counties across the country, had implemented the PMC reform (see Figure 2 for the spatial distribution).

This reform may have impacts on local environmental outcomes by changing county governments' fiscal behaviors. After the PMC reform, reformed counties enjoyed more fiscal autonomy, but the down side of this reform is that the monitoring from upper level governments became less stringent, since the provincial government had to monitor more subordinates than before.[1] This monitoring issue may change the incentive of local officials, and the way they allocate fiscal resources. Since in the sample period (2000-2012), the role of environmental protection in the evaluation and promotion of local officials was not as important as later, local governments might spend less on local environmental protection and divert funds for other purposes. If so, we may see a deterioration in environmental outcomes after a county enrolled in the PMC reform.

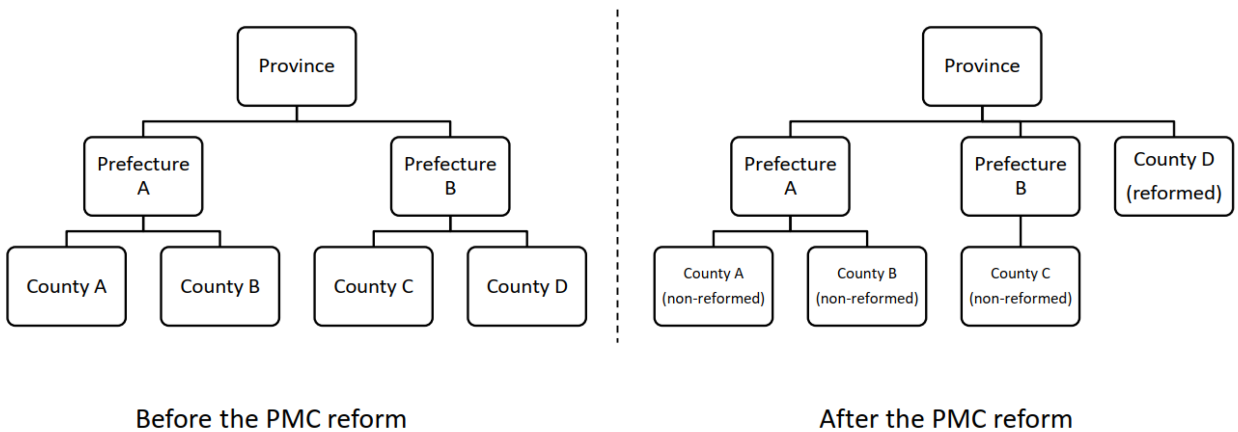

Figure 1: The hierarchical change in the PMC reform.

3 Estimation strategy and data

3.1 Empirical specification

I estimate the impact of the PMC reform using a DiD approach. The treatment units are the counties participated in the PMC reform during the sample period (for example, County D in Figure 1). Non-reformed counties serve as control units (for example, Counties A, B and C in Figure 1).

[1]Before the PMC reform, the average number of subordinate governments under the direct supervision of a provincial government was about 12, while this number increased dramatically to about 52 after the reform.

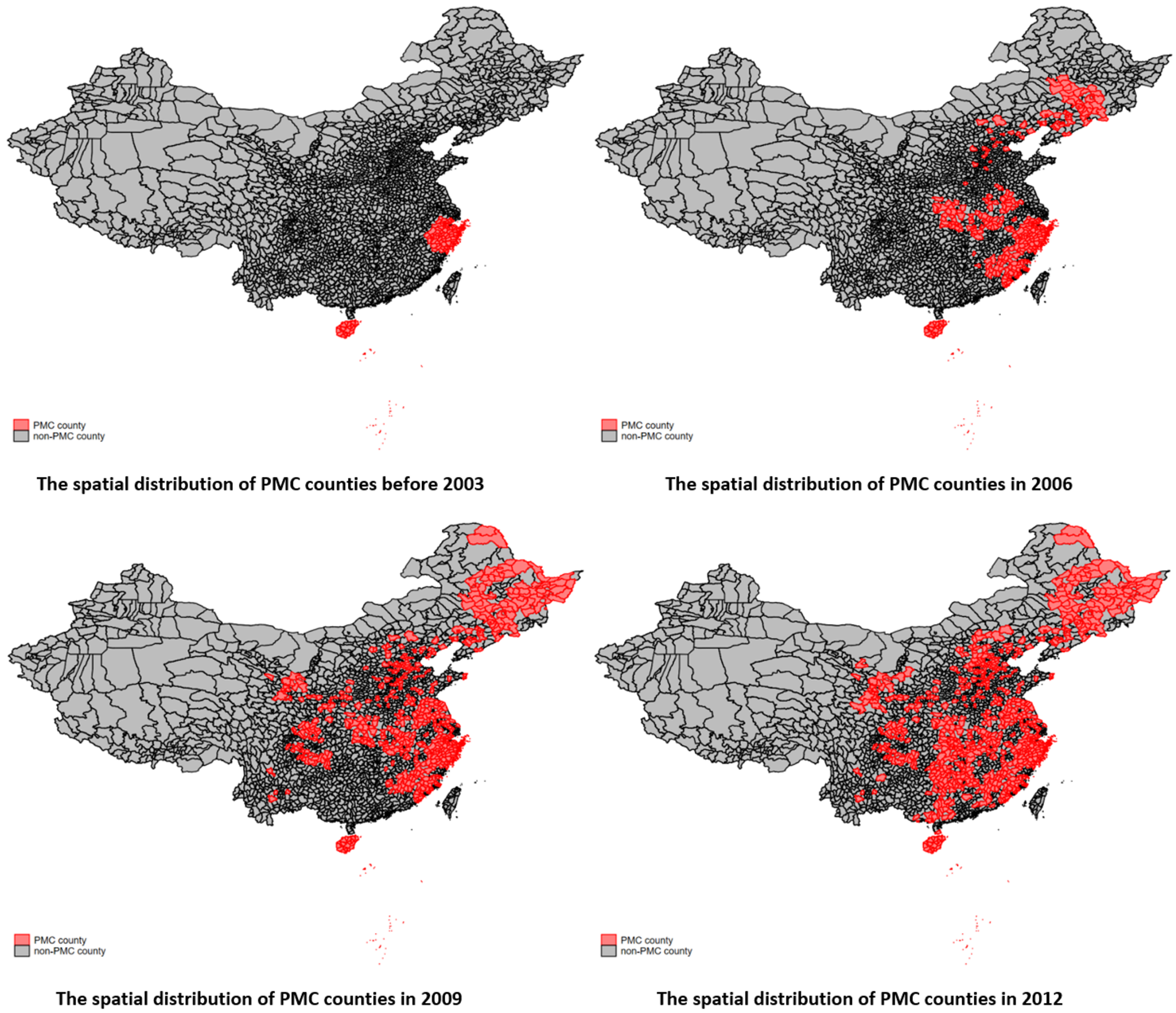

Figure 2: The expansion of the PMC reform.

The baseline DiD specification takes the following form:

$$y_{ct} = \beta_1 PMC_{ct} + \alpha_c + \gamma_t + \epsilon_{ct} \tag{1}$$

where c and t indicate county and year, respectively. y_{ct} represents the outcome variable, i.e., the logarithm of annual average PM2.5 concentration of county c in year t. PMC_{ct} is the variable of interest. Specifically, $PMC_{ct} = treatment_c * post_{ct}$, where $treatment_c = 1$ if county c participated in the PMC reform during the sample period, and 0 otherwise. $post_{ct}$ is a post-treatment indicator, taking a value of 1 if $t \geq t_{c0}$, where t_{c0} is the time when county

c was engaged in the reform. α_c is the county fixed effects, capturing all the time-invariant characteristics of the county. γ_t is the year fixed effects, controlling for time trend that affect the outcome variable. ϵ_{ct} is an error term. Standard errors are clustered at county level.

As extensions to the baseline model, I further add some control variables to get relatively precise estimates of the policy effect. As atmospheric conditions have significant explanatory power for air pollution (Zheng et al., 2019), I also include the annual average temperature, annual average precipitation and annual ventilation coefficient[2] of each county as control variables. In addition, treatment-specific linear time trends ($treatment_c * t$) are also added to control for the differences in time trends between the treatment and control groups. Moreover, since the selection of counties to adopt the PMC reform was not random, following Li et al. (2016), I control for eight criteria that provinces used in selecting the PMC counties, including county-level city, national poor county, major food-producing county, provincial boundary county, altitude, average slope, fiscal gap, and urbanization rate. Lastly, the County-Power-Expansion (CPE) reform was in place in some counties during the sample period. This reform decentralized some economic administrative powers and authorities, such as over infrastructure construction, technology upgrading, and foreign-invested projects, from the prefectural government to county governments. To rule out the effect from the CPE reform, another dummy variable indicating whether a county government carried out a CPE reform is included as an additional control variable. The definitions of these variables are summarized in Table 1.

3.2 Identifying assumptions

Although the selection of PMC counties is not random, the selection criteria as mentioned above are not directly related to any environmental outcomes. Therefore, in this sense the

[2] Ventilation coefficient is defined as the product of wind speed, which determines the horizontal dispersion of pollution, and mixing height, which determines the height at which pollutants disperse in the atmosphere. It reflects the meteorological conditions that influence the speed of dispersion of pollutants in the air (Hering and Poncet, 2014), and it is a more comprehensive measure of the impact of wind on the dispersion of pollutants.

Table 1: Summary of variables.

Variable	Definition
Indicators for local air quality:	
PM2.5	Annual average of PM2.5 concentration (µg/m^3)
Core explanation variable:	
PMC	=1 if a county is a PMC county in year t; =0 otherwise
Selection criteria:	
city	=1 if a county is a county-level city; =0 otherwise
poor	=1 if a county is a national poverty county; =0 otherwise
food	=1 if a county is a national food or cotton production county; =0 otherwise
boundary	=1 if a county is a provincial boundary county; =0 otherwise
altitude	County seat's altitude (km)
slope	Average county slope (degrees)
fsratio	Ratio of fiscal expenditure to fiscal revenue in year 1999
ubrate	Percentage of non-agricultural population in the total population in year 2000
Weather controls:	
temperature	Annual average temperature of county c ($^\circ C$)
precipitation	Annual average precipitation of county c (mm)
ventilation coefficient	The product of annual average wind speed and mixing height
Other variables:	
cpe	=1 if a county adopted County-Power-Expansion reform in year t and afterwards; =0 otherwise
gdp_pc	Real GDP per capita of county c
span2000	Initial span of control of the provincial government that county c is managed by.
County fiscal variables:	
totalexp_pc	Total budget expenditure per capita of county c (yuan)
adm_pc	Administration expenditure per capita of county c (yuan)
grth_pc	Productive investment expenditure per capita of county c (yuan)
env_pc	Environmental protection expenditure per capita of county c (yuan)
Firm-level variables:	
coal_p	Fuel coal consumption per unit output (ton per ten thousand yuan)
oil_p	Fuel oil consumption per unit output (ton per ten thousand yuan)
smoke_p	untreated smoke per unit output (cubic meter per ten thousand yuan)
facility	Number of exhaust treatment facilities

selection of the treatment group is random in terms of the outcome variable. Even though there might be some unobservable factors that could affect the PM2.5 concentration of a county, this type of endogeneity can be addressed by including county and year fixed effects.

As a formal test for the parallel trend assumption, I conduct an event study following Jacobson et al. (1993). Annual changes in local PM2.5 concentration before and after adopting the PMC reform are estimated using the following specification:

$$y_{ct} = \beta_i \sum_{i=-4}^{4} D_{t_{c0+i}} + \theta_1 treatment_c * t + \theta_2 weather_{ct} + \theta_3 CPE_{ct} + S * \gamma_t + \alpha_c + \gamma_t + \epsilon_{ct} \quad (2)$$

where $D_{t_{c0+i}}$ is a series of dummies indicating whether $t - t_{c0} = i$, where $i = -4, -3...3, 4$, and t_{c0} represents the year when county c was engaged in the PMC reform. So, the sum of dummy variables $\sum_{i=-4}^{4} D_{tc0+i}$ represents an 8-year window spanning from 4 years before the reform until 4 years after the reform. For those years that are outside the window period, following Beck et al. (2010) and He et al. (2020c), I set D_{tc0-4} equals 1 for all years that are 4 or more years before the reform and D_{tc0+4} equals 1 for all years that are 4 or more years after the reform. The omitted time category is $i = -1$, so the coefficient of interest β_i captures the treatment effect relative to the year just before adopting the PMC reform. S is a matrix of selection variables. Other variables are defined in the same way as Equation 1 and standard errors are clustered at county level.

If the coefficients of the lead terms are not significantly different from zero, then we can say that the parallel trend assumption holds. As will be discussed in Section 4.2, the hypothesis that reformed and non-reformed counties follow similar trends in terms of annual average PM2.5 before the PMC reform cannot be rejected.

One may also suspect that other counties that were in the same prefecture as reformed counties before, like County C in Figure 1, could be affected indirectly by the PMC reform, so including these counties in the control group may violate the non-spillover assumption. However, as we will see in Section 4.4, such spillover effect is not statistically significant. Therefore, the non-spillover assumption still holds.

Another concern about the identifying assumption is that the timing of adopting the PMC reform may be endogenous. On the one hand, counties with specific characteristics might be selected to participate in the reform earlier than other counties. On the other hand, a province with greater initial span of control might adopt the PMC reform later than a province with smaller initial span of control. To address the first county-level concern, I focus on counties that eventually participated in the PMC reform. I then regress the number of years before a county became a reformed county on eight selection criteria, respectively. The results are shown in columns 1 to 8 of Table 2. As we can see, none of the selection criteria is significantly correlated with the timing of the PMC reform at county level. For the province-level concern, I regress the order of adopting the PMC reform for each province on the ranking of initial span of control. As shown in column 9 of Table 2, the estimated coefficient is negative (-0.182) but statistically insignificant, which means that the timing of the reform at province level is not determined by the initial span of control.

Table 2: Testing the endogeneity of the timing of the PMC reform

	County-level timing								Province-level timing
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
city	0.209								
	(0.129)								
poor		-0.046							
		(0.130)							
food			0.291						
			(0.241)						
boundary				0.027					
				(0.068)					
slope					-0.015				
					(0.011)				
altitude						-0.182			
						(0.185)			
urbrate							0.013		
							(0.015)		
fsratio								-0.041	
								(0.037)	
Ranking of initial span of control									-0.182
									(0.149)
Obs.	6,761	6,761	6,761	6,761	6,761	6,761	6,761	6,761	19
R^2	0.0018	0.0001	0.0053	0.0001	0.0024	0.0041	0.0048	0.0012	0.0801

Notes: This table tests the endogeneity of the timing of the PMC reform. Columns 1 to 8 are based on treated sample and show the correlation coefficients between each selection criterion and the number of years before a county participated in the reform. These results address the county-level endogenous timing concern. Column 9 regresses the order of adopting the PMC reform for each province on the rank of initial span of control. This result addresses the province-level endogenous timing concern.

3.3 Data

The dataset consists of 1807 counties across China from 2000 to 2012. As the analyses focus on the PMC reform in the 2000s, Hainan and Zhejiang provinces are not included, since these two provinces adopted the PMC system in the late 1980s. All the counties in Tibet are excluded due to missing data. Counties from four province-level municipalities are also excluded.

One major challenge to evaluating the environmental impact of the PMC reform on counties is that county-level pollution data is very scarce, especially for the sample period. I solve this challenge by using annual average PM2.5 concentration as the measure for local air quality. This data is provided by the Atmospheric Composition Analysis Group at Dalhousie University. They estimated ground-level fine particulate matter (PM2.5) over China by combining Aerosol Optical Depth (AOD) retrievals from the NASA MODIS, MISR, and SeaWIFS instruments with the GEOS-Chem chemical transport model, and subsequently calibrated to regional ground-based observations using Geographically Weighted Regression (GWR). [3] The higher the value of PM2.5, the lower the air quality is.

County-level financial information comes from the National Prefecture and County Finance Statistics (NPCFS). Other characteristics of counties come from annual statistical yearbooks of provinces. All money values are deflated using the price deflator with Beijing as the base province and 2000 as the base year. Weather controls are from China Weather Website. Firm-level data is from China Industrial Enterprise Database, matched with information on firms' pollution.[4]

The summary statistics of all variables are shown in Table 3. The last column suggests that the treatment group and control group are different, so controlling for country fixed effects is necessary.

4 Empirical results

4.1 Impact of the PMC reform on local PM2.5

Table 4 presents the impact of the PMC reform on local PM2.5. The regression result of the baseline specification is shown in column 1. In column 2, weather controls are included. Treatment-specific linear time trends are included in column 3 to control for the differences in time trends between the PMC and non-PMC counties. Column 4 further considers eight selection criteria of the PMC counties and thus controls for interactions of these variables

[3] More information about this dataset is available at http://fizz.phys.dal.ca/~atmos/martin/?page_id=140

[4] China Industrial Enterprise Database covers the basic information about finance, production and sales of all industrial enterprises with sales of more than 5 million yuan (more than 20 million yuan since 2011) in mainland China. It is one of the most comprehensive database for firm-level analysis.

Table 3: Summary statistics.

Variable	*Treatment group*			*Control group*			Difference in mean
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.	
PM2.5	11566	44.79	20.216	9573	36.73	23.863	8.056***
city	11566	0.190	0.392	9573	0.168	0.374	0.022***
poor	11566	0.296	0.456	9573	0.351	0.477	-0.056***
food	11566	0.340	0.474	9573	0.198	0.399	0.142***
boundary	11566	0.406	0.491	9573	0.365	0.481	0.041***
altitude	11566	0.473	0.626	9573	0.983	0.977	-0.510***
slope	11566	8.341	6.234	9573	10.05	7.530	-1.708***
fsratio	11566	2.101	1.496	9573	3.179	3.423	-1.077***
ubrate	11566	15.48	9.847	9573	17.08	13.229	-1.600***
temperature	11566	13.95	4.480	9573	12.47	4.211	1.477***
precipitation	11566	9909	4270.78	9573	7821	3748.787	2097***
ventilation coefficient	11566	1860	169.771	9573	1857	165.297	2.852
cpe	11566	0.300	0.458	9573	0.100	0.300	0.200***
gdp_pc	11566	9732	11375.39	9573	12000	16670.26	-1949.352***
span2000	11566	10.370	4.464	9573	10.162	3.750	0.208***
totalexpenditure	4635	575	487.059	3830	854.189	906.810	-278.611***
adm_pc	4359	75.862	58.733	3516	125.654	134.109	-49.792***
grow_pc	4635	74.533	111.671	3830	144.865	277.142	-70.333***
env_pc	810	27.332	42.817	634	90.025	178.026	-62.693***
coal_p	44016	4.693	69.313	109643	2.346	248.326	2.347
oil_p	18300	0.027	0.511	73752	0.041	3.822	-0.014
smoke_p	53004	29.289	4863.403	121483	11.734	374.996	17.555***
facility	43793	2.957	8.483	110456	2.803	9.652	0.154***
total assets	107329	199925.3	1240893	66834	266305.9	1908352	-66380.6***
ownership	107322	4.457	2.858	66847	4.313	2.934	0.144***
firm age	107204	11.785	24.529	66814	12.420	51.341	-0.635***

Notes: The difference in mean is calculated using the mean of the treatment group minus the mean of the control group. Two-sample t-tests with unequal variances are conducted to check whether the difference is significantly different from 0. ***denotes significance at 1% level.

with the year dummy. The effect from the CPE reform is controlled in column 5. From Table 4, we can find that although the magnitude of the treatment effect decreases as more control variables are added to the baseline specification, the PMC reform resulted in a statistically significant increase in the local PM2.5 concentration. To be specific, according to the result for the preferred specification in column 5, adopting the PMC reform caused the annual average PM2.5 concentration of reformed counties to increase by 1.3%. The average PM2.5 concentration for sample period is about 41 µg/m^3, so this coefficient can be translated into a 0.53 µg/m^3 increase per year. Existing literature has shown that a 10 µg/m^3 increase in PM2.5 increases mortality by 3.25% (He et al., 2020a). Given this fact, this small but significant result can not be ignored.

Table 4: The impact of the PMC reform on PM2.5.

Dependent variable	*log(PM2.5)*				
	(1)	(2)	(3)	(4)	(5)
PMC	0.027***	0.026***	0.019***	0.015**	0.013**
	(0.005)	(0.005)	(0.007)	(0.006)	(0.006)
County fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Weather controls		Yes	Yes	Yes	Yes
Treatment trend			Yes	Yes	Yes
Selection criteria×Year dummy				Yes	Yes
CPE dummy					Yes
Adj. R-squared	0.918	0.919	0.919	0.936	0.936
Obs.	21,180	21,180	21,180	21,180	21,180

Notes: This table estimates the impacts of the PMC reform on local PM2.5 concentration with different specifications. Weather controls include annual average temperature, precipitation and ventilation coefficient for each county. Treatment trend refers to treatment-specific linear time trends ($treatment_c * t$). Selection criteria includes county-level city, national poor county, major food-producing county, provincial boundary county, altitude, average slope, fiscal gap, and urbanization rate. CPE is a dummy variable indicating whether a county government carried out County-Power-Expansion reform in the sample period. Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Dynamic impact of the PMC reform

This subsection analyzes the dynamic impact of the PMC reform using event study approach (Jacobson et al., 1993). Figure 3 plots the results, with the horizontal axis shows the number of years from the PMC reform. The negative numbers denote years before a county participated in the PMC reform (pre-treatment periods), while the positive numbers refer to years after the reform (post-treatment periods). The year just before adopting the PMC reform is set as the benchmark. As we can see, the coefficients of pre-treatment periods are close to 0, which suggests that reformed and non-reformed counties followed similar time trends before the adoption of the PMC reform. For post-treatment periods, Figure 3 shows that the coefficients of the PMC reform become significantly positive and gradually increase in magnitude except for a slight drop in the second year. This implies that the negative effect of the PMC reform on local air quality is more pronounced in the long run.

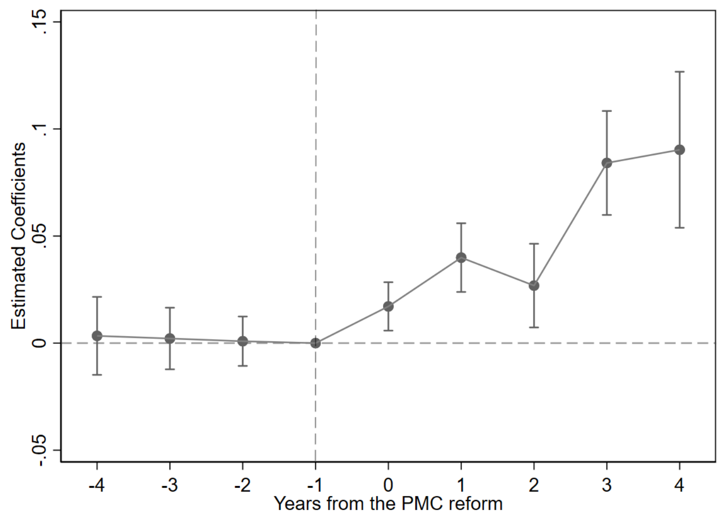

Figure 3: Dynamic impact of the PMC reform.

Notes: This figure plots the estimated coefficients and their 95% confidence intervals based on Equation 2. The horizontal axis denotes years before or after a county adopted the PMC reform, where the year just before adopting the PMC reform is the benchmark.

4.3 Robustness checks

I check the robustness of the main finding in several ways. First, based on the baseline specification in column 1 of Table 4 and the specification with all control variables in column 5 of Table 4, I further control for provincial trends in columns 1 and 3 of Table 5. The results show that the estimated coefficients of PMC are still positive and are significant at 1% and 10% levels, respectively. Then, to control for annual common shocks to all counties in a province, such as provincial abatement policies, I follow He et al. (2020c) and add province-year fixed effects to the specifications from Table 4. In columns 2 of Table 5, although the estimated coefficients of the treatment variable drop in magnitude compared with the baseline result, the sign of the effect is still the same and remains significant. After adding all control variables, as shown in column 4 of Table 5, the estimated coefficient of PMC is close to that of column 5 of Table 4 (0.013) and is statistically significant at 1% level. These results confirm the robustness of the main finding.

Table 5: Robustness checks.

Dependent variable	*log(PM2.5)*			
	(1)	(2)	(3)	(4)
PMC	0.027***	0.014***	0.010*	0.016***
	(0.005)	(0.005)	(0.006)	(0.006)
County fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes		Yes	
Provincial trends	Yes		Yes	
Province-year fixed effects		Yes		Yes
Weather controls			Yes	Yes
Treatment trend			Yes	Yes
Selection criteria×Year dummy			Yes	Yes
CPE dummy			Yes	Yes
Adj. R-squared	0.918	0.959	0.936	0.960
Obs.	21,180	21,180	21,180	21,180

Notes: This table summarizes the results of robustness checks. Columns 1 and 3 control for provincial trends. Columns 2 and 4 control for province-year fixed effects. Weather controls include annual average temperature, precipitation and ventilation coefficient for each county. Treatment trend refers to treatment-specific linear time trends ($treatment_c * t$). Selection criteria includes county-level city, national poor county, major food-producing county, provincial boundary county, altitude, average slope, fiscal gap, and urbanization rate. CPE is a dummy variable indicating whether a county government carried out County-Power-Expansion reform in the sample period. Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.4 Tests addressing other concerns

First, to check the extent to which the results are influenced by any omitted variables, following Li et al. (2016), I conduct a placebo test by randomly assigning the PMC status to counties. To preserve the fact that eight years in the sample periods have positive number of reformed counties ({58,113,75,15,256,201,153,87}), eight years between 2001 and 2011 are selected at random.[5] Then, within each year the same number of counties as listed above are randomly selected as the treatment group without replacement. After having this randomly generated dataset, a placebo DiD estimation is conducted using the specification in column 5 of Table 4. Since the PMC status is now randomly designated, the estimated coefficient of PMC should be close to zero. Otherwise, it would indicate a misspecification problem. To

[5]As required by the DiD method, there should be at least one year before and one year after the PMC adoption.

increase the identification power of this placebo test, it is repeated 500 times.

Figure 4 plots the distribution of the estimated coefficient of PMC from the 500 runs, together with the benchmark estimate (0.013) from column 5 of Table 4 (represented by the vertical line). The distribution of estimates is clearly centered around zero, and the standard deviation of the estimates is 0.00628, indicating no effect with the randomly assigned PMC status. The benchmark result is at 95 percentile of the distribution and we can also observe that most of the estimates are distributed to the left of the benchmark estimate. Taken together, the placebo test suggests that the effect of the PMC reform on PM2.5 is not driven by unobserved factors.

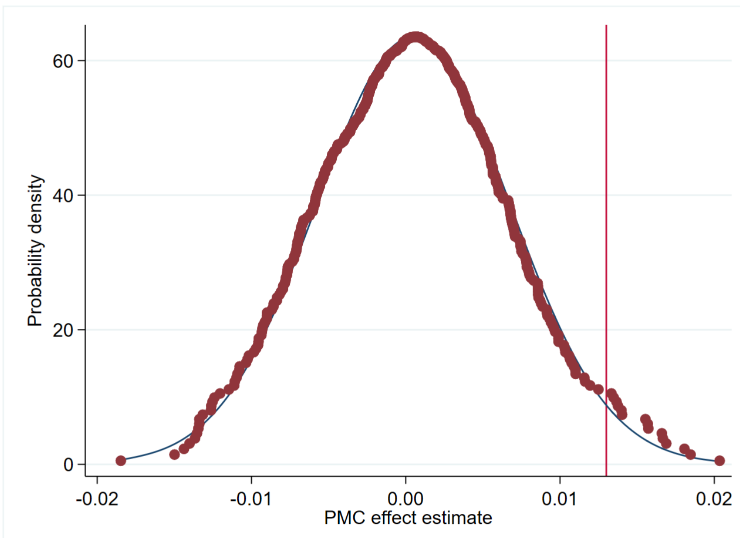

Figure 4: The distribution of the estimates from the placebo test.

Notes: This figure shows the cumulative distribution density of the estimated coefficients from 500 simulations randomly assigning the PMC status to counties. The vertical line is the benchmark estimate from column 5 of Table 4.

Next, I test for two types of spillover effects, which could confound the main result. The first one is about geographical spillovers. The deterioration of air quality in a reformed county might affect its neighboring non-reformed counties. To address this concern, following He et al. (2020c), I focus on counties that never participated in the PMC reform in the sample and estimate the relative impacts of having reformed neighboring counties on the annual average PM2.5 concentration of these non-reformed counties. To be specific, I create a new dummy variable $Neigbor_{ct}$ which equals 1 if at least one of the neighboring counties of county c participated in the reform in year t, and 0 otherwise. Then, PMC_{ct} in Equation 1 is replaced with $Neigbor_{ct}$, the coefficient of which captures the spillover effect of concern. As shown in columns 1 and 2 of Table 6, although having reformed neighboring counties increases a non-reformed county's annual average PM2.5 concentration, but this spillover effect is statistically insignificant. In other words, there is no evidence of geographical spillover effects.

Table 6: Testing spillover effects.

Dependent variable	*log(PM2.5)*			
	(1)	(2)	(3)	(4)
Neighbor	0.007	0.007		
	(0.009)	(0.009)		
PMC			0.034***	0.013**
			(0.011)	(0.006)
PMCShare			-0.009	-0.001
			(0.012)	(0.001)
County fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Weather controls		Yes		Yes
Treatment trend				Yes
Selection criteria×Year dummy				Yes
CPE dummy		Yes		Yes
Adj. R-squared	0.958	0.958	0.918	0.936
Obs.	6,200	6,200	21,180	21,180

Notes: This table summarizes the regression results about spillover effects. Columns 1 and 2 test geographical spillovers. Columns 3 and 4 test if unreformed counties that were in the same prefecture as reformed counties could be affected by the PMC reform. $Neighbor$ equals 1 if at least one of the neighboring counties of county c participates in the reform in year t. $PMCShare$ measures the proportion of reformed counties among all counties in a prefecture. Weather controls include annual average temperature, precipitation and ventilation coefficient for each county. Treatment trend refers to treatment-specific linear time trends ($treatment_c * t$). Selection criteria includes county-level city, national poor county, major food-producing county, provincial boundary county, altitude, average slope, fiscal gap, and urbanization rate. CPE is a dummy variable indicating whether a county government carried out County-Power-Expansion reform in the sample period. Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Another concern of the spillover effects is that, as previously mentioned, other counties that were in the same prefecture as reformed counties before, like County C in Figure 1, could also be affected by the PMC reform since the prefectural government had less subordinates than before and thus might allocate more attention on non-reformed counties. To test if such spillover effects exist, I introduce a new variable $PMCShare_{ct}$ to measure the proportion of reformed counties among all counties in a prefecture. This variable characterizes the reform intensity of a prefecture, and it varies with time as more non-reformed counties are selected into the reform. Following the strategy of Miguel and Kremer (2004) and Bo et al. (2020), I add this variable to the specifications in columns 1 and 5 of Table 4 and its coefficient captures the spillover effect within prefectures. As columns 3 and 4 of Table 6 shows, the estimated coefficients of $PMCShare_{ct}$ are statistically insignificant, suggesting that conditional on a county's own PMC status, it was not affected by the share of reformed counties in the same prefecture. Therefore, the spillover effects within prefectures are less of a concern.

Lastly, recent advances in econometric theory argue that under staggered treatment settings, the two-way fixed effects (TWFE) estimator might be biased when already treated units are used as the control group and the treatment effects are heterogeneous (Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Borusyak et al., 2021). To test if such problems threat the main result of this paper, I re-estimate the specifications in columns 1 and 5 of Table 4 using two newly developed methods. The first one is the estimation method proposed by Callaway and Sant'Anna (2021). The basic idea of this method is to first estimate group-time average treatment effects and then aggregate them into a single overall treatment effect parameter (Callaway and Sant'Anna, 2021). The corresponding results are reported in columns 1 and 2 of Table 7. The second one is the imputation approach proposed by Borusyak et al. (2021), which imputes potential outcomes for the treated units using data for the untreated units. The replication results are shown in columns 3 and 4 of Table 7. The estimated coefficients of PMC using both methods are close to those in Table 4 (0.027 and 0.013), suggesting that there is little bias induced by the mismatch between comparison groups and the heterogeneity of the treatment effects in the baseline estimates.

Table 7: Checking for the validity of TWFE estimates.

Dependent variable	*log(PM2.5)*			
	(1)	(2)	(3)	(4)
PMC	0.020***	0.018**	0.021***	0.017***
	(0.006)	(0.007)	(0.004)	(0.004)
County fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Weather controls		Yes		Yes
Treatment trend		Yes		Yes
Selection criteria×Year dummy		Yes		Yes
CPE dummy		Yes		Yes
Obs.	21,138	21,138	21,139	21,139

Notes: This table re-estimates specifications in columns 1 and 5 of Table 4. Columns 1 and 2 correspond to the method proposed by Callaway and Sant'Anna (2021), and columns 3 and 4 correspond to the imputation approach proposed by Borusyak et al. (2021). Adjusted R-squared or R-squared is not given by both methods. Weather controls include annual average temperature, precipitation and ventilation coefficient for each county. Treatment trend refers to treatment-specific linear time trends ($treatment_c * t$). Selection criteria includes county-level city, national poor county, major food-producing county, provincial boundary county, altitude, average slope, fiscal gap, and urbanization rate. CPE is a dummy variable indicating whether a county government carried out County-Power-Expansion reform in the sample period. Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5 Further discussion on the impact of the PMC reform on economic performance

As we have seen from the previous analyses, the PMC reform led to a deterioration in local air quality. A natural question is that, did the reformed counties sacrifice local environment in exchange for higher GDP? In this section, I try to answer this question by investigating the impact of the PMC reform on counties' economic performance, which is measured by GDP per capita.

Table 8 reports the average effect of the PMC reform on GDP per capita of reformed counties. The empirical specifications are the same as those used with PM2.5 concentration as the outcome variable, but do not include weather controls. The results show that adopting the reform had a significantly negative impact on counties' economic performance. The preferred specification (column 4 of Table 8) suggests that adopting the reform decreased a county's GDP per capita by 3.1% on average. This negative effect of the PMC reform on counties' economic perform is in line with the findings of some existing literature, for example, Li et al. (2016).

This finding, together with the main result, indicates that the deterioration in local air quality was not a consequence of better economic perfromance. This counter-intuitive finding implies that we should investigate some factors that are not environmentally specific to understand these results.

Table 8: The impact of the PMC reform on counties' economic performance.

Dependent variable	*log(gdp_pc)*			
	(1)	(2)	(3)	(4)
PMC	-0.065***	-0.019**	-0.025***	-0.031***
	(0.011)	(0.008)	(0.008)	(0.009)
County fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Treatment trend		Yes	Yes	Yes
Selection criteria×Year dummy			Yes	Yes
CPE dummy				Yes
Adj. R-squared	0.943	0.943	0.946	0.946
Obs.	23,491	23,491	23,491	23,491

Notes: This table estimates the impact of the PMC reform on counties' GDP per capita (in logarithm). GDP data is deflated using Beijing as the base province and 2000 as the base year. $Span2000$ measures the initial span of control of a provincial government. Treatment trend refers to treatment-specific linear time trends ($treatment_c * t$). Selection criteria includes county-level city, national poor county, major food-producing county, provincial boundary county, altitude, average slope, fiscal gap, and urbanization rate. CPE is a dummy variable indicating whether a county government carried out County-Power-Expansion reform in the sample period. Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6 The role of initial span of control

6.1 Heterogeneous impacts of the PMC reform on local PM2.5

One of the most substantive features of the PMC reform is the flattened government hierarchies. Theoretically speaking, horizontal hierarchies feature less delay, but more communication and coordination costs compared with vertical hierarchies (Garicano and Van Zandt, 2012). More importantly, the incentive theory suggests that when supervision is used to mitigate moral hazard in a hierarchical organization, the probability of monitoring decreases with the span (Qian, 1994). In the context of the PMC reform, if initially the span of control of a provincial government is relatively small, adding more subordinates may not be so harmful. However, if the initial span of control is already large, the increase in the number of subordinates makes the negative side of a horizontal hierarchy more severe. Therefore, the impact of the PMC reform may be more negative when the initial span of control of provincial governments is greater.

To test this argument formally, I introduce a new variable $Span_{c,2000}$ to quantify the initial span of control of the provincial government that county c is administrated by in 2000. To be specific, $Span_{c,2000}$ measures the number of subordinates of a provincial government in 2000, including prefectures and counties that are under its direct supervision. For example, in Figure 1, the initial span of control of the provincial government before the PMC reform is 2 (Prefecture A and Prefecture B). Then, I estimate the following equation:

$$y_{ct} = \beta_1 PMC_{ct} + \beta_2 PMC_{ct} * Span_{c,2000} + \alpha_c + \gamma_t + \epsilon_{ct} \tag{3}$$

where β_2 is the coefficient of interest. If the above argument holds, β_2 is expected to be positive when the outcome variable is PM2.5 concentration and negative when the outcome variable is GDP per capita.

Table 9 summarizes the heterogeneous impacts of the PMC reform on local PM2.5. In column 1, the baseline result shows that the estimated coefficient of the interaction term between the reform dummy and the initial span of control is positive and significant at 1% level. This result is consistent across specifications with different controls, as shown in columns 2 to 5. This finding implies that the impact of the PMC reform on local PM2.5 was more adverse for counties in provinces which had greater initial span of control.

Table 9: Heterogeneous impacts of the PMC reform on local PM2.5 across the initial span of control of provincial governments.

Dependent variable	*log(PM2.5)*				
	(1)	(2)	(3)	(4)	(5)
PMC	-0.025**	-0.024**	-0.031***	-0.031***	-0.036***
	(0.010)	(0.010)	(0.011)	(0.011)	(0.010)
PMC×Span2000	0.005***	0.005***	0.005***	0.004***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
County fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Weather controls		Yes	Yes	Yes	Yes
Treatment trend			Yes	Yes	Yes
Selection criteria×Year dummy				Yes	Yes
CPE dummy					Yes
Adj. R-squared	0.918	0.919	0.919	0.936	0.936
Obs.	21,180	21,180	21,180	21,180	21,180

Notes: This table summarizes the regression results about the heterogeneous impacts of the PMC reform on local PM2.5 concentration with different specifications. $Span2000$ measures the initial span of control of a provincial government. Weather controls include annual average temperature, precipitation and ventilation coefficient for each county. Treatment trend refers to treatment-specific linear time trends ($treatment_c * t$). Selection criteria includes county-level city, national poor county, major food-producing county, provincial boundary county, altitude, average slope, fiscal gap, and urbanization rate. CPE is a dummy variable indicating whether a county government carried out County-Power-Expansion reform in the sample period. Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Based on the distribution of the initial span of control of provincial governments, I further predict the relative impacts of the PMC reform at different quantiles of initial span of control. Figure 5 shows the predicted relative impacts with 95% confidence intervals using the specification in column 5 of Table 9. We can see that the PMC reform reduced local PM2.5 concentration in provinces of which the initial span of control ranks at the first 20% quantile, but this effect is not statistically significant. In provinces with greater initial span of control (the third 20% quantile and above), the PMC reform significantly increased local PM2.5 concentration. The PMC reform did not have significant effect on local PM2.5 when provinces' initial span of control ranks at the second 20% quantile.

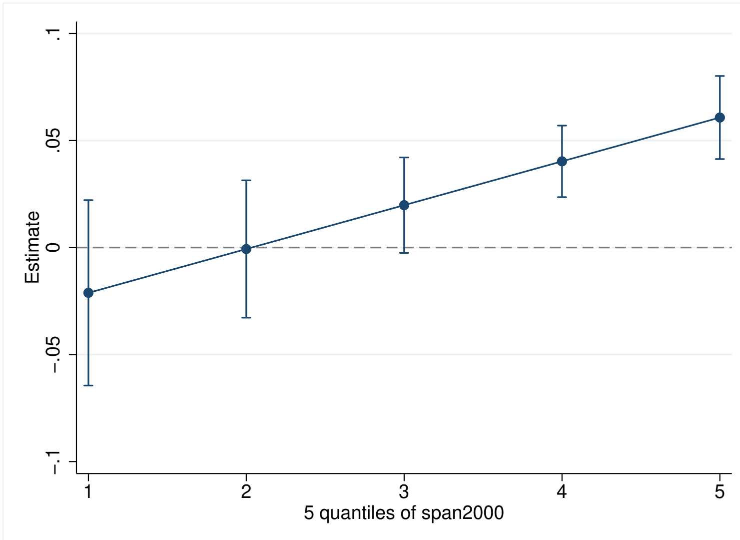

Figure 5: Predicted heterogeneous impacts of the PMC reform on PM2.5.

Notes: This figure shows the predicted impacts of the PMC reform on local PM2.5 concentration at different quantiles of the initial span of control of provincial governments, and their 95% confidence intervals. The prediction is based on the specification in column 5 of Table 9.

6.2 Understanding the heterogeneity

One may argue that this heterogeneity across the initial span of control of provincial governments might result from the initial difference in the manufacturing sector. This concern could be reasonable because any heterogeneity of the impact on local air quality across initial span of control could carry some heterogeneity of the impact across initial pollution or initial polluting capacity, if the initial span of control and the initial pollution or polluting capacity happen to be correlated with each other. To address this concern, I add (i) the interaction term between the PMC reform dummy and the initial number of manufacturing firms in each county, and (ii) the interaction term between the PMC reform dummy and the initial share of manufacturing output in GDP for each county, to the specification in column 5 of Table 9. The results are shown in Table 10. For comparison, the result in column 5 of Table 9 is listed in column 1 of Table 10. We can find that the estimated coefficients of the interaction term between the PMC reform dummy and initial span of control are always robust. Therefore, the heterogeneous impact of the PMC reform on local PM2.5 across the initial span of control of provincial governments cannot be explained by the factors related to the manufacturing sector.

Table 10: Understanding the heterogeneity.

Dependent variable	*log(PM2.5)*			
	(1)	(2)	(3)	(4)
PMC	-0.036***	-0.023	-0.026	0.015
	(0.010)	(0.020)	(0.018)	(0.033)
PMC×Span2000	0.005***	0.005***	0.005***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
PMC×log(firm_number2000)		-0.004		-0.008
		(0.005)		(0.005)
PMC×log(manufacturing_share2000)			0.008	0.020
			(0.013)	(0.014)
County fixed effect	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Treatment trend	Yes	Yes	Yes	Yes
Selection criteria×Year dummy	Yes	Yes	Yes	Yes
CPE dummy	Yes	Yes	Yes	Yes
Adj. R-squared	0.936	0.936	0.936	0.936
Obs.	21,180	21,074	21,139	21,074

Notes: This table shows whether the heterogeneity effects across initial span of control of provincial governments on local PM2.5 concentration can be explained by the initial difference in the manufacturing sector. Column 1 replicates the results in column 5 of Table 9 to facilitate comparison. $Span2000$ measures the initial span of control of a provincial government. $firm\_number2000$ is the initial number of manufacturing firms in a county. $manufacturing\_share2000$ is the initial ratio of manufacturing value-added of a county to its GDP. Weather controls include annual average temperature, precipitation and ventilation coefficient for each county. Treatment trend refers to treatment-specific linear time trends ($treatment_c * t$). Selection criteria includes county-level city, national poor county, major food-producing county, provincial boundary county, altitude, average slope, fiscal gap, and urbanization rate. CPE is a dummy variable indicating whether a county government carried out County-Power-Expansion reform in the sample period. Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6.3 Heterogeneous impacts of the PMC reform on economic performance

In this subsection, I examine if the same heterogeneity pattern applies to economic performance. Table 11 presents the results. The impact of the PMC reform on counties' economic performance also shows a strong heterogeneity across the initial span of control of provincial governments, as with its impact on local PM2.5 concentration. We can find that the impact of adopting the PMC reform on GDP per capita is more negative for counties of which the provincial governments had greater initial span of control. Figure 6 shows the predicted relative impacts of the PMC reform on economic performance with 95% confidence intervals. Although the relative impact of the PMC reform on economic performance is positive when the initial span of control of provincial governments ranks at the first 20%, this effect is not statistically significant. As the initial span of control increases, the relative impact of the PMC reform on economic performance becomes negative and increases in magnitude.

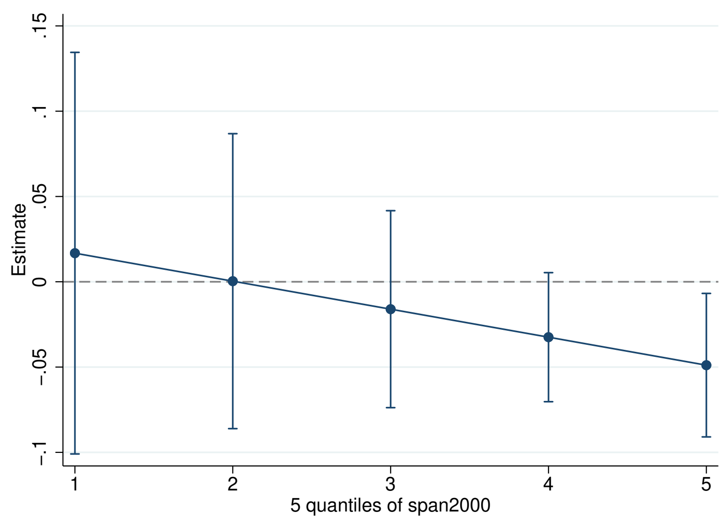

Figure 6: Predicted heterogeneous impacts of the PMC reform on economic performance.

Notes: This figure shows the predicted impacts of the PMC reform on GDP per capita at different quantiles of the initial span of control of provincial governments, and their 95% confidence intervals. The prediction is based on the specification in column 4 of Table 11.

Taken together, the results for both PM2.5 and GDP per capita indicate the important

Table 11: Heterogeneous impacts of the PMC reform on counties' economic performance across the initial span of control of provincial governments.

Dependent variable	*log(gdp_pc)*			
	(1)	(2)	(3)	(4)
PMC	0.033	0.079***	0.062***	0.052**
	(0.022)	(0.021)	(0.021)	(0.021)
PMC×Span2000	-0.010***	-0.009***	-0.008***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)
County fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Treatment trend		Yes	Yes	Yes
Selection criteria×Year dummy			Yes	Yes
CPE dummy				Yes
Adj. R-squared	0.943	0.943	0.946	0.946
Obs.	23,491	23,491	23,491	23,491

Notes: This table estimates the heterogeneous impacts of the PMC reform on counties' GDP per capita (in logarithm). GDP data is deflated using Beijing as the base province and 1999 as the base year. $Span2000$ measures the initial span of control of a provincial government. Treatment trend refers to treatment-specific linear time trends ($treatment_c * t$). Selection criteria includes county-level city, national poor county, major food-producing county, provincial boundary county, altitude, average slope, fiscal gap, and urbanization rate. CPE is a dummy variable indicating whether a county government carried out County-Power-Expansion reform in the sample period. Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

role of initial span of control in understanding the negative impacts of the PMC reform. Greater initial span of control of provincial governments makes the impacts of the PMC reform on environmental quality and economic performance more negative. This evidence motivates a further discussion on the potential mechanisms.

7 Potential Mechanisms

7.1 PMC reform and county government expenditures

To shed further light on the undesirable impacts of the PMC reform on environmental and economic outcomes, and the heterogeneity patterns in the previous section, I examine the relationship between the PMC reform, initial span of control and county government expenditures in this subsection.[6] Specifically, I focus on county-level aggregate expenditures and three specific government spending: (i) productive investments—expenditures for capital construction, and expenditures supporting rural production, agriculture, forestry, water management and meteorology, (ii) administrative expenditures, and (iii) environmental protection expenditures.

It should be noted that the county-level expenditure data is only publicized up to 2007. In addition, since in 2007 the central government implemented a substantial reform to the classifications of government revenues and expenditures, the items shown on the statistical yearbook for 2007 are different from those for previous years. One salient change is that environmental protection expenditures of county governments are listed separately in 2007, while this information is not available before that. Therefore, for aggregate expenditures, productive investments and administrative expenditures, the sample period is from 2000 to 2006. For environmental protection expenditures, the sample period is 2007.

Table 12 reports the results. Columns 1, 3, 5 and 7 show the impacts of the PMC reform on county government expenditure per capita. In column 1, the estimated coefficient of PMC is significantly negative, implying that the PMC reform decreased county governments' aggregate expenditures. In terms of specific expenditure items, we can see that the estimated coefficient of PMC is significantly positive for administrative expenditures (adm_pc), while that for productive investments ($grth\_pc$) and environmental protection expenditures (env_pc) are significantly negative. These results suggest that the PMC reform increased administrative expenditures but decreased expenditures on economic development. The PMC reform was also associated with the decrease in county-level environmental protection expenditures.

Columns 2, 4, 6 and 8 show the heterogeneous impacts of the PMC reform on county

[6] I focus on county government expenditures for two reasons. First, as mentioned in the background section, the PMC reform flattens government hierarchies in terms of fiscal relations. The outcome variables in this paper are more relevant to the expenditure side, rather than the revenue side. Second, almost all the literature has shown that county government's fiscal revenue increases after the reform (see Li et al. (2016) for a comprehensive analysis on transfer and revenue), so it is unlikely that more fiscal revenue leads to the undesirable outcomes found in this paper.

government expenditure per capita across the initial span of control of provincial governments. For aggregate expenditure, the estimated coefficient of the interaction term ($PMC * Span2000$) is negative but not statistically significant, suggesting no strong heterogeneity across initial span of control. But the coefficient of the interaction term is significantly positive for administrative expenditures, and is significantly negative for productive investments and environmental protection expenditures. From these results, we can know that in provinces with greater initial span of control, the PMC reform led to even more administrative expenditures, and even fewer expenditures on economic development at the county level. In addition, greater initial span of control was also associated with fewer county-level environmental protection expenditures.

Table 12: Impacts of the PMC reform on county government expenditures and the heterogeneity pattern.

Dependent variable	*log(totalexp_pc)*		*log(adm_pc)*		*log(grth_pc)*		*log(env_pc)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PMC	-0.077***	-0.042	0.087***	-0.057	-0.066*	0.127	-0.318***	-0.381***
	(0.014)	(0.031)	(0.019)	(0.050)	(0.036)	(0.091)	(0.073)	(0.083)
PMC×Span2000		-0.004		0.004***		-0.020**		-0.013**
		(0.003)		(0.001)		(0.009)		(0.006)
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes		
Treatment trend	Yes	Yes	Yes	Yes	Yes	Yes		
Selection criteria×Year dummy	Yes	Yes	Yes	Yes	Yes	Yes		
CPE dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.930	0.930	0.921	0.922	0.871	0.871	0.256	0.257
Obs.	8,460	8,460	8,460	8,460	8,460	8,460	1,284	1,284

Notes: This table shows the effects and the heterogeneity pattern of the PMC reform on counties' total expenditures, and expenditures on administration, economic development and environmental protection. All expenditures (measured in per capita term) are in logarithm and are deflated using Beijing as the base province and 2000 as the base year. $Span2000$ measures the initial span of control of a provincial government. Due to the availability of expenditure data, the sample period for administration expenditures and productive investments is from 2000 to 2006, while the sample period for environmental expenditures is 2007. Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

From the previous analyses, we know that the PMC reform generated negative impacts on reformed counties' environmental and economic outcomes, and these impacts are more negative for counties in provinces which had greater initial span of control. The results in this subsection provide evidence on the changes in county government expenditures which is consistent with the previous findings: the PMC reform had negative effects on county governments' expenditures on environmental protection and economic development, and these effects were more negative for counties in provinces with greater initial span of control. Taken together, all these results are consistent with the reading that the impacts of the PMC reform work through county government expenditures and that is related to the initial span of control of provincial governments. In addition, since administrative expenditures are typically used as the proxy for on-the-job consumption (Wang and Xu, 2017),[7] and existing literature has shown that county-level amount of misused fiscal funds increased after the PMC reform (Bo et al., 2020),[8] one interpretation for the results here could be that more fiscal funds were diverted from environmental protection and economic development to on-the-job consumption due to the less effective monitoring caused by the PMC reform and this diversion was more severe for counties in provinces with greater initial span of control.

7.2 PMC reform and firm-level pollution

In this subsection, I further investigate the source of PM2.5. Because the combustion of fossil fuel has been shown by scientific researches to be a large contributor to the PM2.5 pollution (Chow and Watson, 2002), I examine if firm-level consumption of fuel coal and fuel oil, as well as the smoke generated are affected by the PMC reform.

To this end, I match firm-level data with county PMC status according to firms' address. The outcome variables of interest are fuel coal consumption per unit output ($coal\_p$), fuel oil consumption per unit output (oil_p), untreated smoke per unit output ($smoke\_p$) and the number of exhaust treatment facilities ($facility$). I also controls for firm characteristics including total assets, firm age and the type of ownership (i.e., state-owned enterprises,

[7] In China's government budget report, there is no separate item measuring on-the-job consumption. Therefore, administrative expenditures are usually used as the proxy for on-the-job consumption. In practice, it is quite common that local officials get reimbursement from administrative funds after on-the-job consumption occurred (Meng, 2009).

[8] In Bo et al. (2020), they find that the county-level amount of misused fiscal funds detected by auditors increased after the PMC reform. Misuse of funds refer to the diversion of earmarked funds by the central government for specific projects to other projects, and the diversion of funds from the budget plans approved by the local People's Congress at the beginning of each year.

private enterprises, foreign enterprises, and collectively-owned enterprises). The results are shown in Table 13. As we can see, the estimated coefficients of PMC are not significant in columns 1 and 3, but are significantly positive in columns 5 and significantly negative in column 7, respectively. These results suggest that although firms in the reformed counties did not significantly change their fossil fuel consumption, their untreated smoke emission increased due to decreased number of exhaust treatment facilities. The heterogeneity analysis presented in columns 2, 4, 6 and 8 further shows that untreated smoke increased more and exhaust treatment facilities decreased more for firms in reformed counties of which the provincial governments had greater initial span of control.

Taken together the evidence on county government's expenditures and the change in firm pollution behaviors, we can infer that since reformed counties spent less on environmental protection, stuff funding and equipment procurement related to environmental enforcement could be negatively affected. Therefore, the effectiveness and intensity of environmental monitoring and enforcement over firms' exhaust emission were weakened. To reduce costs, firms spent less on exhaust treatment and thus discharged more untreated smoke. In counties where the effects of the PMC reform were more negative due to greater initial span of control of provincial governments, the effectiveness of environmental enforcement was further weakened, so firms generated more pollution. These evidence provide a possible channel to understand how the PMC reform led to a deterioration in local air quality.

8 Conclusion

Using the PMC reform starting in 2003 in China as a natural experiment, this paper investigates the unintended impact of fiscal hierarchy reforms on local air quality. The results indicate that local air quality measured by annual average PM2.5 concentration was negatively affected by the PMC reform. This negative effect was more pronounced in the long run. The deterioration in air quality was not caused by better economic performance, since

Table 13: The impacts of the PMC reform on firm-level pollution and the heterogeneity patterns.

Dependent variable	*log(coal_p)*		*log(oil_p)*		*log(smoke_p)*		*log(facility)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PMC	0.606	-0.784	0.006	0.023	0.195***	0.318	-0.125**	0.023
	(1.902)	(3.560)	(0.085)	(0.210)	(0.040)	(3.090)	(0.059)	(0.107)
PMC×Span2000		0.014		-0.001		0.091***		-0.165***
		(0.213)		(0.013)		(0.010)		(0.032)
Firm characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-squared	0.125	0.124	0.704	0.704	0.124	0.146	0.230	0.230
Obs.	36,104	36,104	36,680	36,680	49,770	49,770	34,911	34,911

Notes: This table shows the effects and the heterogeneity pattern of the PMC reform on firm fossil fuel consumption and smoke treatment. All outcome variables are in logarithm. $Span2000$ measures the initial span of control of a provincial government. Firm characteristics include total assets, firm age and ownership (such as state-owned enterprises, private enterprises, foreign enterprises, and collectively-owned enterprises). Standard errors are clustered at the county level and reported in the parentheses below the estimated coefficients. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

the GDP per capita of reformed counties also witnessed a significant drop. For counties in provinces with greater initial span of control, these effects were more negative. Further analysis shows that these undesirable consequences of the PMC reform were accompanied by the changes in county government expenditures. The PMC reform reduced county-level aggregate expenditures and expenditures on environmental protection and economic development, but increased administrative expenditures. The impact on administrative expenditure was more positive, while the impacts on environmental protection and economic development expenditures were more negative for counties in provinces with greater initial span of control. This could result from the problem that less effective monitoring became worse when the initial span of control of provincial governments was greater. Less environmental spending weakened environmental monitoring and enforcement, so firms in reformed counties used fewer exhaust treatment facilities and emitted more untreated smoke, which resulted in deteriorated local air quality.

It should be admitted that due to the availability of data, the mechanism that county government expenditures change after the PMC reform just reveal one possible channel linking the PMC reform and its undesirable consequences. Other alternative mechanisms still need to be investigated if relevant data can be available in the future. For example, how the coordination between counties in terms of environmental protection was affected by the PMC reform.

References

Beck, T., Levine, R., and Levkov, A. (2010). Big bad banks? the winners and losers from bank deregulation in the united states. *The Journal of Finance*, 65(5):1637–1667.

Bo, S. (2021). Environmental regulations, political incentives and local economic activities: Evidence from china. *Oxford Bulletin of Economics and Statistics*, 83(3):812–835.

Bo, S., Wu, Y., and Zhong, L. (2020). Flattening of government hierarchies and misuse of public funds: Evidence from audit programs in china. *Journal of Economic Behavior & Organization*, 179:141–151.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Calvo, G. A. and Wellisz, S. (1978). Supervision, loss of control, and the optimum size of the firm. *Journal of Political Economy*, 86(5):943–952.

Chen, C. and Suen, W. (2019). The comparative statics of optimal hierarchies. *American Economic Journal: Microeconomics*, 11(2):1–25.

Chen, Y. J., Li, P., and Lu, Y. (2018). Career concerns and multitasking local bureaucrats: Evidence of a target-based performance evaluation system in china. *Journal of Development Economics*, 133:84–101.

Chow, J. C. and Watson, J. G. (2002). Review of pm2. 5 and pm10 apportionment for fossil fuel combustion and other sources by the chemical mass balance receptor model. *Energy & Fuels*, 16(2):222–260.

Garicano, L. and Van Zandt, T. (2012). 15. hierarchies and the division of labor. In *The handbook of organizational economics*, pages 604–654. Princeton University Press.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

He, G., Liu, T., and Zhou, M. (2020a). Straw burning, pm2. 5, and death: evidence from china. *Journal of Development Economics*, 145:102468.

He, G., Wang, S., and Zhang, B. (2020b). Watering down environmental regulation in china. *The Quarterly Journal of Economics*, 135(4):2135–2185.

He, G., Xie, Y., and Zhang, B. (2020c). Expressways, gdp, and the environment: The case of china. *Journal of Development Economics*, 145:102485.

Hering, L. and Poncet, S. (2014). Environmental policy and exports: Evidence from chinese cities. *Journal of Environmental Economics and Management*, 68(2):296–318.

Hooghe, L., Marks, G., Schakel, A. H., Osterkatz, S. C., Niedzwiecki, S., and Shair-Rosenfield, S. (2016). *Measuring regional authority: A postfunctionalist theory of governance, Volume I.* Oxford University Press.

Jacobson, L. S., LaLonde, R. J., and Sullivan, D. G. (1993). Earnings losses of displaced workers. *The American economic review*, pages 685–709.

Jia, R. (2017). Pollution for promotion. *21st Century China Center Research Paper*, (2017-05).

Kahn, M. E., Li, P., and Zhao, D. (2015). Water pollution progress at borders: the role of changes in china's political promotion incentives. *American Economic Journal: Economic Policy*, 7(4):223–42.

Li, P., Lu, Y., and Wang, J. (2016). Does flattening government improve economic performance? evidence from china. *Journal of Development Economics*, 123:18–37.

Ma, G. and Mao, J. (2018). Fiscal decentralisation and local economic growth: Evidence from a fiscal reform in china. *Fiscal Studies*, 39(1):159–187.

Meng, Q. (2009). The traceability, current situation and improvements on china's on-thejob consumption system. *Finance and Trade Economics*, (12):77–81 (in Chinese).

Miguel, E. and Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1):159–217.

Qian, Y. (1994). Incentives and loss of control in an optimal hierarchy. *The Review of Economic Studies*, 61(3):527–544.

Rohde, R. A. and Muller, R. A. (2015). Air pollution in china: mapping of concentrations and sources. *PloS one*, 10(8):e0135749.

Simon, H. A. (2013). *Administrative behavior*. Simon and Schuster.

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):334–356.

Wang, X. and Xu, J. (2017). Does the "province managing county" reform discipline the lower-level civil servants' position-related consumption? *Finance and Trade Economics*, 38(6):17–32 (in Chinese).

Zheng, S., Zhang, X., Sun, W., and Wang, J. (2019). The effect of a new subway line on local air quality: A case study in changsha. *Transportation research part D: transport and environment*, 68:26–38.