Università Commerciale "Luigi Bocconi"

PhD School

PhD program in: Economics and Finance

Cycle: 35

Disciplinary Field (code): SECS-P/01

Essays on the Political Economy of Development

Advisor: Guido Tabellini

Co-Advisor: Stefano Fiorin

PhD Thesis by

Yongwei Nian

ID number: 3111127

Year: 2024

Abstract

This thesis consists of three chapters on the political economy of development. The regional focus is on China, but the issues studied are widespread around the world. In Chapter 1, I study the manipulation of economic statistics by government officials. Leveraging a unique reform in China, I show that a combination of top-down monitoring and punishment could effectively reduce such manipulation; furthermore, reducing such manipulation could elicit officials' effort in developing the economy, as manifested in both policy changes and downstream impacts. In Chapter 2, I focus on agricultural fires, a major source of air pollution in the rural areas induced by farmers burning crop residues after harvest. I show that providing economic incentives to farmers could effectively reduce such burnings, while a command-and-control that merely forbids such burnings fails to do so. This work has been published in the *Journal of Development Economics* (March 2023). In Chapter 3, joint with Chunyang Wang (Peking University), we examine a widespread practice of interregional rotation of local officials in China, whose initial purpose is to curb corruption. We document an interesting pattern of "go with the politician": firms follow politicians' move to purchase land in the politicians' new constituents, with cheaper prices but lower land usage efficiency after purchase. This pattern thus indicates the continuation of corruption after such rotation. This work has been published in the *American Economic Journal: Economic Policy* (May 2023).

Acknowledgements

I am deeply indebted to my advisors, Professors Guido Tabellini and Stefano Fiorin, for their invaluable guidance and instruction throughout my PhD journey. Their exceptional mentoring was crucial to the completion of this thesis. My interactions with them have always been intellectually stimulating and deeply inspiring.

I extend my sincere thanks to my committee member, Professor Paolo Pinotti, and to the numerous professors at Bocconi University who have enriched my work with insightful comments and suggestions. In particular, I would like to thank Jan Bakker, Luca Braghieri, Alexia Delfino, Erika Deserranno, Sarah Eichmeyer, Rafael Jimenez-Duran, Massimo Morelli, Marco Ottaviani, and Carlo Schwarz.

My special thanks also go to the PhD course coordinators, Mariano Massimiliano Croce and Marco Ottaviani, and to my colleagues, for their great guidance and support in navigating the PhD program.

Finally, I wish to express my deepest gratitude to my family for their unwavering support and understanding throughout my pursuit of a PhD. Their constant encouragement has been a cornerstone of my academic journey.

Contents

Chapter 1: Curbing Bureaucratic Information Manipulation 1

Chapter 2: Incentives, Penalties, and Rural Air Pollution 133

Chapter 3: Go with the Politician 278

Chapter 1

Curbing Bureaucratic Information Manipulation: Evidence from a Statistical Reform in China

Yongwei Nian

Abstract

Bureaucrats are often incentivized to manipulate information, which may have real consequences. Leveraging China's 2009 reform punishing economic data manipulation and counties' quasi-random reform exposure, this paper provides rich causal evidence showing that the reform led to: (1) a decrease in GDP growth manipulation amounting to 5% of reported GDP growth, driven by a *reputational discipline* effect; (2) an increase in local officials' development effort manifested in both policy changes and downstream impacts, consistent with an *effort reallocation* effect. These results thus highlight the far-reaching costs of such manipulation and the welfare implications of curbing it.

JEL classification: H70, O10, P20
Keywords: GDP manipulation; Information distortion; Bureaucrats; China

1 Introduction

Bureaucracies are pivotal to economic development, a view that dates back at least to Max Weber and is further stressed by a growing body of economics literature (Besley et al., 2021). However, lower-level bureaucrats may have incentives to manipulate information, especially when it is instrumental to their career advancement but imperfectly observed, thereby undermining bureaucratic effectiveness. One example is the overreporting of economic statistics by local officials in China, where their career advancement is closely tied to economic performance (Wallace, 2016; Xiong, 2018; Lyu et al., 2018). Examples outside China and the domain of economic statistics also abound (Bossuroy, Delavallade and Pons, 2019; Martinez, 2022; Kofanov et al., 2023). Such manipulation may not only create information distortion in the bureaucracy, but may also divert bureaucrats' effort from their designated missions from a multitasking perspective (Holmstrom and Milgrom, 1991). Despite its persuasiveness and potential costs, there is scant evidence on how to effectively curtail such manipulation, and even less is known about the economic benefits of potential interventions.

This paper sheds some of the first light on these issues leveraging China's 2009 reform aimed at combating economic data manipulation by local officials. The reform, initiated by the National Bureau of Statistics (NBS) and other central authorities, increased punishment for manipulation nationwide. However, the detection mainly relied on pre-deployed survey teams in 40% of counties, creating cross-sectional variations in counties' exposure to the reform. When manipulation was detected, the survey teams could directly report to the central authorities, typically leading to reputational warnings or demerits as punishment. Notably, these survey teams were deployed in 2005 and initially tasked with economic surveys, making their deployment largely unrelated to local data manipulation. This unique setup reduces concerns of positive selection common in government interventions (Al-Ubaydli, List and Suskind, 2019; Wang and Yang, 2021). In addition, the teams operated independently from local political influence under the leadership of the NBS, a key feature for effective monitoring (Kofman and Lawarrée, 1993; Olken, 2007; Vannutelli, 2022).

To identify the effects of the reform, I assemble a county-level dataset from 2005 to 2018 and employ a difference-in-differences design comparing counties with these survey teams (treatment) to counties without these survey teams (control) before and after 2009. I focus on the manipulation of GDP growth given its central role in dictating local officials' career advancement (Li and Zhou, 2005). I find that the reform significantly decreased GDP growth manipulation, which is measured as the discrepancy between GDP growth reported by local statistics bureaus and nighttime light intensity growth inferred from satellite observations, following the seminal framework of Henderson, Storeygard and Weil (2012). In terms of economic magnitude, in the most stringent specification with a set of baseline demographic, economic, and geographic controls interacted with the post-reform dummy, I find a 0.58 percentage points drop in GDP growth manipulation in treatment counties relative to control counties, which amounts to 5.3% of the mean of reported GDP growth. I further find no pre-trends in manipulation between treatment and control counties using an event study specification, which is consistent with the pre-deployment nature of the survey teams.

I conduct various tests to ensure the robustness of the baseline findings. First, in the spirit of Martinez (2022), I show that the estimates on GDP growth manipulation remain virtually unchanged when I allow the mapping between economic growth and light growth to be nonlinear or depend on various temporal and spatial characteristics. Second, I show that the results are robust to accounting for potential imbalance between treatment and control counties, by achieving covariate balance through entropy balancing Hainmueller (2012) or coarsened exact matching (Iacus, King and Porro, 2012), among other conventional methods. Third, I conduct a placebo test by examining the dynamic effect around 2005 when the survey teams were deployed but had not conducted any disciplining actions. This test helps to further rule out pre-trends and alleviate additional concerns that these teams per se may affect my outcomes, even in the absence of the reform in 2009. Reassuringly, I find no diverging trends between treatment and control counties until 2009.

To further address concerns about unobserved heterogeneity, I leverage a unique institutional feature to conduct an instrumental variable estimation. By checking local gazetteers, I find that most of these survey teams launched in 2005 were restructured from earlier rural survey teams set up in 1984. As I describe later, these earlier teams were mainly used to collect information on agricultural output and their assignment at that time was done through a systematic random sampling of counties within provinces. This random assignment hence forms a valid instrument for counties with survey teams deployed in 2005. The validity of this instrument is further supported by balance tests showing that counties with these earlier rural survey teams were no different from other counties on a battery of baseline county characteristics. In addition, counties with these earlier rural survey teams were in parallel with other counties in terms of GDP growth manipulation until 2009. The difference-in-differences estimates using this instrument have no substantial changes, suggesting that the baseline findings are unlikely to be driven by unobserved heterogeneity between treatment and control counties.

I investigate several potential mechanisms that could account for the drop in GDP growth manipulation, following the predictions from a conceptual framework. Specifically, the reform could decrease GDP growth manipulation through both a *discipline* effect and a *selection* effect. The discipline effect refers to behavioral changes within local officials' terms due to reputational and promotional costs, while the selection effect stems from the removal of those involved in manipulation or the appointment of more competent successors. To distinguish between these two mechanisms, I focus on a sample in which there was no change of local officials in the treatment counties. The estimates are similar to those in the full sample. I also show that the reform had no effects on personnel turnovers and official characteristics. Hence, the findings can only be explained by the reform generating a discipline effect on local officials. This discipline effect could be further decomposed into a reputational discipline effect and a promotional discipline effect. To disentangle these two mechanisms, I show that the drop in GDP growth manipulation is no different for local officials with greater promotion incentives. Hence, the findings so far are most consistent with a reputational discipline effect caused by the exposure of manipulation within the bureaucracy upon detection. This result resonates with a growing literature showing both theoretically and empirically the role of reputational concerns in shaping truth-telling (Benabou and Tirole, 2011; Gneezy, Kajackaite and Sobel, 2018; Abeler, Nosenzo and Raymond, 2019). I also rule out some alternative explanations. First, I show that the reduction in manipulation was not driven by potential soft information acquired by the survey teams, which may dampen the role of GDP growth in performance evaluation. Second, I show that the reduction in manipulation was not driven by improvements in local statistical capacity, which may close the gap between reported GDP growth and light growth. Finally, I show that the findings did not capture the effects of other concurrent reforms that may also strengthen the monitoring of local officials, such as the anti-corruption inspections.

I then test the effect on local officials' effort in developing the economy. Conceptually, local officials should allocate more effort into developing the economy given the relatively higher cost of GDP growth manipulation after the reform. Along this line, I first show through a textual analysis of county government work reports that government policies shifted in directions conducive to economic growth after the reform. In particular, I find that local officials put more greater emphasis on business attraction and market reform, which are critical to China's recent economic success (Xu, 2011). To alleviate concerns about cheap talk in these reports and shed light on whether these policy shifts translated into any real impacts, I further examine two downstream economic outcomes closely related to local officials' effort. The first is bank credit, over which local officials have substantial discretion (Ru, 2018; Cong et al., 2019). I show that the reform generated positive impacts on various measures of bank credit, and such effects were more pronounced for loans to small firms and credit from banks controlled by local governments. The second is firm entry, which is subject to lengthy bureaucratic procedures in China (WorldBank, 2008; Jia, Lan and Padró i Miquel, 2021). I also find that the reform boosted firm entry, especially for those with higher productivity. Finally, I investigate citizens' trust in local officials and evaluation of local government performance, two common measures of the performance of government officials (Bertrand et al., 2020; Martinez-Bravo et al., 2022). I find that citizens' attitudes towards local officials improved after the reform; as a placebo, their trust in most people or evaluation of others' health had no changes. Overall, these results are consistent with local officials exerting more effort in economic development after the reform. I also rule out several alternative explanations. First, I show that local officials' differential performance did not result from varying fiscal transfers from the central government, which may have similar development effects. Second, I show that better policies adopted by local officials was not driven by the reform facilitating policy diffusion among counties or from upper-level governments. Third, I show that local officials did not switch to other short-termist behaviors that may boost the economy in the short run, such as overleverage. Finally, using both corruption convictions and perceived corruption by citizens, I show that their improved performance was not due to a decrease in corruption due to perceived higher corruption cost after the reform.

This paper relates to several strands of literature. First, it contributes to the nascent literature documenting the manipulation of statistical data by career-minded government officials across various domains and political regimes. A common form of such manipulation is the inflation of economic statistics in authoritarian countries (Wallace, 2016; Lyu et al., 2018; Xiong, 2018; Martinez, 2022). Additionally, such manipulation exists in other settings, such as the under-reporting of pandemic statistics in Russia (Kofanov et al., 2023), the falsification of regional electricity consumption bills in India (Mahadevan, 2023), and the manipulation of air pollution data in China (Ghanem and Zhang, 2014; Greenstone et al., 2022). Despite the pervasiveness, there is surprisingly little evidence on how to curb such manipulation. One possible reason is the underestimation of the potential costs. In certain cases, the costs are obvious, such as profit losses by utility providers (Mahadevan, 2023) or the underinvestment of defensive goods against pollution by citizens (Greenstone et al., 2022). In the domain of economic data manipulation, the costs are not obvious. My paper thus advances this literature in two ways. First, to my best, this is the first to estimate the causal impacts of a large-scale intervention on the exaggeration of GDP statistics by local officials, a typical form of such manipulation. Second, the rich micro-level data allows me to further estimate the downstream impacts of this intervention. By highlighting the real consequences of such manipulation, this paper thus has implications for the design of relevant interventions targeting economic data manipulation, which are still scarce.

Second, this paper also adds to a growing literature on disciplining local officials. In the spirit of Becker and Stigler (1974), existing research shows that the combination of monitoring and punishment could reduce unaccounted expenditures in Indonesian villages (Olken, 2007), curb the misuse of federal funds in Brazilian municipalities (Avis, Ferraz and Finan, 2018), and enhance fiscal performance in Italian local governments (Vannutelli, 2022). Underpinning the success of such practices are two noteworthy features. On the monitoring side, independent audit in a top-down fashion is critical. Lack of independence may lead to collusion between monitors and local officials (Kofman and Lawarrée, 1993). My study features such a setting where the survey teams were deployed from the central and insulated from local interference. On the punishment side, extant research highlights the role of electoral and judicial punishment. My study differs from this, however, by showing that the findings are most consistent with a reputational discipline effect. In this vein, my findings join a small literature showing the role of reputational concerns in shaping truthtelling (Benabou and Tirole, 2011; Gneezy, Kajackaite and Sobel, 2018; Abeler, Nosenzo and Raymond, 2019). This feature has implications for many settings plagued by bureaucratic misconduct, where effective accountability mechanisms are either lacking or difficult to enforce. As far as I know, such reputational discipline effect is rarely explored at the local official level.

Third, this paper relates to the literature estimating the real impacts of corruption. The distortions created by such misconduct have long been recognized (Krueger, 1974; Shleifer and Vishny, 1993), but well-identified empirical evidence is still inadequate due to the difficulty in measuring corruption (Banerjee, Mullainathan and Hanna, 2012). As such, the literature mostly estimates the real impacts of relevant interventions targeting corruption. Extant research in this vein finds that curbing corruption improves resource allocation and spurs economic activity (Giannetti et al., 2021; Colonnelli and Prem, 2022). My paper studies a less examined aspect of bureaucratic misconduct—the manipulation of economic statistics. By diverting local officials' effort from economic development, such misconduct may be similarly costly. Curbing such misconduct thus improved local officials' performance and spurred economic activities. In some domains such as credit allocation and firm entry, the impacts are comparable to those of anti-corruption campaigns as found both in China and Brazil. Hence, manipulation of economic statistics should be given similar emphasis as corruption, although in reality its salience is ignored.

Finally, while this paper focuses on the bureaucracy, the manipulation of information by local officials echoes earnings manipulation by corporate managers (Stein, 1989; Fischer and Verrecchia, 2000; Goldman and Slezak, 2006; Benmelech, Kandel and Veronesi, 2010; Agarwal, Daniel and Naik, 2011; Ma, Pan and Stubben, 2020). In this vein, the findings have implications beyond the bureaucratic setting. Different from corporate managers, local officials often wield broader influence over the local economy (Xu, 2011). Thus, their manipulation carries widespread welfare consequences.

The remainder of the paper is organized as follows. Section 2 introduces the institutional background of GDP growth manipulation, the reform, and a conceptual framework illustrating the impacts of the reform. Section 3 describes the data and discusses the empirical strategy. Section 4 presents the main results on GDP growth manipulation, robustness tests, and mechanisms. Section 5 investigates the impact on local officials' effort in developing the economy. Section 6 concludes.

2 Background and conceptual framework

2.1 Institutional background

In China, GDP is calculated using the value-added method; that is, by summing the value added in all sectors in a region, with the county as the lowest level of regions for GDP calculation. The calculation is done by the local statistical bureaus, which are controlled by local officials in terms of personnel and funding. It is worth noting that a Chinese county is co-led by two leaders: the party secretary controlling personnel and other political affairs, and the magistrate running the economy (Xu, 2011). However, both leaders are evaluated heavily on GDP growth and hence have incentives to manipulate (Yao and Zhang, 2015). These leaders have a couple of ways to manipulate GDP: directly asking local statistical bureaus to make up numbers, requiring firms to overstate income or pay additional "tax" and return later, or double counting firms' non-local subsidiaries, among others.

In terms of the statistical reform, it was initiated by the National Bureau of Statistics (NBS), joint with other central authorities, in May 2009 with the goal of disciplining misreporting of local governments in processing statistical data. It mainly targeted local officials who falsified statistical data by themselves, forced or instructed other agents to manipulate, retaliated against those detecting manipulation, or failed to find severe distortion in local statistical data. The last clause means that local officials were still punished even if there was no evidence of their direct manipulation, alleviating concerns about the local statistical bureau acting as scapegoats upon detection. In addition, other agents participating in manipulation, such as the staff in local statistical bureaus, were also punished. The punishment was enforced by the relevant authorities (local officials superiors and the supervisory organs) and mostly took the form of reputational warning or demerit. Specifically, if a local official was found to manipulate economic statistics, a warning or demerit would be issued within the bureaucracy. In serious circumstances, demotion or dismissal would be issued.[1] Nevertheless, no legal actions were specified, which differentiates this reform from the commonly used random audits in Western countries that carry legal consequences (Avis, Ferraz and Finan, 2018).

In terms of the detection, both the local statistical bureau in each county and centrally managed survey teams in some counties, which I will describe their deployment shortly, were responsible. However, the local statistical bureau shared aligned incentives with local officials because they were appointed and funded by local officials; in contrast, the survey teams had a higher probability of detecting manipulation, as they were appointed and funded centrally. As emphasized in the literature, this type of independence is the key to the effectiveness of monitoring (Kofman and Lawarrée, 1993; Olken, 2007; Vannutelli, 2022). Upon detection, the survey teams could directly report to the NBS, and the NBS would take actions together with other authorities. The main takeaway so far is that the survey teams, which only existed in part of the counties, create the key source of variation in each county's exposure to the reform, enabling me to identify the reform' effect through difference-in-differences identification strategy.

What is crucial to my difference-in-differences strategy is when and how the survey teams were deployed. They were deployed by the NBS in 2005 in 40% of counties. Their initial job was to conduct sampling surveys to collect information on CPI, household income, grain output, and micro-firm dynamics. As these variables were frequently published and updated by the NBS, the survey teams could lighten the workload of the generally understaffed local statistical bureaus. Starting from 2009 when the aforementioned reform was launched, these teams also began to detect the manipulation of statistical data. While the NBS did not officially reveal the criteria regarding the selection of counties with these survey teams, the initial goal—to generate nationally representatively information—suggests that the allocation of these survey teams to

[1] Given that demotion or dismissal of local officials was extremely rare in reality, these punishments were unlikely to be enforced, which is confirmed by my subsequent analysis in which I show the reform had no effect on personnel turnovers and official characteristics.

counties should be largely orthogonal to local economic data manipulation. Furthermore, by checking local gazetteers, I find that most of these survey teams launched in 2005 were restructured from earlier rural survey teams set up in 1984. At that time, these earlier rural survey teams were mainly used to collect information on agricultural output, and counties with them were selected randomly within a province. This unique feature could allow me to further address potential concerns about endogenous allocation, by utilizing exogenous variations from the random assignment of these earlier rural survey teams. I leave the detailed discussion of these earlier rural survey teams in Section 4.2.3, where they are used as an instrument. Unless explicitly noted, the survey teams refer to those launched in 2005 in my subsequent analysis.

2.2 Conceptual framework

In this part, I will leverage a simple economic tournament model to illustrate the sources of GDP growth manipulation under China's unique promotion rule and generate some testable predictions on the effect of the reform. Similar to the game setting in Lazear and Rosen (1981), I consider a single-period tournament without discounting. There are two county leaders indexed by $i = 1,2$ competing for promotion, which is decided by the principal (the upper-level government) based on their reported economic performance. Leader i can manipulate GDP growth with effort m_i and stimulate the economy with effort e_i, subject to a constraint $m_i + e_i \leq \bar{C}$.[2] Conceptually, the effort exerted in manipulating GDP growth involves time, energy, and resources spent on cooking the book or persuading potential dissenters, among others. The payoffs to these two types of effort are $h(\cdot)$ and $g(\cdot)$, respectively, which are increasing and concave. $h(\cdot)$ is concave because, at higher levels of manipulation, an additional unit increase in manipulation is more likely to be detected and to incur greater dissent, thus requiring more effort. Furthermore, I assume that to make the manipulation less detectable, a leader conducts manipulation simultaneously with stimulating the economy, instead of after observing true GDP growth at

[2] While the manipulation of GDP growth may also be done by the staff in local statistical bureaus, I only model the behaviors of local officials as local statistical bureaus are controlled by local officials and thus act in concert with local officials.

the end of the period.[3] Hence, the reported GDP growth is given by

$$G_i = h(m_i) + g(e_i) + \varepsilon_i, \quad \varepsilon_i - \varepsilon_{-i} \sim \mathrm{U}[-\frac{1}{2\phi}, \frac{1}{2\phi}]$$

where the reported GDP growth is the sum of the payoffs of both types of effort, plus an idiosyncratic shock ε_i. I assume the difference of the shocks between the two counties is uniformly distributed with mean 0 and density ϕ. Such distribution is known to all, but the realized values of the shocks are only known at the end of the period.

In addition, manipulating GDP would be detected with probability p, where $p \in (0,1)$ denotes the exogenous rate of identifying manipulation. As discussed in the previous section, a warning or demerit will be issued within the bureaucracy upon detection. For severe manipulation, a demotion or dismissal will be issued. Thus, the leader suffers from a direct reputational loss due to the exposure of manipulation within the bureaucracy, which takes the linear form of $\lambda h(m_i)$. This type of reputational cost in shaping truth-telling has both theoretical and empirical foundations (Benabou and Tirole, 2011; Gneezy, Kajackaite and Sobel, 2018; Abeler, Nosenzo and Raymond, 2019). In addition, the leader could also suffer from a promotion cost, which I will model in the promotion part. Finally, legal cost is not modelled as the reform involves no legal actions. Then leader i's payoff is given by:

$$U_i = \mathbb{1}_{\{i \text{ promoted}\}} u(R) + (1 - \mathbb{1}_{\{i \text{ promoted}\}}) u(r) + u(\Omega) - \mathbb{1}_{\{i \text{ detected}\}} \lambda h(m_i)$$

where $\mathbb{1}_{\{i \text{ detected}\}}$ is an indicator equal to 1 if leader i is detected for manipulation and 0 otherwise, and $\mathbb{1}_{\{i \text{ promoted}\}}$ is an indicator equal to 1 if leader i is promoted and 0 otherwise. The utility function $u(\cdot)$ is increasing and concave. Leader i receives reward R if promoted and r if not, where $R \gg r > 0$.[4]. To generate sharp predictions on GDP growth manipulation, I abstract from corruption by assuming that leader i extracts a fixed amount of rents Ω from the current office.[5]

[3]This is also supported by anecdote evidence showing that local officials asked firms to overstate income in the middle of a year.

[4]The positive reward r captures the fact that in China most local officials would still serve in a similar position even if not promoted, instead of exiting the bureaucracy.

[5]I support such assumption by empirically showing that the reform did not affect corruption.

Promotion rule The promotion rule posits that the principal promotes the county leader with the highest reported GDP growth. This rule has a widely-acknowledged theoretical foundation, as the high comparability across subnational units in China makes economic tournaments particularly suitable for promoting regional leaders (Maskin, Qian and Xu, 2000). It is also verified by a growing literature showing that GDP growth is positively related to local officials' career advancement at various levels of governments (Li and Zhou, 2005; Xu, 2011; Jia, Kudamatsu and Seim, 2015; Landry, Lü and Duan, 2018). In addition, scholarship on political selection in China also emphasizes the role of connections with upper-level leaders (Shih, Adolph and Liu, 2012; Jia, Kudamatsu and Seim, 2015; Meyer, Shih and Lee, 2016). I abstract from this factor as it mainly matters for promotion at higher level. In particular, Landry, Lü and Duan (2018) show that at the county level, GDP growth has a significant impact on promotion while political connections do not. To capture the potential punishment on promotion upon detection of manipulation, I assume that the principal subtracts an amount of $\delta h(m_i)$ from a leader's reported GDP growth. Here, δ represents a promotional punishment, but how large it is remains an empirical question. So leader i is promoted if

$$G_i - \mathbb{1}_{\{i \text{ detected}\}} \delta h(m_i) > G_{-i} - \mathbb{1}_{\{-i \text{ detected}\}} \delta h(m_{-i})$$

Timing The timing of events in this tournament is summarized as follows:

1. Both leaders simultaneously choose effort in manipulating GDP growth and stimulating the economy, before knowing the realization of ε_i.

2. ε_i is realized and all uncertainty is resolved.

3. The principal detects manipulation, punishes the involved, and makes promotion decision based on the aforementioned promotion rule.

Equilibrium The equilibrium concept is a pure strategy Nash equilibrium.[6] To solve it, note that leader i maximizes expected payoff taking leader $-i$'s choice as given. As shown in Ap-

[6] I focus on pure strategies as it is empirically obscure to interpret mixed strategies in manipulating GDP growth and stimulating the economy in a static game.

pendix A, through usual maximization (assume interior solution), one can solve for equilibrium m^* (effort in manipulating GDP growth) and e^* (effort in stimulating the economy):

$$\begin{aligned} m^* &= K\left[\frac{V\phi}{V\phi(1-p\delta)-\lambda p}\right] \\ e^* &= \bar{C} - m^* \end{aligned}$$

where $K(\cdot)$ is the inverse function of $h'(\cdot)/g'(\bar{C}-\cdot)$. V is equal to $u(R)-u(r)$, which measures the utility gains from promotion.

The reform's effect Conceptually, the reform could decrease local leaders' effort in GDP growth manipulation m^* in three ways. First, by exposing manipulation upon detection, the reform could increase local leaders' reputational cost of manipulation, which is captured by λp. Second, by imposing possible penalties on local officials' promotion prospects upon detection, the reform could increase the promotional cost of manipulation, which is captured by δp. Third, the reform may also decrease m^* by removing those involved in manipulation upon detection, leading to a change of the shape of $K(\cdot)$. Regarding local leaders' effort in developing the economy e^*, the reform could increase e^* as it would be relatively less costly to develop the economy.

To empirically test the reform's effect, recall that the reform mainly relied on the survey teams deployed in some counties to detect manipulation. In counties without the survey teams, the detection of manipulation relied on local statistical bureaus. As local statistical bureaus share aligned incentives with local officials and are also controlled by local officials in terms of personnel and funding, they are essentially dysfunctional in terms of detection. As can be seen from the expression of m^*, these counties would essentially be unaffected by the reform without effective detection (i.e., $p = 0$). One could thus test the reform's effect through a difference-in-differences framework by comparing counties with and without the survey teams before and after 2009.

3 Data and empirical strategy

3.1 Main data

Below, I briefly describe the main data used in this paper. More details about these data, along with descriptions of additional data, will be provided when the data first appears in the paper.

The list of counties with survey teams The list of counties with survey teams deployed in 2005 is collected from the annual reports published by various levels of statistical bureaus, and supplemented by local gazetteers.

County-level outcomes and covariates County-level data on GDP and other variables are collected from county statistical yearbooks. County-level data on harmonized nighttime light intensity are collected from Li et al. (2020). These data will be used to construct proxies for GDP growth manipulation. Other county data on demographic, economic, and geographic characteristics, which are used to conduct balance tests and serve as controls, are collected from multiple sources including the 2010 population census, the National Oceanic and Atmospheric Administration (NOAA), and the United States Geological Survey (USGS).

Local leader résumés Local leader résumés for the party secretary and magistrates are collected from various government websites, Baidu Baike (China's equivalent of Wikipedia), and occasionally complemented by online news reports. They are used to construct variables on leader characteristics.

County-level government annual work reports County-level government annual work reports, which outline a county's development policies, are collected from the government websites of each county. This data will be used to examine local officials' policy changes.

Bank credit County-level data on bank loans and branches for various types of banks is collected from the China Banking Regulatory Commission. This data will be used to examine credit allocation.

Firm entry Firm registration data, initially at the firm level for various types of firms (private firms, state-owned enterprises (SOEs), foreign-owned firms, and collectively owned firms), is collected from (Dong et al., 2021). This data will be used to examine firm entry.

Household surveys Household survey data (the China Family Panel Studies) is collected from the Institute of Social Science Survey (ISSS) maintained by Peking University. This survey is a nationally representative survey and will be used to examine citizens' attitudes towards local officials.

My main analysis focuses on a county-level panel from 2005 to 2018. To construct the county sample, I exclude the following special types of counties following the convention in the literature (Li, Lu and Wang, 2016; Chen et al., 2020): (1) counties in the four centrally-managed cities (*Beijing*, *Shanghai*, *Tianjin*, and *Chongqing*). These counties have a higher political status than others and, therefore, are not comparable to other counties; (2) urban districts, which are more developed economically but less independent administratively;[7] (3) counties in Tibet where data is unavailable; (4) counties outside mainland China.[8] I then define treatment counties as those with the surveys teams deployed in 2005 and control counties as those without. In the end, I have 1,779 counties in total, of which 40% are treated. The spatial distribution of treatment and control counties can be found in Figure 1, which is quite even across space.

3.2 Empirical strategy

Deriving estimation equation To derive the estimation equation, I incorporate GDP growth manipulation and the reform into the framework of Henderson, Storeygard and Weil (2012), who establish the positive relationship between nighttime light intensity and real economic activities. First, denote the reported GDP growth (with manipulation), true GDP growth (unobservable), and nighttime light intensity growth in county c and year t as z_{ct}, y_{ct}, and l_{ct}, respectively. Assume the degree of manipulation is m_{ct}. Then the GDP growth observed by the local statistical bureau is $z_{ct} - m_{ct}$ (without manipulation). According to Henderson, Storey-

[7] Both urban districts and counties are county-level divisions under a prefecture-level city. Some key differences are: (1) urban districts are the core areas of a prefecture-level city and are thus more developed; (2) urban districts have strong dependency on the city in terms of administrative functions (such as land development, urban planning, fiscal expenditure, etc.).

[8] Specifically, these include counties in Hongkong, Macau, and Taiwan. They are excluded due to institutional and administrative differences from mainland China.

gard and Weil (2012), the mapping from $z_{ct} - m_{ct}$ to y_{ct}, and the mapping from l_{ct} to y_{ct} can be written respectively as:

$$z_{ct} - m_{ct} = y_{ct} + \varepsilon^z_{ct} \tag{1}$$

$$l_{ct} = \gamma y_{ct} + \varepsilon^l_{ct} \tag{2}$$

Combining equations (1) and (2), the degree of manipulation can be written as:

$$m_{ct} = z_{ct} - \frac{1}{\gamma} l_{ct} + \varepsilon^m_{ct} \tag{3}$$

where ε^m_{ct} is a combination of the error terms ε^z_{ct} and ε^l_{ct}. Then the difference-in-differences equation to test the effects of the reform on GDP growth manipulation can be written as:

$$m_{ct} = \beta Treat_c \times Post_t + \delta_c + \lambda_t + \varepsilon^m_{ct} \tag{4}$$

where $Treat_c$ and $Post_t$ are dummy variables for treatment counties (the 40% aforementioned counties with survey teams deployed in 2005) and post-reform years (years after 2009), respectively. δ_c denotes county fixed effects, controlling for time-invariant factors at the county level that may correlate with the treatment or the outcome; λ_t denotes year fixed effects, controlling for time-varying shocks common to all counties. As one cannot directly observe m_{ct}, substituting equation (3) into equation (4) and rearranging generates:

$$\underbrace{z_{ct}}_{\text{Reported GDP Growth}} = \frac{1}{\gamma} \underbrace{l_{ct}}_{\text{Light Growth}} + \beta Treat_c \times Post_t + \delta_c + \lambda_t + \varepsilon^z_{ct} \tag{5}$$

where ε^z_{ct} is a combination of ε^m_{ct} and ε^m_{ct}. One can then estimate this equation. Since the treatment varies at the county level, I cluster the standard errors by county (Abadie et al., 2023) and assess robustness using alternative inference procedures, such as clustering at different levels (city and province), correcting for spatial correlation (Conley, 1999), and employing randomization inference (Young, 2019). I expect the coefficient of interest β to be negative, which implies that counties with the survey teams would engage in less manipulation relative to other counties after the reform. Note that both local statistical bureaus and the survey teams could engage in detecting manipulation after the reform, which may affect the interpretation of β. Under the assumption that local statistical bureaus are dysfunctional in detection, which is plausible as they are controlled by local officials, the coefficient β could be well interpreted as the overall effect of the reform. If this assumption is not true, the coefficient β is a lower bound of the effect of the reform, but this is still meaningful. In addition, β may also capture the lower bound effect in the presence of spillover effects among counties, which I will rule out empirically.

Identification concerns The identification assumption is that, reported GDP growth, after adjusting for light growth, should evolve in parallel between treatment and control counties in the absence of the reform in 2009. This assumption is essentially unverifiable. Pre-reform parallel trends between treatment and control counties, which is commonly estimated using event study specifications, can lend support to this assumption but cannot fully verify it. One still needs to address two types of concerns: first, the relationship between economic growth and light growth may differ across counties or years, which is specific to my setting; second, treatment counties may differ significantly from control counties ex ante, a common concern in difference-in-differences designs. To address the first concern, in the spirit of Martinez (2022), I will allow the effect of light to vary by a host of spatial and temporal characteristics to check the sensitivity of the estimates. For the second concern, while perfect covariate balance ex ante is not necessarily required in such designs, significant imbalance may cast doubt on the validity of using the control groups as counterfactual. To check this, Table 1 provides a balance test a-long various baseline county covariates, which shows that treatment counties were quite similar to control counties ex ante, except for the levels of population and GDP. The pre-deployment nature of the survey teams implies that such imbalance should not be endogenously related to the reform and therefore should be largely orthogonal to my outcomes. Indeed, there was no significant difference for pre-reform GDP growth and light growth, implying that treatment and control counties were similar in terms of GDP growth manipulation (or lack thereof). In the robustness checks, I also adopt various methods to address concerns about covariate imbalance, such as flexibly controlling for size and other baseline covariates, allowing for county-specific trends, and achieving covariate balance through entropy balancing (Hainmueller, 2012) and coarsened exact matching (Iacus, King and Porro, 2012). I further leverage institutional knowledge to design two additional tests to address concerns about unobserved heterogeneity: the first is a placebo event study around 2005 when the survey teams were launched but had not yet undertaken any disciplining actions; the second is an instrumental variable estimation using the randomly assigned rural survey teams in the 1980s as an instrument, which I will elaborate on later.

4 Results on GDP growth manipulation

4.1 Main results

Event study Figure 2 shows the dynamic effect of the reform estimated using an event study specification (6), with baseline county covariates included gradually from panel (a) to panel (d). The year before the reform, 2008, is omitted as the reference year. The coefficient estimates in the pre-reform period, namely, β_js for $j < 2009$, are essentially small in magnitude and statistically insignificant. F-tests of joint significance of all the pre-reform estimates generate p-values larger than 0.9 in all specifications, implying that the parallel trends assumption is plausibly satisfied. Note that recent econometric literature shows that this type of pretests may be underpowered to detect a diverging pre-trend (Roth, 2022). However, as I show in subsequent analysis, the results are robust to accounting for potential pre-trends using a couple of methods including an instrumental variable approach. In the post-reform period, there is an immediate and persistent negative effect, suggesting that the reform decreased GDP growth manipulation, which confirms the prediction from the conceptual framework. In Appendix Figure A1, I further show a decomposition of the effect of the reform, by checking the dynamic effect on reported GDP growth and light growth separately.[9] The results further confirm that the reform decreased GDP growth manipulation: there is a sharp drop in reported GDP growth but little change in

[9] Specifically, I estimate the following equation, where Y_{ct} denotes either reported GDP growth or light growth:

$$Y_{ct} = \sum_{j=2005,\ j\neq 2008}^{j=2018} \beta_j Treat_c \times 1_{\{t=j\}} + \delta_c + \lambda_t + \varepsilon_{ct}$$

light growth in the post-reform period.

$$ReportedGDPGrowth_{ct} = \alpha LightGrowth_{ct} + \sum_{j=2005,\ j\neq 2008}^{j=2018} \beta_j Treat_c \times 1_{\{t=j\}} + \delta_c + \lambda_t + \varepsilon_{ct} \tag{6}$$

Average effect Table 2 summarizes the dynamic treatment effect above into an average treatment effect. Column (1) reports the results using equation (5), controlling for only county and year fixed effects, and light growth. The point estimate on Treat $\times$ Post is negative and statistically significant (coef.=-0.751, s.e.=0.316), implying that relative to the control counties, treatment counties experienced a 0.751 percentage points drop in GDP growth manipulation after the reform. The estimate on light growth is also consistent with that in the literature (Martinez, 2022).[10] Through columns (2)-(4), I gradually introduce a set of baseline demographic, economic, and geographic controls (interacted with the post-reform dummy), which are presented in the balance tests. The precision of the estimates improves, although the size drops slightly. In the most stringent specification in column (4) with all the county controls, the estimate shows a 0.576 (s.e.=0.161) percentage points drop in GDP growth manipulation in treatment counties relative to control counties after the reform. This drop is also economically substantial, which amounts to 5.3% of the mean of the reported GDP growth. In sum, these findings suggest the effectiveness of the combination of monitoring and punishment in reducing bureaucratic misconduct in data processing, which resonates with the key insights from Becker and Stigler (1974). In Appendix Table A1, I also show the estimates on reported GDP growth and light growth separately, and find a similar decrease in reported GDP growth (coef.=-0.576, s.e.=0.161). The effect on light growth is small and statistically insignificant (coef.=0.254, s.e.=0.294). This means that the reform reduced GDP growth manipulation primarily through a reduction in reported GDP growth.[11]

Spillover Having established the negative effect of the reform on GDP growth manipulation,

[10] Specifically, in a similar specification in Martinez (2022)'s cross-country analysis, the coefficient estimate on light growth is about 0.027-0.039 (s.e.=0.006-0.007), and in my setting, it is about 0.017-0.023 (s.e.=0.005).

[11] If any, the positive yet insignificant effect on light growth implies my estimates may be a lower bound of the true impact on GDP growth manipulation.

I turn to check if there exists any spillover effect, which could bias my baseline estimation even if the treatment is exogenous.[12] To this end, I estimate equation (7), where $Spillover_c$ denotes the strength of spillover to county c from other counties. Thus, β^{Direct} captures the direct effect of the reform while $\beta^{Spillover}$ captures the spillover effect. Following Avis, Ferraz and Finan (2018) and Huber (2023), I use the number of treatment counties among a county's neighbors to proxy for the strength of spillover to that county, where neighbors are defined as other counties sharing a common boundary segment with that county.[13] The results are reported in Table 3. Column (1) reproduces the baseline estiamtes. In column (2), the estimated spillover effect is small and statistically insignificant (coef.=-0.01, s.e.=0.071). Considering the average number of treatment neighbors for a county is 2, such estimates imply that moving from a county with no treatment neighbors to the average county would decrease GDP growth manipulation by an additional 0.02 percentage points. Given the direct effect of about 0.58 percentage points, the spillover effect is thus economically negligible. In contrast, the direct effect remains virtually unchanged compared to the baseline effect in column (1). In the remaining two columns, I use dummies to indicate counties with at least one treatment neighbor or with treatment neighbors higher than the median, and the results have no substantial changes. In Appendix Table A2, I further show that the results are robust to using alternative definitions of neighbors or weighting the treatment neighbors by their sizes. The lack of a spillover effect can be well reconciled with two facts: (1) the coverage of treatment counties is fixed over time, as they hinges on the pre-deployed survey teams; (2) the evaluation of local officials may be among similar counties (Xu, 2011), namely, within either treatment or control counties.

$$\begin{aligned} ReportedGDPGrowth_{ct} = \alpha LightGrowth_{ct} + {} & \beta^{Direct} Treat_c \times Post_t \\ & + \beta^{Spillover} Spillover_c \times Post_t + \delta_c + \lambda_t + \varepsilon_{ct} \end{aligned} \tag{7}$$

[12]To the extent that there is a spillover effect to control counties, my estimates would underestimate the true reform effect.

[13]To alleviate concerns that this number captures the geographic centrality of a county, I include neighbor number fixed effects interacted with the post-reform dummy throughout the estimation of spillover effect. Results are essentially unaffected by such fixed effects.

4.2 Robustness checks

4.2.1 Alternative specifications

In this section, I show the robustness of the baseline results to a host of alternative specifications that alleviate the aforementioned identification concerns. The first concern is about heterogeneous mapping between nighttime light intensity and true economic activities. As previously discussed, the baseline equation (5) assumes a uniform and linear relationship between these two variables. However, such relationship may be nonlinear and change across counties or over years. To alleviate this concern, I allow the effect of light to: (1) be non-linear by including a 3rd-order polynomial of light; (2) vary by county longitude and latitude; (3) vary by county area; (4) vary by baseline GDP (5) vary by baseline population (6) vary by baseline urbanization rate; (7) vary by baseline economic structure (proxied by share of population in the primary and secondary sectors); (8) vary by year; (9) vary by province; (10) vary by both province and year; (11) vary by treatment status; (12) vary by both treatment status and linearly by year. To better examine the sensitivity to a specific modification, I estimate a variant of the baseline equation (5) each time according to one of the modifications above, and plot the results in panel (a) of Figure 3. The results are essentially unaffected by these alternative specifications, which suggests that my baseline findings are not an artifact of heterogeneous light effect. In Panel (a) of Appendix Figure A2, I futher control for baseline county covariates and also estimate a specification incorporating all the modifications above. The patterns are similar.

The second concern is about covariate imbalance. As shown in Table 1, treatment counties were larger in size than control counties ex ante and therefore may differ significantly from control counties later on, leading to potential violations of the parallel trends assumption. To alleviate this concern, I estimate the following alternative specifications: (1) I flexibly control for size effects by including county size decile bin fixed effects interacted with year fixed effects, where county size is proxied by baseline GDP, population, or area; (2) I add county-specific time trends that allow treatment and control counties to be on differential linear trajectories (Angrist and Pischke, 2014). This could relax the identification assumption, although the precision of the estimates may decrease;[14] (3) I add province $\times$ year fixed effects. In this way, I

[14] Using linear time trends in DiD specifications could absorb part of the effect and the treatment variation, which

am only comparing counties in the same province and year, and the covariates should be more balanced; (4) I include all the baseline county covariates interacted with year fixed effects to allow treatment and control counties to trend differentially depending on the covariates; (5) I select the most relevant covariates using the Double LASSO method (Belloni, Chernozhukov and Hansen, 2014); (6) I re-weight observations to make treatment and control counties similar in terms of observables using the entropy balancing method (Hainmueller, 2012); (7) I adopt the coarsened exact matching (CEM) method to match treatment counties to control counties within groups defined by all intersections of the deciles of baseline GDP, population, and area (Iacus, King and Porro, 2012). To better examine the sensitivity to a specific modification, I estimate a variant of the baseline equation (5) each time according to one of the modifications above, and plot the results in panel (b) of Figure 3.[15] These estimates have no substantial changes compared to the baseline estimate, implying that my findings are not driven by possible differential trends caused by covariate imbalance. I further control for baseline county covariates when appropriate in panel (b) of Appendix Figure A2 and find similar patterns.

I also address some other concerns in the remaining panels of Appendix Figure A2. Panel (c) shows that the results are not driven by a particular region, by conducting estimations leaving out each province individually. Panel (d) shows that the results are not driven by a few marginal counties by weighting the regression by county size (e.g., population or GDP). Panel (e) shows that the results are robust to alternative levels of clustering (by city or province) and spatial correlation correction (Conley, 1999).[16] Panel (f) shows that the results are robust to randomization inference, which has better finite sample properties and is also insensitive to high-leverage observations (Young, 2019).[17]

leads to less precise estimates (Goodman-Bacon, 2021).

[15]Note that unlike in panel (a), here a specification incorporating all the modifications is unfeasible.

[16]For province-level clustering with a small number of clusters of 26, I also report the wild bootstrap p-value with 2,000 replications (Roodman et al., 2019). For Conley standard errors, I account for serial correlation spanning all years and spatial correlation within distances of 250 km, 500 km, 750 km, and 1,000 km.

[17]Following the recommendation by Young (2019), I use 2,000 permutations as the marginal gain from additional permutations is minimal. The randomization inference p-value from this exercise is 0.001.

4.2.2 Placebo reform: The launch of the survey teams in 2005

I corroborate the previous results by conducting a placebo event study around 2005 when the survey teams were launched but had not conducted any disciplining actions. This could further help to examine if there existed any pre-trends. In addition, it could alleviate further concerns that the survey teams per se may affect the outcomes, even without the reform in 2009. The specification is equation (8), which is similar to the baseline event study specification in equation (6) except that the sample period here is from 2001 to 2008, with the year 2004 omitted as the reference year. Figure 4 shows the event study estimates. In contrast to the sharp drop in reported GDP growth after 2009 in Figure 2, there was no discernible change in reported GDP growth around 2005. This pattern persisted until 2009. Furthermore, there were no diverging trends before 2005. Such results suggest that the reform effect is unlikely to be driven by baseline differences between treatment and control counties or differential effects (net of the reform effect) generated by the survey teams per se.

$$
\begin{aligned}
ReportedGDPGrowth_{ct} = \alpha LightGrowth_{ct} + \sum_{j=2001,\ j\neq 2004}^{j=2008} \beta_j Treat_c \times 1_{\{t=j\}} \\
+ \delta_c + \lambda_t + \varepsilon_{ct} \qquad (8)
\end{aligned}
$$

4.2.3 Instrumental variable strategy

While the results presented so far could alleviate most concerns about covariate unbalance, unobserved heterogeneity is still possible. For example, to ensure the quality of information collected, counties with the survey teams may be those good counties with a potentially downward and time-varying trend in GDP growth manipulation. In this case, the OLS estimates would overestimate the true impact of the reform. This may be legitimate concern given the positive selection nature of most policy experimentations in China (Wang and Yang, 2021). To alleviate this concern, in this section I leverage a unique institutional feature to construct an instrumental variable for the treatment counties and conduct an instrumented difference-in-differences estimation.

Background In 1984, to gauge agricultural production, the National Bureau of Statistics

(NBS) set up a group of teams called *rural survey teams* in part of the counties. At that time, China was essentially an agricultural country. The counties with these rural survey teams were chosen randomly within a province. In particular, the NBS adopted a commonly used probability sampling method called *systematic random sampling*. Under this sampling method, one first selects a random starting point in a sequence of counties and then chooses counties at fixed and periodic intervals. I collect the list of counties eventually selected from provincial gazetteers.[18] In theory, counties within the same province should have the same probability of being selected, leading to perfect within-province randomness of assignment of counties with the rural survey teams. In practice, the randomness may be affected by particular patterns in the county sequence or the limited number of counties in some provinces,[19] and I will provide several tests to check the randomness. In terms of specific work, these rural survey teams were guided by the NBS, but in terms of personnel and funding, they were controlled by local officials. Given the dramatic change in economic structures caused by market reform in recent years, in 2005, these rural survey teams were abolished, and most of them were restructured into more comprehensive and independent survey teams led solely by the NBS, which are the survey teams examined in the previous parts. In sum, the unique random assignment feature of these earlier rural survey teams suggests that they could be used as an instrument for treatment counties, with the validity formally examined below.

Relevance I define an instrumental variable $Treat_c^{1984}$, which is a dummy equal to 1 if county c had a rural survey team in 1984. Given the previous discussion, I expect this instrument to be strongly correlated with $Treat_c$. Panel (a) of Figure 5 shows the distribution of the rural survey teams launched in 1984. The significant overlap with the treatment counties suggests the high relevance of the instrument. To formally assess the strength of this instrument, I report

[18]Gazetteers are called *Difangzhi* in Chinese and are a series of encyclopedias covering a wide range of topics: history, geography, economics, politics, culture, social sciences, etc. They are compiled by local officials and noted literati in each county and updated every dozens of years.

[19]In my sample, the number of counties in each province ranges from 13 to 128, with an average of 80.

the first-stage regression results in panel (a) of Table 4 according to the following equation:

$$
\begin{aligned}
Treat_c \times Post_t = \theta LightGrowth_{ct} + \ \tau Treat_c^{1984} \times Post_t \\
+ \gamma_p \times Post_t + \delta_c + \lambda_t + \varepsilon_{ct}
\end{aligned}
\tag{9}
$$

I include province fixed effects γ_p interacted with the post-reform dummy $Post_t$ to account for the fact that the random assignment is stratified by province (Duflo, Glennerster and Kremer, 2007; Bruhn and McKenzie, 2009). The coefficient estimates on $Treat_c^{1984} \times Post_t$ are positive and highly statistically significant across specifications, indicating the instrument's strong relevance. Following the suggestion by Andrews, Stock and Sun (2019), I report the effective F-statistic to assess the strength of the instrument (Montiel Olea and Pflueger, 2013).[20] Across all specifications, the effective F-statistics are around 2,000, which far exceeds both the rule-of-thumb value of 10 and the 5% critical value of 37.4.(Montiel Olea and Pflueger, 2013). In addition, Lee et al. (2022) argue that inference relying on the first-stage F-statistic exceeding a certain threshold may still be distorted, unless the F-statistic is larger than 104.7.[21] In this sense, the large F-statistics in my case are sufficient to guarantee correct inference in the second stage. In addition to the high relevance, it is worth noting that the coefficients remain highly stable when various county characteristics are included. This stability points to the random assignment nature of the instrument, which I will further test below.

Exogeneity For the instrument $Treat_c^{1984}$ to be valid, it must also be uncorrelated with any other determinants of my outcomes, except the treatment, namely, the launch of a survey team in 2005. While this condition is essentially untestable, the random assignment nature of the instrument suggests that this condition should be plausibly satisfied. To check this, I run separate univariate regressions of various baseline demographic, economic, and geographic county characteristics on the instrument with province fixed effects included. Panel (b) of Figure 5 plots the standardized coefficients from this exercise. These coefficients are not only statistically insignificant but also centered around zero with small magnitudes, strongly supporting the

[20] In just-identified cases, the effective F-statistic is equivalent to the conventional Kleibergen-Paap F-statistic.

[21] With F-statistics smaller than 104.7, Lee et al. (2022) shows that an adjustment factor should be applied to inflate the second-stage standard errors to deliver correct inference, which is not needed in my case.

randomness of the instrument.

To provide further evidence, I compute the standardized differences between counties with and without the rural survey teams.[22] Since the randomization is stratified by province, I first calculate the standardized differences within each province and then calculate a weighted average using the number of counties within each province as weights. Appendix Figure A3 plots the distribution of the standardized differences across county characteristics. The absolute values of the standardized differences never exceed 8%, which is far below the threshold of 25% for covariate balance as suggested by Imbens and Rubin (2015) .

Despite the strong evidence on the randomness of the instrument, one remaining concern is about potential *legacy effect* generated by the rural survey teams. Specifically, these rural survey teams, albeit abolished in 2005, may still have a lasting impact on the outcomes examined in my study period. A possible cause is that local officials may develop a cautious character due to the existence of the rural survey teams. If this character is passed on to subsequent leaders and leads to less manipulation, then the exclusion restriction could be violated. Conceptually, I view this as very unlikely due to two facts. First, the rural survey teams did not conduct any disciplining actions and were essentially led by local officials. Second, it is very rare that two consecutive leaders could co-work in the same county for a long period. In most turnover years, the overlap of their reigns is no longer than a few months, meaning that the legacy effect, if any, is unlikely to be persistent.

Another possible cause of the legacy effect is that upper-level governments may acquire better information about the agricultural sector in the counties with the rural survey teams. If so, it would be difficult for these counties to manipulate data on agricultural production even after the abolition of the rural survey teams. However, this should lead to differential shares of the agricultural sectors, which is not the case as one could see from the balance test. This is likely due to the frequent turnovers of upper-level leaders and significant changes in economic structures, rendering information acquired decades ago less useful.

While there may still exist other sources of the legacy effect, it is worth noting that the legacy effect could be *differenced out* by my DID strategy around 2009, as long as it did not

[22] The standardized difference between two groups within a province is calculated as the difference between the sample means normalized by the square root of the average of the sample variances.

change over time, which seems plausible. In other words, only time-varying legacy effect that is correlated with my outcomes would lead to a violation of the exclusion restriction. One way to rule out a time-varying legacy effect is to check whether there exist any pre-trends before 2009 using the reduced-form event study specification as below:

$$ReportedGDPGrowth_{ct} = \alpha LightGrowth_{ct} + \sum_{j=2005,\ j\neq 2008}^{j=2018} \beta_j Treat_c^{1984} \times 1_{\{t=j\}}$$
$$+ \gamma_p \times Post_t + \delta_c + \lambda_t + \varepsilon_{ct} \tag{10}$$

where the only difference of this specification with the baseline event study specification is that the $Treat_c$ there is replaced by the instrument $Treat_c^{1984}$. β_js are the coefficients of interest and are expected to be different from zero in the pre-reform period if there exists a time-varying legacy effect. Panel (c) of Figure 5 plots the coefficient estimates from this exercise, which are small and statistically insignificant in the pre-reform period. This rules out the possibility of a time-varying legacy effect directly affecting manipulation in my study period.

2SLS estimates Given the relevance and exogeneity of the instrument, I perform 2SLS estimation with the following second-stage equation:

$$ReportedGDPGrowth_{ct} = \vartheta LightGrowth_{ct} + \beta \widehat{Treat_c \times Post_t}$$
$$+ \gamma_p \times Post_t + \delta_c + \lambda_t + \varepsilon_{ct} \tag{11}$$

where all variables are defined as previously, and the coefficient of interest is β, which captures the average causal effect of the reform. To check the dynamics, I first conduct an instrumented event study and plot the coefficient estimates in panel (d) of Figure 5. Similar to the baseline OLS event study (Figure 2), the figure shows no pre-trends in the pre-reform period and a persistent drop in GDP growth manipulation in the post-reform period, albeit with less precision. To summarize the dynamic effect, I report the second-stage results estimated based on equation (11) in panel B of Table 4. The results are also similar to OLS estimates both in terms of economic magnitude and statistical significance. Considering the estimate in the last column with the inclusion of baseline county covariates interacted with the post-reform dummy, the estimate shows a 0.59 percentage points drop in GDP growth manipulation in treatment counties relative to control counties after the reform in 2009 (coef.=-0.590, s.e.=0.203), which accounts for 5.4% of the mean of reported GDP growth and is nearly identical to the baseline OLS magnitude of 5.3%. Such similar results imply that the baseline findings are unlikely to be confounded by selection bias. This is consistent with the aforementioned fact that treatment counties were determined several years before the reform; therefore, any preexisting differences between treatment and control counties should be largely orthogonal to GDP growth manipulation. In Appendix Table A1, I further show that the reduction in GDP growth manipulation is mainly driven by a reduction in reported GDP growth, by providing IV estimates on reported GDP growth and light growth separately.[23]

Difference between IV and OLS While the similar magnitudes found in the IV and OLS estimations imply that selection bias is a minimal concern, it could also be that the IV estimates only reflect the treatment effect on compliers, namely, counties affected by the instrument. In the presence of heterogeneous treatment effects across counties, the treatment effect on compliers may differ from that on the full sample.[24] I take several steps to alleviate this concern. First, I calcuate the proportion of compliers and profile their characteristics using the method developed by Marbach and Hangartner (2020). The detailed procedures are described in Appendix B. I find that about three quarters of the counties are compliers. These counties also look similar to the full sample on all observable dimensions as shown in Appendix Figure A4. Second, I adopt a reweighting method in Appendix Table A4 to adjust the OLS estimates to match the sample of compliers (Bhuller et al., 2020; Agan, Doleac and Harvey, 2023).[25] The reweighted OLS estimates are quite similar to the unweighted OLS estimates. These results suggest that the similarity between the IV and OLS estimates is unlikely to be confounded by heterogeneous

[23]Specifically, the table shows a similar decrease in reported GDP growth (coef.=-0.592, s.e.=0.203). The effect on light growth, while positive, is small and statistically insignificant (coef.=0.162, s.e.=0.355).

[24]Under heterogeneity treatment effects, the monotonic condition is also required, which means that counties affected by the instrument should be affected in the same way. A testable prediction of this condition is that the first-stage results should hold qualitatively across all subgroups. In Appendix Table A3, I provide suggestive evidence that this is satisfied by showing that the first-stage results are similar across subgroups divided by the medians of baseline county covariates.

[25]Specifically, I split the sample into multiple groups of equal size based on the baseline county covariates, and then reweight the OLS estimates using the complier share in each group as weights. See Appendix Table A4 for details.

treatment effects.

4.3 Mechanisms and alternative explanations

After demonstrating the robustness of the baseline findings, in this section I explore the underlying mechanisms. I show that the drop in GDP growth manipulation is most consistent with a reputational discipline effect generated by the reform on local officials. Other mechanisms outlined in the conceptual framework, such as promotional discipline effect and personnel changes, play a minimal role. Additionally, I rule out several alternative explanations that could generate similar patterns, namely, enhanced soft information, improved local statistical capacity, and other concurrent reforms that may also strengthen the monitoring of local officials.

4.3.1 Reputational discipline effect as the key mechanism

As discussed in the conceptual framework in Section 2.2, the reform could decrease GDP growth manipulation through two broad mechanisms: a *discipline* effect and a *selection* effect. The discipline effect suggests that local officials would refrain from manipulation within their terms due to reputational and promotional costs. The selection effect implies a reduction in manipulation as a result of personnel changes among local officials, stemming from the dismissal of those found to be involved in manipulation or the appointment of more competent successors.

To distinguish between the discipline effect and the selection effect, I construct a trimmed sample in which I require local officials' terms to straddle 2009 in the treatment counties.[26] Given the lack of personnel changes in the treatment counties in this sample, one should observe a smaller reduction in GDP growth manipulation in this sample if the reform partly worked through a selection effect. As presented in column (2) of Appendix Table A5, the reduction manipulation (coef.=-0.617, s.e.=0.207) in this trimmed sample is similar to the reduction (coef.=-0.576, s.e.=0.161) in the full sample as shown in column (1). If anything, the reform effect is slightly larger in this trimmed sample. This means that the selection effect cannot explain the drop in manipulation. In column (3), I further require local officials to stay at least two years both before and after the reform in the treatment counties, and the estimate is similar.

[26] As a county is co-led by the party secretary and the magistrate, I require the terms of both types of officials to straddle 2009 when constructing the trimmed sample.

In Appendix Table A6, I also test the effects of the reform on personnel turnovers and personnel traits, which are generally small and statistically insignificant. In sum, these results imply a minimal role played by the selection effect. This null effect is consistent with the fact that only the most severe cases of manipulation would result in the dismissal of local officials, which is itself a rare occurrence.

Given the absence of a selection effect, the reform likely worked through a discipline effect, which could be further decomposed into reputational and promotional discipline effects. For the latter effect, one should expect the reduction in GDP growth manipulation to vary by local officials' promotion incentives. In contrast, the former effect should not lead to such heterogeneity. To distinguish between these two mechanisms, I start with testing whether the reform effect is larger for local officials with greater promotion incentives. I measure local officials' promotion incentives utilizing age restrictions on promotion. Specifically, county-level leaders are generally unable to be promoted to the next level once their ages reach 52 (Kou and Tsai, 2014), which creates a sharp drop in promotion incentives at this age cutoff. In addition, as a county is co-led by both the party secretary and the magistrate, I adopt a specification allowing for their incentives to separately affect GDP growth manipulation (Yao and Zhang, 2015).[27] As shown in Appendix Table A8, I do not find a statistically significant differential effect for local officials older than 52 who would have lower promotion incentives. As another measure of promotion incentives, I also estimate local officials' ex ante likelihood of promotion based on their start ages, years of schooling, and political connections with upper-level leaders (Avis, Ferraz and Finan, 2018; Wang, Zhang and Zhou, 2020).[28] I still do not find a statistically significant differential effect for local officials with higher promotion incentives as shown in Appendix Table

[27] Specifically, following Yao and Zhang (2015), I treat the party secretary and the magistrate in a county as if they work in two different but identical counties. Empirically, this means that for each county-year, I generate two parallel observations that are identical except for one distinction: one includes only the party secretary, while the other includes only the magistrate.

[28] Specifically, I estimate the following Probit model based on local officials' ages when they started their terms, years of schooling, political connections with upper-level leaders, and all two-way interactions between these three variables. I estimate it separately for the party secretary and the magistrate.

$$\Phi^{-1}[P(Promotion)] = \beta_0 + \beta_1 StartAge_i + \beta_2 Education_i + \beta_3 Connection_i + \beta_4 StartAge_i \times StartAge_i + \beta_5 Education_i \times Connection_i + \beta_6 StartAge_i \times Connection_i + \varepsilon_i$$

Note that other performance variables, such as GDP growth, are intentionally excluded from this regression, so the estimated probabilities capture the ex ante likelihood of promotion (Wang, Zhang and Zhou, 2020). The estimates are reported in Appendix Table A7. The estimates are similar if I instead use a linear probability model.

A8. These results mean that promotional punishment cannot explain the drop in GDP growth manipulation. This could be possible if in reality the survey teams were only able to detect manipulation but cannot precisely determine the total amount of manipulation. Consequently, imposing a promotion-based punishment becomes impractical.

Taken together, the evidence so far is most consistent with the reform generating a reputational discipline effect on local officials. This echoes the insights from a growing literature highlighting the role of reputational concerns in shaping truthtelling (Benabou and Tirole, 2011; Gneezy, Kajackaite and Sobel, 2018; Abeler, Nosenzo and Raymond, 2019). I further rule out some other explanations in the following sections.

4.3.2 Alternative explanations

Soft information The survey teams may assist the upper-level government in achieving soft information about the performance of local officials, thereby dampening the role of GDP growth in promotion (Hart, 1995; Aghion and Tirole, 1997; Stein, 2002). Consequently, local officials may be less inclined to manipulate GDP growth after the reform. To explore this possibility, I focus on counties closer to the upper-level government or counties where the leaders are socially connected to the upper-level government. As well documented in the literature, shorter distances or social connections could also facilitate the flow of soft information.[29] As a result, one should expect to see a smaller reform effect as the soft information provided by the survey teams should be less instrumental in such counties. I follow the literature to measure social connections using shared hometown or education background between county leaders and leaders in the upper-level government (Shih, Adolph and Liu, 2012; Jia, Kudamatsu and Seim, 2015; Fisman et al., 2020).[30] As shown in Appendix Table A9, I do not find a statistically significant differential effect for these counties, suggesting that soft information is unlikely to be a driving force of the baseline results.

Statistical capacity Local statistical bureaus may improve their statistical capacity through

[29] See for example Agarwal and Hauswald (2010), Petersen and Rajan (2002), Bandiera, Barankay and Rasul (2009), and Fisman, Paravisini and Vig (2017).

[30] Specifically, I create a dummy variable named Connection that equals 1 if the party secretary or magistrate in a county shares the same hometown or educational background with upper-level leaders, and 0 otherwise.

interactions with the survey teams, and hence could more accurately measure economic activity (Martinez, 2022). This may also lead to a drop in the gap between reported GDP growth and light growth. Given the difficulty in directly measuring the statistical capacity of a county, I utilize an award from the National Bureau of Statistics (NBS) for counties' outstanding performance in coordinating and conducting economic censuses, which are initiated every 4 or 5 years by the NBS. The award involves no material rewards. Data on recipient counties of such award is collected from the NBS and is available for the years 2004, 2008, 2013, and 2018. In Appendix Table A10, I show that treatment counties did not receive more such award after the reform, suggesting that the baseline results are unlikely driven by improvement in local statistical capacity. This null effect is further supported by an event study presented in Appendix Figure A5.

Concurrent reforms The baseline results could be confounded by concurrent reforms that may also strengthen the monitoring of local officials. The inclusion of province $\times$ year fixed effects in the robustness checks could rule out all confounding reforms at the province level. In this section, I examine two prominent reforms at the county level. The first reform is the province-managing-county (PMC) reform since 2003 (Li, Lu and Wang, 2016). This PMC reform stipulated that the provincial government could bypass the prefectural government and directly administer the county government in fiscal matters (Fiscal PMC), or even in all aspects for a few counties (Full PMC). This data is collected from various government websites. The second reform pertains to the top-down inspections during the recent anti-corruption campaign launched in 2013. These inspections mainly focused on curbing corruption but may also create discipline effect on all aspects. This data is complied by Wang (2021), who collects the detailed timings and sites of inspections from government websites and newspapers. Appendix Figure A6 presents the rollout of these reforms. Appendix Table A11 shows that the results have no substantial changes after accounting for these reforms.

5 Results on effort reallocation

Up to now, I have shown that the reform could address the information problem faced by upper-level governments, namely, GDP growth manipulation by local governments. As outlined in the conceptual framework, an increase in the cost of manipulation should create a shift of local officials' effort from manipulation to economic development after the reform. In this section, I test this conjecture by empirically examining various outcomes reflecting local officials' development effort, which can be grouped into four domains: government policies, bank credit, firm entry, and citizen attitudes towards local governments.[31] Each of these four domains captures a certain aspect of local officials' effort and complements the others. To minimize the risk of finding false positives across multiple outcomes, I focus on relatively aggregate outcomes in each domain, and also correct for multiple hypothesis testing for the aggregate outcomes following Anderson (2008).[32] As I demonstrate below, the results across various outcomes provide consistent and complementary evidence on a beneficial shift in local officials' effort following the reform.

5.1 Government policies

My key measure of local officials' development effort is their emphasis on policies beneficial to the economy, considering the significant latitude they have in shaping local development (Xu, 2011). To construct the measure, I conduct a textual analysis of government annual work reports. These reports are typically issued by local governments at the beginning of each year and are subject to approval by the People's Congress at the same level through anonymous voting. Each report contains two parts: a summary of the government's achievements in the last year and a work plan for the year ahead, which contains detailed and well-structured development policies. The emphasis on each policy area could vary significantly both cross-sectionally

[31]One concern with these outcomes is that they could also be manipulated. This concern should be minimal for two reasons. First, local officials are not evaluated by these outcomes, so their incentives to manipulate them are minimal. Second, outcomes in the latter domains (bank credit, firm entry, and citizen attitudes) are largely immune to interference by local governments as they are collected and maintained by either the central authorities or independent entities.

[32]Due to the different time spans and identification strategies used, further aggregation of the outcomes in different domains is not conducted. Instead, I report the sharpened q-values for the aggregate outcomes in Appendix Table A22 to adjust for multiple hypothesis testing (Anderson, 2008). The findings are essentially unaffected by this correction.

and temporarily, as the reports are essentially at the discretion of local officials. Hence, the reports are well-suited to examine local government policy changes (Jiang, Meng and Zhang, 2019; Campante, Chor and Li, 2022).[33]

I create an original dataset on county-level government annual work reports collected from the official websites of each county. To ease both collection and computational burdens,[34] I randomly select three provinces: *Guangdong*, *Shaanxi*, and *Zhejiang*.[35] The final sample includes 97 counties from 2005 to 2018, corresponding to 1,155 reports and over 320,000 sentences in total. I then define four policies beneficial to the economy: *business attraction*, *infrastructure investment*, *market reform*, and *policy experimentation*, which are major contributory factors to China's recent economic success and are also frequently mentioned in the reports (Li and Zhou, 2005; Xu, 2011; Jiang, Meng and Zhang, 2019).[36] In addition, to alleviate concerns about multiple hypothesis testing, I create a standardized index following Kling, Liebman and Katz (2007), which is constructed as follows. First, I standardize each of the four policies to have a mean of 0 and standard deviation of 1. Second, I take an equally weighted average of the four standardized measures. Third, I standardize the weighted average again to have a mean of 0 and standardized deviation of 1, which serves as my main outcome of interest in this section.

I adopt two approaches to measure the emphasis on each policy, as detailed in Appendix C. The first is a keywords frequency approach. To this end, I first define a list of keywords corresponding to each policy and then count the total number of mentions of these keywords in each report. The list of keywords for each policy can be found in Appendix Table A12. I then normalize the keyword count by the total number of words in each report to account for differential length of each report. My second approach is a supervised machine learning approach. To this end, I randomly select 25% of the sentences from all reports and manually label them

[33] Jiang, Meng and Zhang (2019) use prefecture-level government work reports and an unsupervised Latent Dirichlet Allocation topic model to examine social welfare policies at the prefecture level. Campante, Chor and Li (2022) also utilize prefecture-level government work reports, employing both a dictionary and a supervised machine learning approach, to measure governments' emphasis on political stability.

[34] Unlike the provincial-level or prefecture-level government websites, the county government websites have no uniform layouts, and much of the collection has to be done manually, which is a laborious process.

[35] This is done using a simple random sampling method. Namely, each of the 26 provinces in my full sample is first assigned a unique number from 1 to 26, then three random numbers between 1 and 26 are generated using a random number generator without replacement. The provinces corresponding to these three numbers are the randomly selected provinces.

[36] For policy experimentation, Xu (2011) uses the launch of special economic zones (SEZs) in *Shenzhen* and *Zhuhai* as an example to illustrate the pivotal role played by policy experimentation in China's economic rise.

as belonging to each policy or not. I then apply two commonly used machine learning algorithms: *random forest (RF)* and *support vector machine (SVM)* (Gentzkow, Kelly and Taddy, 2019). These algorithms predict a binary policy score at the sentence level. I then construct a report-level policy score by taking the average of these scores weighted by sentence length.

Table 5 reports the effect of the reform on local government policies. The estimates are similar across both the key words frequency policy measures (panel A) and the machine learning-based policy measures (RF in panel B and SVM in panel C). When evaluated using the standardized policy index in column (1), the estimates indicate a roughly 50% standard deviation increase in local governments' emphasis on policies conducive to economic development after the reform. The estimates on individual policies in columns (2)-(5) reveal that this policy shift is driven by increased emphasis on business attraction and market reform, instead of infrastructure investment and policy experimentation. The null effect on infrastructure investment is consistent with criticisms about the sustainability of infrastructure-driven growth, considering the already substantial infrastructure stock (Zilibotti, 2017) in the 2010s. The null effect on policy experimentation is consistent with the discouraging effect of improved economic statistics on policy experimentation (Binswanger and Oechslin, 2020).[37] The event study graphs in Figure 6 confirm these patterns, showing no pre-trends before the reform and a subsequent positive and sustained shift in both the policy index and the two individual policies (business attraction and market reform).

Overall, these results suggest a shift of government policies in directions conducive to economic development after the reform. One remaining concern is that this policy shift may only reflect local officials' visions instead of tangible actions. I will address this concern in subsequent sections by further demonstrating positive effects on bank credit, firm entry, and citizen attitudes towards local officials.

Robustness I conduct two robustness tests in the Appendix. First, I employ a randomization inference procedure with 2,000 permutations to alleviate concerns about the small sample size, which may lead to distortions in conventional inference (Young, 2019). As shown in Appendix

[37] Specifically, a local government undertaking a policy experiment is less likely to receive the *benefit of the doubt* if the true numbers revealed by improved economic statistics suggest a failure of past attempts (Binswanger and Oechslin, 2020).

Figure A8, the randomization inference p-values are similar to conventional p-values. Second, I present IV estimates in Appendix Table A13 (average effect) and Figure A9 (event study) using the aforementioned randomly assigned rural survey teams in 1984 as an instrument for the treatment. The IV estimates have no substantial changes compared to the OLS estimates.

5.2 Bank credit

In many parts of the world including China, one important way that local officials could affect the economy is through their intervention in financial markets, particularly in the allocation of bank credit to firms (La Porta, Lopez-de Silanes and Shleifer, 2002; Dinç, 2005; Carvalho, 2014; Ru, 2018; Cong et al., 2019). To corroborate the previous findings on improvements in government policies, I turn to examining bank credit allocation following the reform in this section.

I collect disaggregated data on bank loans and branches at the county level from the China Banking Regulatory Commission for various types of banks. One drawback of such data is that it only covers the period 2006 to 2011, which limits the study of long-term effect. Nevertheless, it would still be reassuring if one finds a short-term effect. I construct four measures of bank credit, which are *total amount of loans*, *amount of loans to small firms*, *number of firms granted loans*, and *number of bank branches with loan services*. My focus is on the total amount of loans, which can be viewed as an aggregate outcome on bank credit, thus alleviating issues with multiple hypothesis testing. As the distributions of these variables are highly skewed, I apply the inverse hyperbolic sine transformation to reduce influences from the tails of the skewed distributions, and check robustness using untransformed variables. This transformation approximates the logarithm transformation but is well defined at zero (Bellemare and Wichman, 2020).[38]

Panel A of Table 6 shows that the reform generated positive effects on all four measures of bank credit. While the estimates are less precise, the economic magnitudes are generally larger than 10% except for the number of bank branches with loan services. To benchmark such magnitudes: Colonnelli and Prem (2022) show that random audits on local governments

[38] The inverse hyperbolic sine transformation of a variable x is: $IHS(X) = \ln(x + \sqrt{x^2 + 1})$.

in Brazil increased bank loans by about 3%. In addition, the effect on loans to small firms is the largest and statistically significant (coef.=0.263, s.e.=0.116). The role of small and medium enterprises (SMEs) in economic development is well documented in the literature and has been instrumental in China's recent economic progress. However, it is also true that SMEs face severe credit constraints (Beck, Demirgüç-Kunt and Maksimovic, 2008; Ayyagari, Demirgüç-Kunt and Maksimovic, 2010). The significant increase in loans to small firms thus speaks to positive policy shifts by local governments. The event study graphs in Figure 7 further confirms such patterns and show no pre-trends before the reform.

One concern with the previous difference-in-differences estimates is that they may capture credit demand instead of a government-led credit supply. To address this concern, I further conduct a difference-in-difference-in-diffferences (DDD) estimation exploiting differential control of banks by local governments across counties. To this end, I construct a county-level index $GovernmentControl_c$ as follows:

$$GovernmentControl_c = \sum_b LoanShare_{cb}^{pre} \times GovernmentControl_b$$

where $LoanShare_{cb}^{pre}$ denotes the share of bank b in county c's loan market prior to the reform, and $GovernmentControl_b$ denotes bank b's degree of control by local governments. I set $GovernmentControl_b$ to be 1 for City Commercial Banks (CCBs) whose controlling shareholders are local governments, and 0 for other banks. The county-level index $GovernmentControl_c$ is further standardized to ease interpretation. I then estimate the following DDD specification:

$$\begin{aligned} CreditOutcome_{ct} = &\ \beta_1 Treat_c \times GovernmentControl_c \times Post_t \\ &+ \beta_2 Treat_c \times Post_t + \beta_3 GovernmentControl_c \times Post_t \\ &+ \delta_c + \lambda_t + \varepsilon_{ct} \end{aligned} \tag{12}$$

where β_1 is the coefficient of interest capturing the differential impact of the reform on credit outcomes across counties with varying preexisting government control over banks. Panel B of Table 6 reports the results. The estimates on $Treat_c \times GovernmentControl_c \times Post_t$ are large in magnitude and also generally statistically significant. In particular, the estimates suggest that moving from the average county to a county with one standard deviation higher government control over banks would increase the reform's effect on the total amount of loans by around one quarter (coef.=0.237, s.e.=0.103). The event study graphs in Figure 7 further confirm this pattern. In contrast, the estimates on $Treat_c \times Post_t$, while mostly positive, are small and statistically insignificant. These results are consistent with an expansion of credit supply from the government side.

Robustness I conduct two robustness tests in the Appendix. First, recent econometric literature shows that the coefficient estimates with log-like transformations of variables, including the inverse hyperbolic sine transformation, are sensitive to variable units (Mullahy and Norton, 2022; Chen and Roth, 2023).[39] In Appendix Table A15, I use untransformed variables. The findings still hold qualitatively. Second, I present IV estimates in Appendix Table A14 (average effect) and Figure A10 (event study) using the aforementioned randomly assigned rural survey teams in 1984 as an instrument for the treatment. The IV estimates are similar to the OLS estimates.

5.3 Firm entry

In China, establishing a firm involves lengthy bureaucratic procedures, over which local officials have substantial discretion (WorldBank, 2008; Jia, Lan and Padró i Miquel, 2021).[40] Therefore, increased development effort by local officials should manifest as higher firm entry. I test this conjecture in this section using the universe of firm registration data from Dong et al. (2021), which is available for the years 2005, 2010, and 2015.[41] For my main analysis, I aggregate the firm-level registration data at the county $\times$ year level and further exploit the microstructure of the data for robustness. I create four variables representing the number of registrations for each of the four firm types classified by ownership: *private firms*, *state-*

[39]This concern is particularly prominent when variables are frequently observed at zero (Mullahy and Norton, 2022). However, in my case, all four variables have positive values for more than 93% of observations.

[40]In the World Bank's Doing Business report, China was ranked 151st out of 182 countries in 2008 in terms of the ease of starting a business. This ranking takes into account factors such as the number of procedures, time spent on registration, and cost relative to income (WorldBank, 2008).

[41]The authors collected the data by web scraping an online system called the National Enterprise Credit Information Publicity System (NECIPS) (see https://www.gsxt.gov.cn/index.html). In China, a newly established firm needs to register at the local Administration for Industry and Commerce by providing detailed information such as firm name and address. After approval, the relevant information is publicized on the NECIPS.

owned enterprises (SOEs), *foreign-owned firms*, and *collectively owned firms*. To reduce the number of tests, I also create an aggregate variable representing the total number of registrations for all firms, which serves as my primary focus. As before, I apply the inverse hyperbolic sine transformation to these five variables to reduce the influence of outliers in the tails of the skewed outcome distributions and check robustness using untransformed variables (Bellemare and Wichman, 2020).

Panel A of Table 7 reports the results. In aggregate, the reform significantly increased firm entry by about 5% (coef.=0.046, s.e.=0.026), as shown in column (1). To put this effect into perspective: Giannetti et al. (2021) show that China's 2013 anti-corruption campaign increased firm entry by 6.7% for a province-industry that was initially one standard deviation more corrupt than the average. Although the specification differs, it nonetheless provides reassurance that my estimate is of considerable economic significance. When examining the effect by ownership in the remaining columns, I also find increased entry for private firms (coef.=0.048, s.e.=0.027), SOEs (coef.=0.169, s.e.=0.062), and foreign firms (coef.=0.041, s.e.=0.051), consistent with their relatively higher productivity and pivotal role in economic growth (Song, Storesletten and Zilibotti, 2011).[42] In contrast, the effect on collectively owned firms is negative, consistent with the fact that these firms are inefficient, although the estimate is imprecise (coef.=-0.035, s.e.=0.062). The event study graphs in Figure 8 further confirm these patterns and show that the increase in firm entry happened immediately after the reform, although the limitation of the data prevents me from examining the pre-trends.

To further tighten identification and alleviate concerns about pre-trends, I exploit the microstructure of the data to conduct a standard regression discontinuity (RD) design across county borders. I use towns as the unit of observation and collapse the firm-level registration data at the town $\times$ year level.[43] Each town is then assigned to the nearest county border. In cases where a county shares its border with multiple counties, the county's border is divided into multiple segments so that there is only one county on each side of a border segment. I then

[42]The positive and larger effect on SOEs may seem puzzling as they are generally considered less productive, but SOE productivity has been converging with that of private firms after nearly a decade of productivity-enhancing reforms in the state sector since the late 1990s (Hsieh and Song, 2015). See Appendix Figure A7 for the dynamics of firm TFP by ownership.

[43]To deal with the relatively frequent consolidations of towns, I map all the outcomes and covariates to towns based on a 2010 map. On average, a county contains about 15 towns in 2010.

focus on towns along county borders with different treatment statuses on each side. I estimate the following local linear regression with a uniform kernel:

$$FirmEntry_i = \beta_1 Treat_i + \beta_2 Treat_i \times Distance_i + \beta_3 Distance_i + \delta_{b(i)} + \varepsilon_i \tag{13}$$
$$s.t. \quad -h < Distance_i < h$$

where $FirmEntry_i$ denotes the number of firm registrations in town i, either aggregate or by ownership.[44] $Treat_i$ is equal to 1 if town i is located in a treatment county and 0 otherwise. $Distance_i$ is the distance from the centroid of town i to the nearest county border, and is negative if $Treat_i$=0. To ensure that treatment and control towns are comparable, I include county border fixed effects $\delta_{b(i)}$ to restrict the comparison to be within a narrowly defined geographic area. I estimate the equation using the optimal bandwidth h proposed by Calonico, Cattaneo and Titiunik (2014) with standard errors clustered at the county border level.

The identification assumption of this RD design is that all other factors affecting firm entry should evolve smoothly across county borders, except for the reform. To check this, Appendix Figure A12 conducts a balance test on preexisting town covariates, which show no significant jumps at county borders. In addition, RD designs utilizing administrative borders may suffer from compound treatments issues if there are other institutional or regulatory differences across the borders (Keele and Titiunik, 2015). However, as long as these potential differences are not related to firm entry, the RD design is still valid. To check this, Appendix Figure A12 further presents a placebo RD estimation using pre-reform firm entry data and finds no discontinuities at county borders. The RD estimates on post-reform firm entry are presented in Figure 8 and Panel B of Table 7. Consistent with the DID estimates, the RD estimates show that the reform boosted firm entry, especially for those with higher productivity, which further corroborates the previous findings of greater effort exerted by local officials in economic development after the reform.

Robustness I conduct some tests to check the robustness of the findings in this section. First,

[44] Bellemare and Wichman (2020) suggest using the inverse hyperbolic sine transformation for large values of outcomes (e.g., larger than 10). Given that the number of firm entry at the town level is in general smaller than 10, I use the raw number here.

for the DID estimation, I show in Appendix Table A16 that the estimates are robust to using untransformed firm entry data. I also present IV estimates in Appendix Table A17 (average effect) and Appendix Figure A11 (event study) using the aforementioned randomly assigned rural survey teams in 1984 as an instrument for the treatment. The patterns are similar. Second, for the RD estimation, I show in Appendix Table A18 that the estimates are robust to alternative bandwidth (Imbens and Kalyanaraman, 2012), quadratic polynomial, and triangular kernel. I also show in Appendix Figure A13 that the RD estimates are robust to a randomization inference procedure to address concerns about inference in RD designs (Ganong and Jäger, 2018).[45]

5.4 Citizen attitudes

Finally, increased development effort by local officials should be reflected in the attitudes of citizens towards local governments. To test this conjecture, I use survey data from the China Family Panel Studies (CFPS) and pool three waves, 2012, 2014 and 2016, together.[46] I focus on citizens' *trust in local officials* and *evaluation of local government performance*, two commonly used measures of the performance of government officials (Bertrand et al., 2020; Martinez-Bravo et al., 2022).[47] For the former, the survey asks: "To what extent do you trust local officials?" The answer is an integer from 0 to 10 with larger values denoting higher trust. For the latter, the survey asks: "What is your overall evaluation of the county government's achievements last year?" The answer is one of the following: significant achievement, some achievement, not much achievement, no achievement, worse than before. To ease interpretation, I create a dummy variable indicating some or significant achievement.[48] To tighten identification, I create two placebo variables utilizing another two questions. The first question asks: "Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?" The answer is binary, either yes or no, and I use a dummy

[45] Specifically, I keep only county borders across which there are no treatment variations, and then randomly create placebo treatment variation to each border. Based on these placebo borders, I then re-estimate equation 13. This process is repeated for 2,000 times.

[46] The baseline wave in 2010 is not used as it may take time for both local governments to take actions and the citizens to change attitudes. In addition, the two key variables I examine below are missing in the 2010 wave.

[47] One issue with these variables is that citizens may not express their opinions faithfully. To alleviate this issue, I drop the top decile of citizens who show the highest concerns about the survey, which are observed and recorded by the investigators. Results are similar if I instead control for citizens' concerns.

[48] Because of the categorical nature of the answer to the second question, I do not summarize these two questions into one. I instead correct for multiple hypothesis testing in Appendix Table A22 and the findings still hold.

variable to denote yes. The second question describes a hypothetical person with minor health issues and then asks: "What do you think about the health condition of the person?" The answer is one of the following: extremely healthy, very healthy, relatively healthy, average, not healthy. As before, I create a dummy variable indicating extremely or very healthy.

Due to the lack of pre-reform survey data, I follow the empirical strategy in Duflo (2001) to estimate a cohort DiD specification that utilizes two sources of variation: (1) treatment counties versus control counties; (2) most affected cohorts versus less unaffected cohorts within the same county. The latter source of variation is built on insights in the psychology and political science literature that citizens' political attitudes are most permeable during teenage years and keep stable since one's 30s (Wolfinger and Rosenstone, 1980; Krosnick and Alwin, 1989). I create three cohort groups: those born in the 1970s, in the 1980s, and in the 1990s.[49] The 1970s cohort are the unaffected group as they would be older than 30 during the reform period. The 1990s cohort are the most affected group as they would be younger than 30 during the reform period, and thus, their political attitudes would be most permeable. The 1980s cohort are defined as the less affected group because some of them would be older than 30 during the reform period. The estimation equation is:

$$Y_{icg} = \sum_{k=1980s,1990s,k\neq 1970s} \beta_k Treat_c \times 1_{\{g=k\}} + \delta_c + \lambda_g + W_i\Omega + X_{cg}\Psi + \varepsilon_{icg} \tag{14}$$

where Y_{icg} denotes attitude measures for citizen i living in county c and born in cohort $g \in \{1970s, 1980s, 1990s\}$. The 1970s cohort is the omitted reference group. W_i denotes a set of citizen controls, including years of schooling and its square, age and its square, dummy for male, dummy for living in the urban area, and dummy for survey wave. δ_c and λ_g are county and cohort fixed effects, which help to partial out county- and cohort-specific time-invariant confounding factors, respectively. To the extent that the treatment $Treat_c$ may be correlated with county characteristics, which may have differential impacts on citizen attitudes, I include in X_{cg} the same set of county controls as before, interacted with cohort fixed effects.

[49] Older cohorts, such as those born in the 1950s and the 1960s, are not used as controls in my main analysis as they grew up in turbulent times when China suffered from several catastrophic events (e.g., the Great Famine and the Cultural Revolution). These events may affect trust formation (Chen and Yang, 2015; Bai and Wu, 2020), making them less comparable to younger cohorts. However, as shown in Appendix Table A19, results are similar if I include these older cohorts as controls.

Table 8 presents the results, which are also visualized in Figure 9. Panel A examines the effect on citizens' trust in local officials. As shown in column (1), the reform significantly increased the 1990s cohort's trust in local officials (coef.=0.452, s.e.=0.154). In terms of magnitude, the estimate indicates a 9.3% increase in trust relative to the mean trust across all three cohorts. In contrast, there is no change in citizens' trust in most people as shown in column (2). The estimates are not only small in magnitude but also statistically insignificant. When examining the effect by survey wave in columns (3)-(5), I find a similar patter. In addition, the estimates, albeit with less precision, indicate that the reform changed citizens' attitudes in only two years. This may seem striking considering that attitudes often change gradually, but can be reconciled with the immediate drop in GDP growth manipulation after the reform as shown previously. Panel B examines the effect on citizens' evaluation of local government performance. As shown in column (1), the reform positively shifted the younger cohorts' attitudes, especially for the 1990s cohort (coef.=0.052, s.e.=0.020). Relative to the sample mean, the estimate indicate a 6.3% increase for the 1990s cohort. The placebo estimates on citizens' evaluation of others' health are again small in magnitude and statistically insignificant as shown in column (2). The remaining columns examine the effect by survey wave and show that the improvement was concentrated in the 2016 wave, which is reasonable as it takes time for local governments to make tangible achievements. Taken together, these results bolster the previous findings that local officials exerted more effort in developing the economy after the reform.

Robustness I provide several tests to check the robustness of these findings. First, I adopt an alternative definition of affected and unaffected cohorts in Appendix Table A19, where affected cohort consist of only those born in the 1990s and unaffected cohort consist of those born in the 1970s or older. Second, I control for citizens' media access in Appendix Table A20, which could alleviate the concern that treatment counties may be better at propaganda that could disproportionately affect the young cohort if they have greater access than other cohorts.[50] Third, I provide IV estimates using the aforementioned randomly assigned rural survey teams

[50] Specifically, I include three variables constructed from the survey regarding media access: (1) the number of days political news was accessed via television in the last week; (2) the number of days political news was accessed via Internet in the last week; (3) whether you have posted comments related to political issues and major national events on Internet in the past 12 months.

in 1984 as an instrument for the treatment, which are presented in Appendix Figure A14 and Appendix Table A21. The results are essentially unchanged.

5.5 Alternative explanations

The results across the four domains examined above are consistent with greater development effort exerted by local officials after the reform. In this section, I examine a few alternative explanations that may generate observationally equivalent results.

Fiscal transfers Local officials' differential performance may result from varying fiscal transfers from upper-level governments, given the well-documented development impacts of such transfers (Litschig and Morrison, 2013; Corbi, Papaioannou and Surico, 2019). While I cannot directly test this possibility due to the lack of data on county-level fiscal transfers after 2007,[51] I provide some suggestive evidence showing that this is unlikely. Conceptually, there are two plausible causes of varying fiscal transfers from upper-level governments after the reform. First, the upper-level government may decrease the transfers as an implicit punishment of subordinates' misconduct, but this would then work against from finding a positive impact on local officials' performance. Second, the upper-level government may increase the transfers if better data provided by the survey teams increased the "creditworthiness" of their subordinate counties. If this is true, then one should also find a similar increase in the transfers around the launch of the survey teams in 2005, as they collected additional information about the county, which should play a similar credit-enhancing role. To check this, I collect data on fiscal transfers up to 2007 from the China Prefecture, City, and County Public Finance Statistics. As shown in Appendix Figure A15 and Appendix Table A23, I find little change in fiscal transfers after 2005.

Policy diffusion The arrival of the survey teams may facilitate policy diffusion, either from upper-level governments or among similar local governments (Shipan and Volden, 2008; Wang and Yang, 2021; DellaVigna and Kim, 2022), which may eventually improve local officials'

[51] Data on county-level fiscal transfers could be collected from the China Prefecture, City, and County Public Finance Statistics published by the Ministry of Finance of China (Jia, Liang and Ma, 2021). However, the publication of such data stopped after 2007.

performance. Conceptually, as the survey teams were deployed years before the reform, one should see a pre-trend in local officials' performance if the previous findings are driven by policy diffusion, which is not the case. To further rule out the possibility of varying policy diffusion after the reform, I examine the similarity of local government work reports across counties, with the premise that greater policy diffusion after the reform should lead to a convergence of these reports among treatment counties. Specifically, I estimate:

$$Similarity_{ijt} = \beta Treat_{ij} \times Post_t + \delta_{ij} + \lambda_{it} + \gamma_{jt} + \varepsilon_{ijt} \tag{15}$$

where ij indicates county pairs ($i \neq j$), with ij being equivalent to ji . $Similarity_{ijt}$ denotes the pairwise textual similarity of government work reports, which is calculated following Kelly et al. (2021) and described in Appendix C. $Treat_{ij}$ equals 1 if both i and j are treatment counties, and 0 otherwise. I include county × year fixed effects (λ_{it} and γ_{jt}) and county pair fixed effects (δ_{ij}) to account for county-specific traits and inherent differences between counties in a pair, respectively. Standard errors are two-way clustered by both counties in a pair. The results are reported in Appendix Table A24. Column (1) considers all county pairs to account for both types of policy diffusion. Column (2) considers county pairs within the same province to examine policy diffusion among local governments. Column (3) considers county pairs spanning different provinces to examine policy diffusion from the central government. The results are generally small and statistically insignificant.[52] Appendix Figure A16 provides the event study graphs, further confirming the absence of policy diffusion.

Short termism Local officials may shift effort to other short-termist behaviors after the reform, such as prioritizing growth over other factors (e.g., social welfare and environmental protection) or overleverage through shadow banking (Xiong, 2018). As a result, one may also observe improvements in local officials' performance in economic development after the reform. Appendix Table A25 and Appendix Figure A17 examine local officials' emphasis on social welfare and environmental protection in government work reports, using both a keywords frequency approach and a machine learning approach, as described in Appendix C. The estimates

[52] If anything, the results suggest a slight decrease in pairwise similarity of county work reports for county pairs across provinces (coef.=-0.002, s.e.=0.001), but is small in magnitude relative to the mean (0.45).

are generally small and statistically insignificant. I then examine local government debt. To this end, I collect data on bond issuance by local government financing vehicles (LGFVs) from the Wind database, which serves as a proxy for local government debt.[53] The earliest issuance at the county level was in 2009. As such, I collapse the bond issuance data by county and estimate a cross-sectional regression using the randomly assigned rural survey teams in 1984 as the instrument for treatment counties. Appendix Table A26 reports the estimates, which are small and statistically insignificant.

Corruption A reduction in corruption, which is possible if the reform increased the *perceived cost* of all types of misconduct, may also generate observationally equivalent results given the distorting effect of corruption on the economy (Giannetti et al., 2021; Colonnelli and Prem, 2022). I examine this possibility using a comprehensive dataset on corruption convictions complied by Wang and Dickson (2022).[54] The dataset contains 10,797 corruption convictions from 2005 to 2016, with a vast majority (10,788) happening after 2012 when China's anti-corruption campaigns began. The few convictions (9) before 2012 were likely caused by the lack of enforcement instead of less corruption, and are dropped from my analysis. I then collapse the data by county and estimate a cross-sectional regression using the randomly assigned rural survey teams in 1984 as the instrument for treatment counties. The results are reported in Appendix Table A27. Throughout the table, I include the number of anti-corruption inspections using the data from Wang (2021) to address concerns about differential anti-corruption enforcement (Glaeser and Saks, 2006; Zhu, 2017). The estimates are generally small and statistically insignificant, both in aggregate (column 1) and by corruption types (columns 2-3).[55] To alleviate concerns about potential lags between corruption and subsequent convictions, which may result in a null reform effect if all convictions reflecting corruption before 2009, I examine more recent

[53]In China, local governments are prohibited from borrowing from banks or issuing bonds directly (Huang, Pagano and Panizza, 2020). Instead, they could set up LGFVs and then issue bonds through them, usually with land offered by local governments as collateral. In addition, these LGFVs could also borrow from banks, but such loans are usually not disclosed.

[54]Wang and Dickson (2022) collect the data from Tencent—the largest Internet company in China. In 2011 Tencent launched a searchable online database of all corruption convictions across China, and the authors scraped the website in August 2016. Unfortunately, the website is closed currently.

[55]The major type of corruption contains bribery and appropriation of public property. The remaining types include other misbehaviours that are also considered as corruption in China, such as sex scandals (but no data manipulation).

convictions in columns (4)-(6), namely, those in 2015 and 2016.[56] The results have no substantial change. To further corroborate these findings, I also utilize the previous CFPS survey and equation (14) to estimate the effect of the reform on citizens' perceived corruption about the government.[57] I still find no effect as shown in Appendix Figure A18 and Appendix Table A28.

Leveraging China's 2009 reform punishing economic data manipulation and counties' quasi-random reform exposure, I provide rich causal evidence showing that the reform led to: (1) a decrease in GDP growth manipulation amounting to 5% of reported GDP growth, driven by a *reputational discipline* effect; (2) an increase in politicians' development effort manifested in both policy rhetoric and downstream impacts, consistent with an *effort reallocation* effect. These results highlight the relevance of reputational punishment in weak institutional settings and the cost of overlooked bureaucratic misconduct distinct from corruption.

6 Conclusion

Manipulation of official statistics by government agents is a common phenomenon in the world, as evident in both anecdote evidence and academic research. A direct consequence of such manipulation is information distortion within the bureaucracy. The indirect consequences, however, are not well recognized. This paper focuses on the China setting to provide causal evidence on how to reduce such manipulation and whether reducing it could generate economic benefits. Utilizing multiple datasets, a unique reform targeting economic data manipulation, and multiple identification strategies including an instrumented difference-in-differences design, this papers shows that a combination of top-down monitoring and punishment could effectively reduce such manipulation. Moreover, curbing such manipulation further elicited local officials' development effort. As such, the reform generated significant downstream impacts. In certain domains such as credit allocation and firm entry, the impacts are comparable to those of an anti-corruption campaign documented in the literature both in China and Brazil (Giannetti et al., 2021; Colonnelli and Prem, 2022).

[56]The corruption convictions in 2015 and 2016 account for 61% of all convictions.

[57]The survey question asks: "In general, how serious do you think the problem of government corruption is in our country?" The answer ranges from 0 to 10, with larger integers denoting higher perceived corruption. This question is only available in the 2014 and 2016 waves.

These striking downstream impacts provide a new perspective to understand how individual local officials could affect the macro-level economic outcomes. Unlike political corruption, which could directly affect resource allocation and further stifle economic development (Krueger, 1974; Shleifer and Vishny, 1993), economic data manipulation has a more subtle impact—it hurts the economy by inducing an unfavorable shift of local officials' effort. This subtle impact may explain why such manipulation is not well recognized as a detrimental bureaucratic misconduct, despite its ubiquity. In this vein, the findings have implications for the design of relevant interventions targeting such manipulation.

This study has some limitations, and I outline some directions for future research. First, the question of how such manipulation affects the entire economy is essentially a general equilibrium question. The research design in this paper, by construction, may only estimate the lower bound and partial equilibrium effect. Future work may explore this impact using structural approaches. Second, future work could enrich our understanding of the impacts of such manipulation by studying the effect on firm performance using firm census data.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2023. "When should you adjust standard errors for clustering?" *The Quarterly Journal of Economics*, 138(1): 1–35.

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond. 2019. "Preferences for truth-telling." *Econometrica*, 87(4): 1115–1153.

Agan, Amanda, Jennifer L Doleac, and Anna Harvey. 2023. "Misdemeanor prosecution." *The Quarterly Journal of Economics*, 138(3): 1453–1505.

Agarwal, Sumit, and Robert Hauswald. 2010. "Distance and private information in lending." *The Review of Financial Studies*, 23(7): 2757–2788.

Agarwal, Vikas, Naveen D Daniel, and Narayan Y Naik. 2011. "Do hedge funds manage their reported returns?" *The Review of Financial Studies*, 24(10): 3281–3320.

Aghion, Philippe, and Jean Tirole. 1997. "Formal and real authority in organizations." *Journal of political economy*, 105(1): 1–29.

Al-Ubaydli, Omar, John A List, and Dana Suskind. 2019. "The science of using science: Towards an understanding of the threats to scaling experiments." National Bureau of Economic Research.

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American statistical Association*, 103(484): 1481–1495.

Andrews, Isaiah, James H Stock, and Liyang Sun. 2019. "Weak instruments in instrumental variables regression: Theory and practice." *Annual Review of Economics*, 11: 727–753.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2014. *Mastering'metrics: The path from cause to effect.* Princeton university press.

Avis, Eric, Claudio Ferraz, and Frederico Finan. 2018. "Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians." *Journal of Political Economy*, 126(5): 1912–1964.

Ayyagari, Meghana, Asli Demirgüç-Kunt, and Vojislav Maksimovic. 2010. "Formal versus informal finance: Evidence from China." *The Review of Financial Studies*, 23(8): 3048–3097.

Bai, Liang, and Lingwei Wu. 2020. "Political movement and trust formation: Evidence from the Cultural Revolution (1966–76)." *European Economic Review*, 122: 103331.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2009. "Social connections and incentives in the workplace: Evidence from personnel data." *Econometrica*, 77(4): 1047–1094.

Banerjee, Abhijit, Sendhil Mullainathan, and Rema Hanna. 2012. "Corruption." National Bureau of economic research.

Becker, Gary S, and George J Stigler. 1974. "Law enforcement, malfeasance, and compensation of enforcers." *The Journal of Legal Studies*, 3(1): 1–18.

Beck, Thorsten, Asli Demirgüç-Kunt, and Vojislav Maksimovic. 2008. "Financing patterns around the world: Are small firms different?" *Journal of financial economics*, 89(3): 467–487.

Bellemare, Marc F, and Casey J Wichman. 2020. "Elasticities and the inverse hyperbolic sine transformation." *Oxford Bulletin of Economics and Statistics*, 82(1): 50–61.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *The Review of Economic Studies*, 81(2): 608–650.

Benabou, Roland, and Jean Tirole. 2011. "Laws and norms." National Bureau of Economic Research.

Benmelech, Efraim, Eugene Kandel, and Pietro Veronesi. 2010. "Stock-based compensation and CEO (dis) incentives." *The Quarterly Journal of Economics*, 125(4): 1769–1820.

Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu. 2020. "The glittering prizes: Career incentives and bureaucrat performance." *The Review of Economic Studies*, 87(2): 626–655.

Besley, Timothy J, Robin Burgess, Adnan Khan, and Guo Xu. 2021. "Bureaucracy and development."

Bhuller, Manudeep, Gordon B Dahl, Katrine V Løken, and Magne Mogstad. 2020. "Incarceration, recidivism, and employment." *Journal of Political Economy*, 128(4): 1269–1324.

Binswanger, Johannes, and Manuel Oechslin. 2020. "Better statistics, better economic policies?" *European Economic Review*, 130.

Bossuroy, Thomas, Clara Delavallade, and Vincent Pons. 2019. "Biometric tracking, healthcare provision, and data quality: experimental evidence from tuberculosis control." National bureau of economic research.

Bruhn, Miriam, and David McKenzie. 2009. "In pursuit of balance: Randomization in practice in development field experiments." *American economic journal: applied economics*, 1(4): 200–232.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2014. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica*, 82(6): 2295–2326.

Campante, Filipe R, Davin Chor, and Bingjing Li. 2022. "The Political Economy Consequences of China's Export Slowdown." Working Paper.

Carvalho, Daniel. 2014. "The real effects of government-owned banks: Evidence from an emerging market." *The Journal of Finance*, 69(2): 577–609.

Chen, Jiafeng, and Jonathan Roth. 2023. "Logs with zeros? Some problems and solutions." Working Paper.

Chen, Yi, Ziying Fan, Xiaomin Gu, and Li-An Zhou. 2020. "Arrival of young talent: The send-down movement and rural education in China." *American Economic Review*, 110(11): 3393–3430.

Chen, Yuyu, and David Yang. 2015. "Historical traumas and the roots of political distrust: Political inference from the Great Chinese Famine." *Available at SSRN 2652587*.

Colonnelli, Emanuele, and Mounu Prem. 2022. "Corruption and Firms." *The Review of Economic Studies*, 89(2): 695–732.

Cong, Lin William, Haoyu Gao, Jacopo Ponticelli, and Xiaoguang Yang. 2019. "Credit allocation under economic stimulus: Evidence from China." *The Review of Financial Studies*, 32(9): 3412–3460.

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of econometrics*, 92(1): 1–45.

Corbi, Raphael, Elias Papaioannou, and Paolo Surico. 2019. "Regional transfer multipliers." *The Review of Economic Studies*, 86(5): 1901–1934.

DellaVigna, Stefano, and Woojin Kim. 2022. "Policy diffusion and polarization across US states." National Bureau of Economic Research.

Dinç, I Serdar. 2005. "Politicians and banks: Political influences on government-owned banks in emerging markets." *Journal of financial economics*, 77(2): 453–479.

Dong, Lei, Xiaohui Yuan, Meng Li, Carlo Ratti, and Yu Liu. 2021. "A gridded establishment dataset as a proxy for economic activity in China." *Scientific Data*, 8(1): 1–9.

Duflo, Esther. 2001. "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment." *American economic review*, 91(4): 795–813.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using randomization in development economics research: A toolkit." *Handbook of development economics*, 4: 3895–3962.

Fischer, Paul E, and Robert E Verrecchia. 2000. "Reporting bias." *The Accounting Review*, 75(2): 229–245.

Fisman, Raymond, Daniel Paravisini, and Vikrant Vig. 2017. "Cultural proximity and loan outcomes." *American Economic Review*, 107(2): 457–492.

Fisman, Raymond, Jing Shi, Yongxiang Wang, and Weixing Wu. 2020. "Social ties and the selection of China's political elite." *American Economic Review*, 110(6): 1752–1781.

Ganong, Peter, and Simon Jäger. 2018. "A permutation test for the regression kink design." *Journal of the American Statistical Association*, 113(522): 494–504.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as data." *Journal of Economic Literature*, 57(3): 535–574.

Ghanem, Dalia, and Junjie Zhang. 2014. "'Effortless Perfection:'Do Chinese cities manipulate air pollution data?" *Journal of Environmental Economics and Management*, 68(2): 203–225.

Giannetti, Mariassunta, Guanmin Liao, Jiaxing You, and Xiaoyun Yu. 2021. "The externalities of corruption: Evidence from entrepreneurial firms in China." *Review of Finance*, 25(3): 629–667.

Glaeser, Edward L, and Raven E Saks. 2006. "Corruption in america." *Journal of public Economics*, 90(6-7): 1053–1072.

Gneezy, Uri, Agne Kajackaite, and Joel Sobel. 2018. "Lying aversion and the size of the lie." *American Economic Review*, 108(2): 419–453.

Goldman, Eitan, and Steve L Slezak. 2006. "An equilibrium model of incentive contracts in the presence of information manipulation." *Journal of Financial Economics*, 80(3): 603–626.

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

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. 2022. "Can technology solve the principal-agent problem? Evidence from China's war on air pollution." *American Economic Review: Insights*, 4(1): 54–70.

Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political Analysis*, 20(1): 25–46.

Hart, Oliver. 1995. *Firms, contracts, and financial structure.* Clarendon press.

Henderson, J Vernon, Adam Storeygard, and David N Weil. 2012. "Measuring economic growth from outer space." *American Economic Review*, 102(2): 994–1028.

Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design." *The Journal of Law, Economics, and Organization*, 7(special_issue): 24–52.

Hsieh, Chang-Tai, and Zheng Michael Song. 2015. "Grasp the large, let go of the small: The transformation of the state sector in China." National Bureau of Economic Research.

Huang, Yi, Marco Pagano, and Ugo Panizza. 2020. "Local crowding-out in China." *The Journal of Finance*, 75(6): 2855–2898.

Huber, Kilian. 2023. "Estimating general equilibrium spillovers of large-scale shocks." *The Review of Financial Studies*, 36(4): 1548–1584.

Iacus, Stefano M, Gary King, and Giuseppe Porro. 2012. "Causal inference without balance checking: Coarsened exact matching." *Political analysis*, 20(1): 1–24.

Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal bandwidth choice for the regression discontinuity estimator." *The Review of Economic Studies*, 79(3): 933–959.

Imbens, Guido W, and Donald B Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences.* Cambridge University Press.

Jia, Junxue, Xuan Liang, and Guangrong Ma. 2021. "Political hierarchy and regional economic development: Evidence from a spatial discontinuity in China." *Journal of Public Economics*, 194: 104352.

Jiang, Junyan, Tianguang Meng, and Qing Zhang. 2019. "From Internet to social safety net: The policy consequences of online participation in China." *Governance*, 32(3): 531–546.

Jia, Ruixue, Masayuki Kudamatsu, and David Seim. 2015. "Political selection in China: The complementary roles of connections and performance." *Journal of the European Economic Association*, 13(4): 631–668.

Jia, Ruixue, Xiaohuan Lan, and Gerard Padró i Miquel. 2021. "Doing business in China: Parental background and government intervention determine who owns busines." *Journal of Development Economics*, 151: 102670.

Keele, Luke J, and Rocio Titiunik. 2015. "Geographic boundaries as regression discontinuities." *Political Analysis*, 23(1): 127–155.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy. 2021. "Measuring technological innovation over the long run." *American Economic Review: Insights*, 3(3): 303–320.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica*, 75(1): 83–119.

Kofanov, Dmitrii, Vladimir Kozlov, Alexander Libman, and Nikita Zakharov. 2023. "Encouraged to Cheat? Federal Incentives and Career Concerns at the Sub-national Level as Determinants of Under-Reporting of COVID-19 Mortality in Russia." *British Journal of Political Science*, 53(3): 835–860.

Kofman, Fred, and Jacques Lawarrée. 1993. "Collusion in hierarchical agency." *Econometrica: Journal of the Econometric Society*, 629–656.

Kou, Chien-wen, and Wen-Hsuan Tsai. 2014. ""Sprinting with small steps" towards promotion: solutions for the age dilemma in the CCP cadre appointment system." *The China Journal*, 71(1): 153–171.

Krosnick, Jon A, and Duane F Alwin. 1989. "Aging and susceptibility to attitude change." *Journal of Personality and Social Psychology*, 57(3): 416.

Krueger, Anne O. 1974. "The political economy of the rent-seeking society." *The American economic review*, 64(3): 291–303.

Landry, Pierre F, Xiaobo Lü, and Haiyan Duan. 2018. "Does performance matter? Evaluating political selection along the Chinese administrative ladder." *Comparative Political Studies*, 51(8): 1074–1105.

La Porta, Rafael, Florencio Lopez-de Silanes, and Andrei Shleifer. 2002. "Government ownership of banks." *The Journal of Finance*, 57(1): 265–301.

Lazear, Edward P, and Sherwin Rosen. 1981. "Rank-order tournaments as optimum labor contracts." *Journal of Political Economy*, 89(5): 841–864.

Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack R Porter. 2022. "Valid t-ratio Inference for IV." *American Economic Review*.

Levinsohn, James, and Amil Petrin. 2003. "Estimating production functions using inputs to control for unobservables." *The review of economic studies*, 70(2): 317–341.

Li, Hongbin, and Li-An Zhou. 2005. "Political turnover and economic performance: the incentive role of personnel control in China." *Journal of Public Economics*, 89(9-10): 1743–1762.

Li, Pei, Yi Lu, and Jin Wang. 2016. "Does flattening government improve economic performance? Evidence from China." *Journal of Development Economics*, 123: 18–37.

Litschig, Stephan, and Kevin M Morrison. 2013. "The impact of intergovernmental transfers on education outcomes and poverty reduction." *American Economic Journal: Applied Economics*, 5(4): 206–240.

Li, Xuecao, Yuyu Zhou, Min Zhao, and Xia Zhao. 2020. "A harmonized global nighttime light dataset 1992–2018." *Scientific Data*, 7(1): 1–9.

Lyu, Changjiang, Kemin Wang, Frank Zhang, and Xin Zhang. 2018. "GDP management to meet or beat growth targets." *Journal of Accounting and Economics*, 66(1): 318–338.

Mahadevan, Meera. 2023. "The price of power: Costs of political corruption in Indian electricity." *Working Paper*.

Ma, Matthew, Jing Pan, and Stephen R Stubben. 2020. "The effect of local tournament incentives on firms' performance, risk-taking decisions, and financial reporting decisions." *The Accounting Review*, 95(2): 283–309.

Marbach, Moritz, and Dominik Hangartner. 2020. "Profiling compliers and noncompliers for instrumental-variable analysis." *Political Analysis*, 28(3): 435–444.

Martinez-Bravo, Monica, Gerard Padró i Miquel, Nancy Qian, and Yang Yao. 2022. "The rise and fall of local elections in China." *American Economic Review*, 112(9): 2921–58.

Martinez, Luis R. 2022. "How much should we trust the dictator's GDP growth estimates?" *Journal of Political Economy*, 130(10): 2731–2769.

Maskin, Eric, Yingyi Qian, and Chenggang Xu. 2000. "Incentives, information, and organizational form." *The Review of Economic Studies*, 67(2): 359–378.

Meyer, David, Victor C Shih, and Jonghyuk Lee. 2016. "Factions of different stripes: gauging the recruitment logics of factions in the reform period." *Journal of East Asian Studies*, 16(1): 43–60.

Montiel Olea, José Luis, and Carolin Pflueger. 2013. "A robust test for weak instruments." *Journal of Business & Economic Statistics*, 31(3): 358–369.

Mullahy, John, and Edward C Norton. 2022. "Why transform Y? A critical assessment of dependent-variable transformations in regression models for skewed and sometimes-zero outcomes." National Bureau of Economic Research.

Olken, Benjamin A. 2007. "Monitoring corruption: evidence from a field experiment in Indonesia." *Journal of Political Economy*, 115(2): 200–249.

Petersen, Mitchell A, and Raghuram G Rajan. 2002. "Does distance still matter? The information revolution in small business lending." *The journal of Finance*, 57(6): 2533–2570.

Roodman, David, Morten Ørregaard Nielsen, James G MacKinnon, and Matthew D Webb. 2019. "Fast and wild: Bootstrap inference in Stata using boottest." *The Stata Journal*, 19(1): 4–60.

Roth, Jonathan. 2022. "Pretest with caution: Event-study estimates after testing for parallel trends." *American Economic Review: Insights*, 4(3): 305–322.

Ru, Hong. 2018. "Government credit, a double-edged sword: Evidence from the China Development Bank." *The Journal of Finance*, 73(1): 275–316.

Shih, Victor, Christopher Adolph, and Mingxing Liu. 2012. "Getting ahead in the communist party: explaining the advancement of central committee members in China." *American political science review*, 106(1): 166–187.

Shipan, Charles R, and Craig Volden. 2008. "The mechanisms of policy diffusion." *American journal of political science*, 52(4): 840–857.

Shleifer, Andrei, and Robert W Vishny. 1993. "Corruption." *The quarterly journal of economics*, 108(3): 599–617.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing like china." *American economic review*, 101(1): 196–233.

Stein, Jeremy C. 1989. "Efficient capital markets, inefficient firms: A model of myopic corporate behavior." *The quarterly journal of economics*, 104(4): 655–669.

Stein, Jeremy C. 2002. "Information production and capital allocation: Decentralized versus hierarchical firms." *The journal of finance*, 57(5): 1891–1921.

Vannutelli, Silvia. 2022. "From Lapdogs to Watchdogs: Random Auditor Assignment and Municipal Fiscal Performance." National Bureau of Economic Research.

Wallace, Jeremy L. 2016. "Juking the stats? Authoritarian information problems in China." *British Journal of Political Science*, 46(1): 11–29.

Wang, Erik H. 2021. "Frightened mandarins: the adverse effects of fighting corruption on local bureaucracy." *Forthcoming, Comparative Political Studies*.

Wang, Shaoda, and David Y Yang. 2021. "Policy experimentation in China: The political economy of policy learning." National Bureau of Economic Research.

Wang, Yuhua, and Bruce J Dickson. 2022. "How corruption investigations undermine regime support: Evidence from China." *Political Science Research and Methods*, 10(1): 33–48.

Wang, Zhi, Qinghua Zhang, and Li-An Zhou. 2020. "Career incentives of city leaders and urban spatial expansion in China." *Review of Economics and Statistics*, 102(5): 897–911.

Wolfinger, Raymond E, and Steven J Rosenstone. 1980. *Who votes?* Yale University Press.

WorldBank. 2008. *Doing business 2009.* The World Bank.

Xiong, Wei. 2018. "The mandarin model of growth." National Bureau of Economic Research.

Xu, Chenggang. 2011. "The fundamental institutions of China's reforms and development." *Journal of Economic Literature*, 49(4): 1076–1151.

Yao, Yang, and Muyang Zhang. 2015. "Subnational leaders and economic growth: evidence from Chinese cities." *Journal of Economic Growth*, 20(4): 405–436.

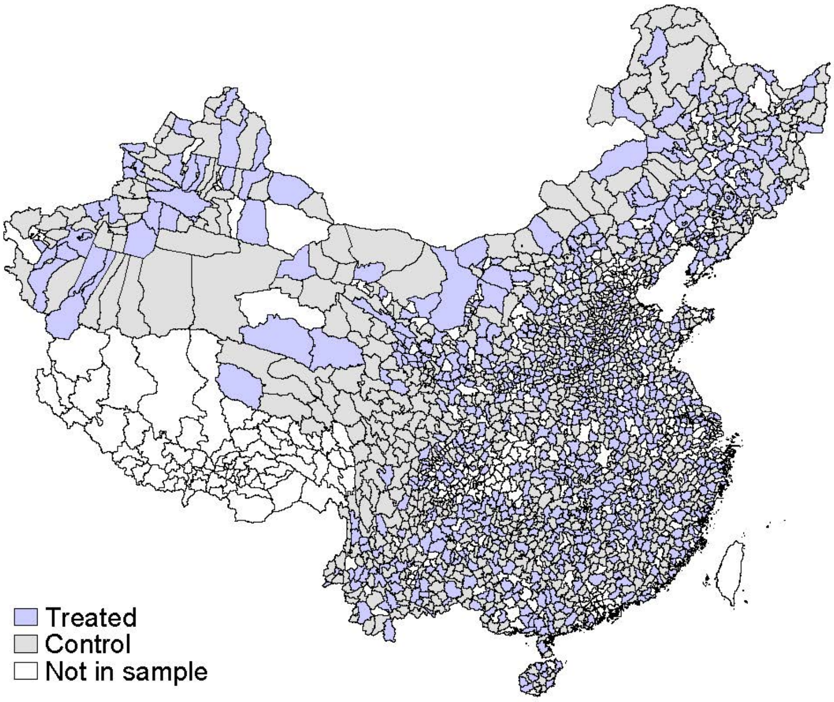
Young, Alwyn. 2019. "Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results." *The quarterly journal of economics*, 134(2): 557–598.

Zhu, Boliang. 2017. "MNCs, rents, and corruption: Evidence from China." *American Journal of Political Science*, 61(1): 84–99.

Zilibotti, Fabrizio. 2017. "Growing and slowing down like China." *Journal of the European Economic Association*, 15(5): 943–988.

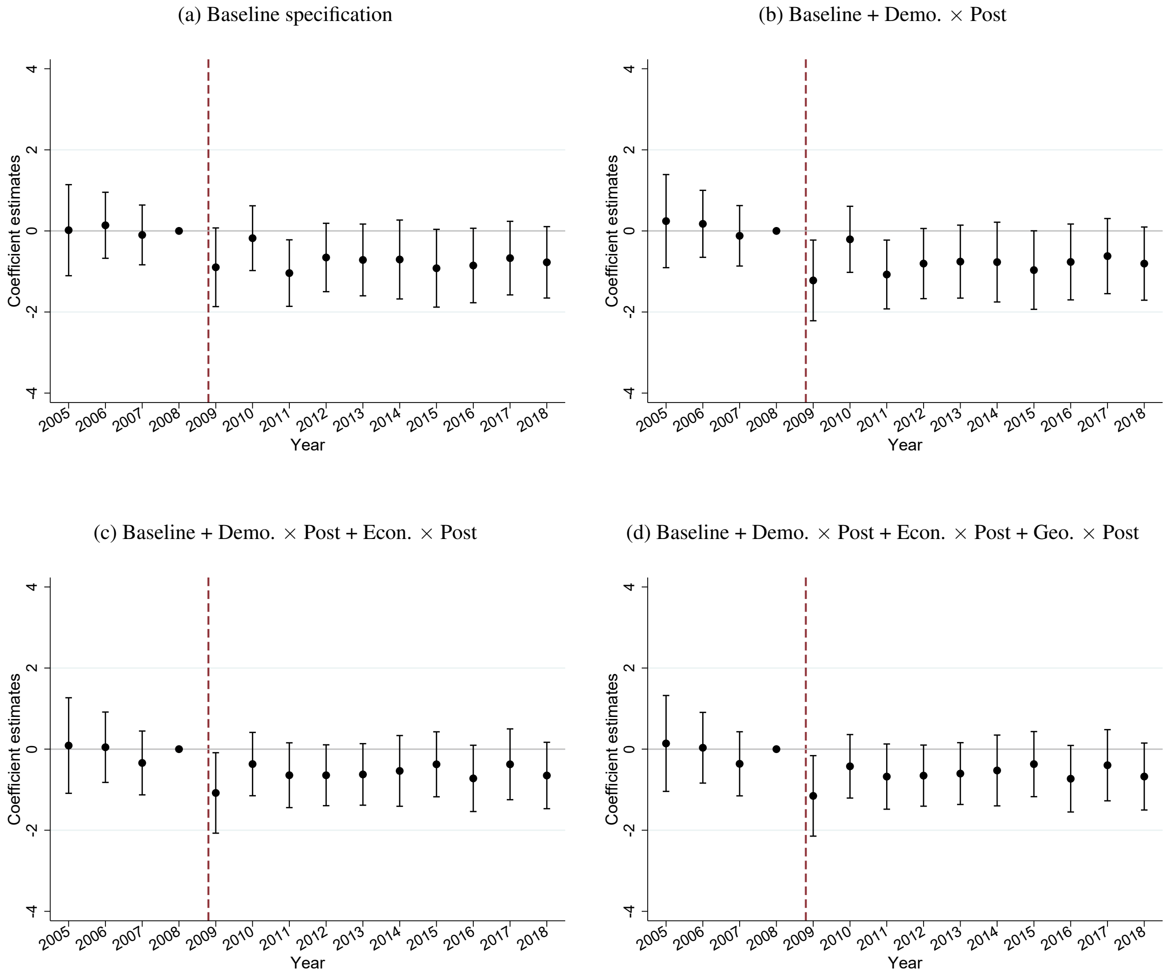
Figures and Tables

Figure 1: Distribution of treatment and control counties



Notes: This figure shows the distribution of the treatment and control counties, where treatment counties are those with the surveys teams deployed in 2005 and control counties are those without. Counties not in sample include: (1) counties in the four centrally-managed cities (*Beijing*, *Shanghai*, *Tianjin*, and *Chongqing*); (2) urban districts, which are more developed economically but less independent administratively; (3) counties in Tibet where data is unavailable; (4) counties outside mainland China.

Figure 2: Dynamic effect on GDP growth manipulation



(a) Baseline specification

(b) Baseline + Demo. × Post

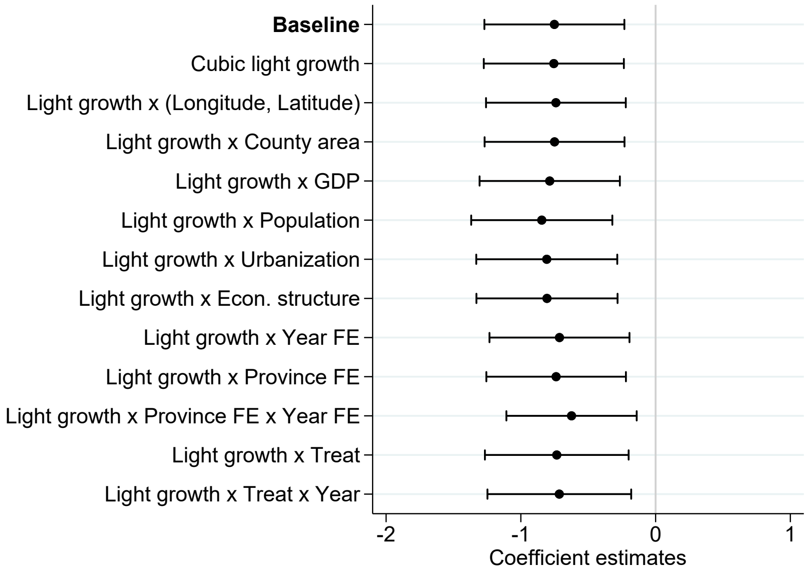
(c) Baseline + Demo. × Post + Econ. × Post

(d) Baseline + Demo. × Post + Econ. × Post + Geo. × Post

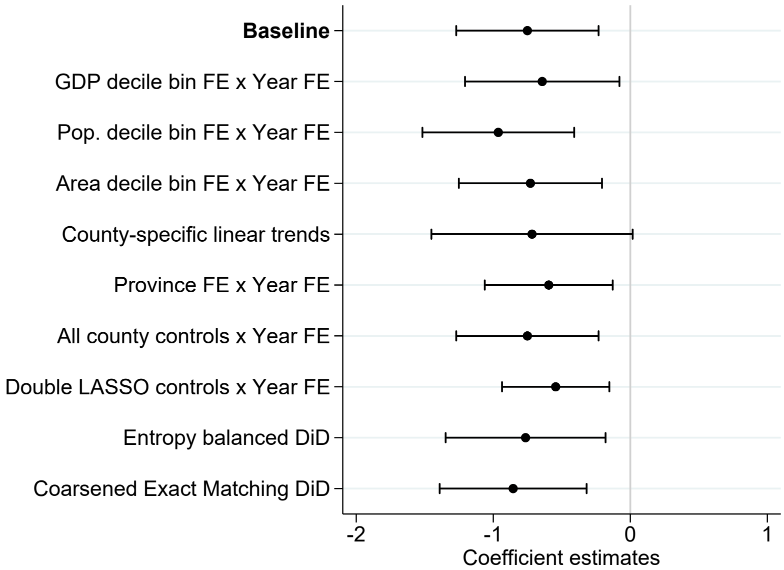
Notes: This figure shows the dynamic effect of the reform on GDP growth manipulation using the baseline event study specification (equation 6) with baseline county covariates included gradually. The year 2008, one year before the reform in 2009, is omitted as the reference year. Standard errors used to construct the 90% confidence intervals, which are denoted by the spikes, are clustered at the county level.

Figure 3: Sensitivity to alternative specifications

(a) Flexible effects of light growth



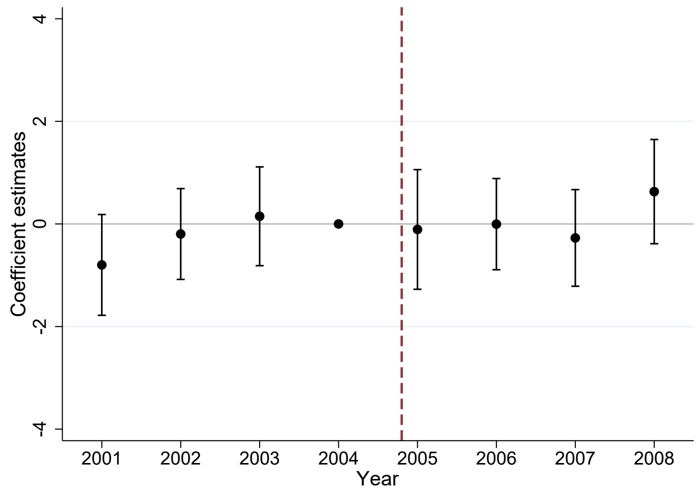
(b) Addressing covariate imbalance



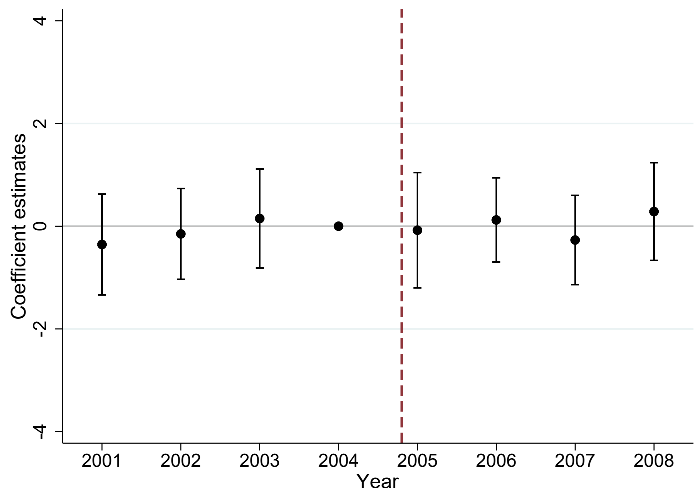
Notes: This figure checks the sensitivity of the baseline results to alternative specifications, and is created by estimating variants of the baseline equation (5). Specifically, panel (a) addresses the concern that the mapping between light growth and economic growth may not be uniform across counties or years, by allowing the mapping to vary flexibly; panel (b) addresses the concern that the results may be confounded by covariate imbalance between treatment counties and control counties, by directly controlling for the sources of imbalance or achieving covariate balance through entropy balancing and coarsened exact matching, among others. Standard errors used to construct the 90% confidence intervals, which are denoted by the spikes, are clustered at the county level.

Figure 4: Using the launch of the survey teams in 2005 as a placebo

(a) Baseline specification



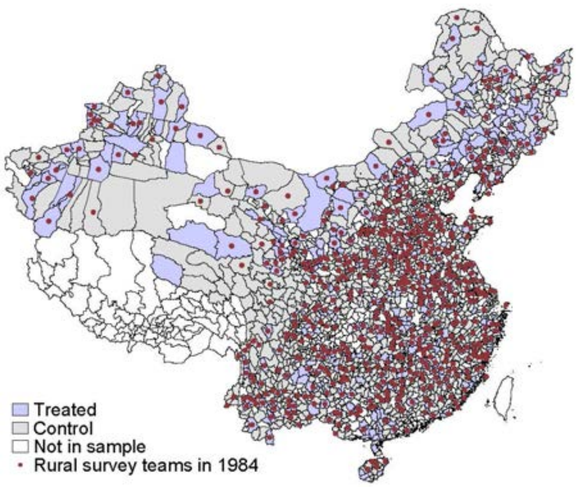
(b) Baseline + County covariates $\times$ Post



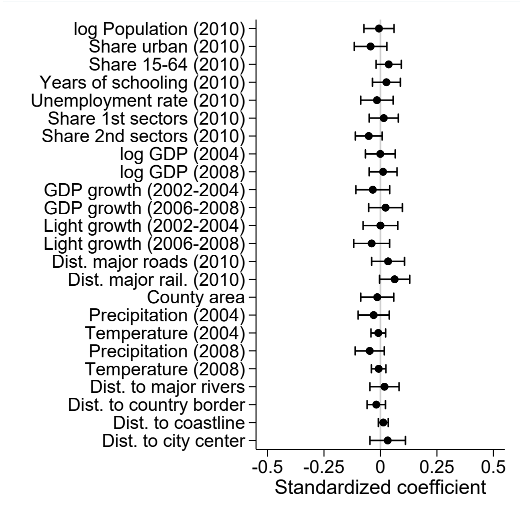
Notes: This figure conducts a placebo event study around 2005, when the survey teams were deployed but had not started disciplining local officials. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure 5: The randomly assigned rural survey teams in 1984 as IV

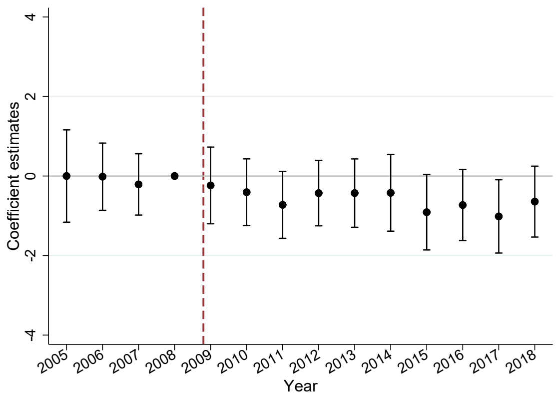
(a) Distribution of the rural survey teams in 1984



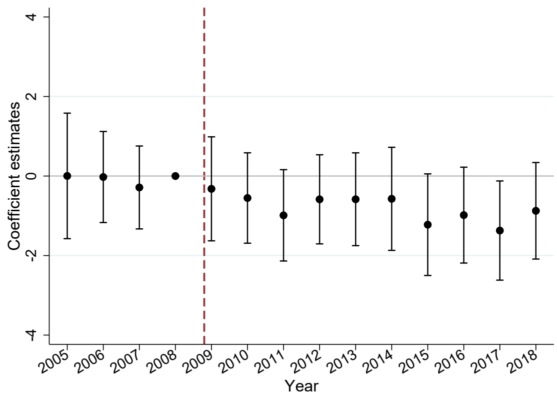
(b) Testing for random assignment of IV



(c) Reduced-form estimates of dynamic effect



(d) Second-stage estimates of dynamic effect

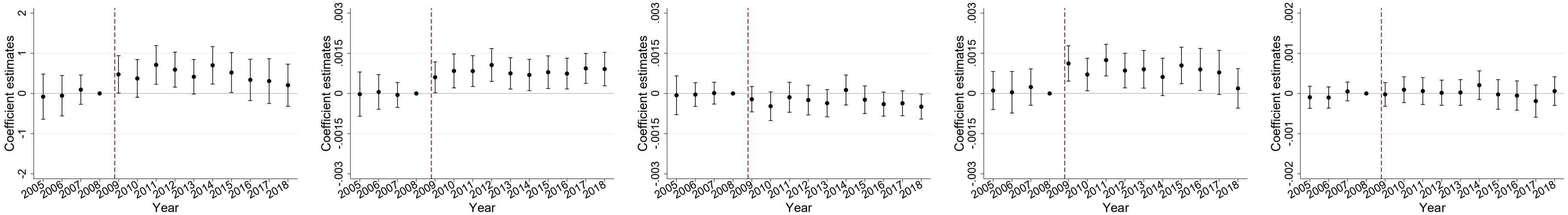


Notes: Panel (a) shows the distribution of the randomly assigned rural survey teams in 1984, which serves as an instrument for treatment counties. Panel (b) tests the randomness of the instrument and plots the standardized coefficients from a univariate regression of baseline county covariates on the instrument, with province fixed effects included as the randomization is stratified by province. Panel (c) shows the reduced-form event study estimated by replacing the treatment in the baseline event study specification (equation 6) with the instrument. Panel (d) shows the second-stage event study estimated by replacing the post-reform dummy in the baseline IV specification (equation 11) with a set of year dummies. The year 2008, which is one year before the reform in 2009, is omitted as the reference year. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure 6: Dynamic effect on government policies

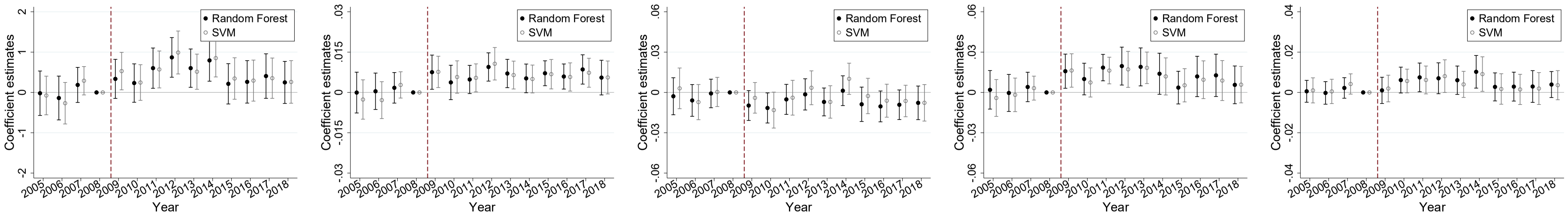
Key words frequency

(a) Standardized index (b) Business attraction (c) Infrastructure investment (d) Market reform (e) Policy experimentation



Machine learning-based topic score

(f) Standardized index (g) Business attraction (h) Infrastructure investment (i) Market reform (j) Policy experimentation

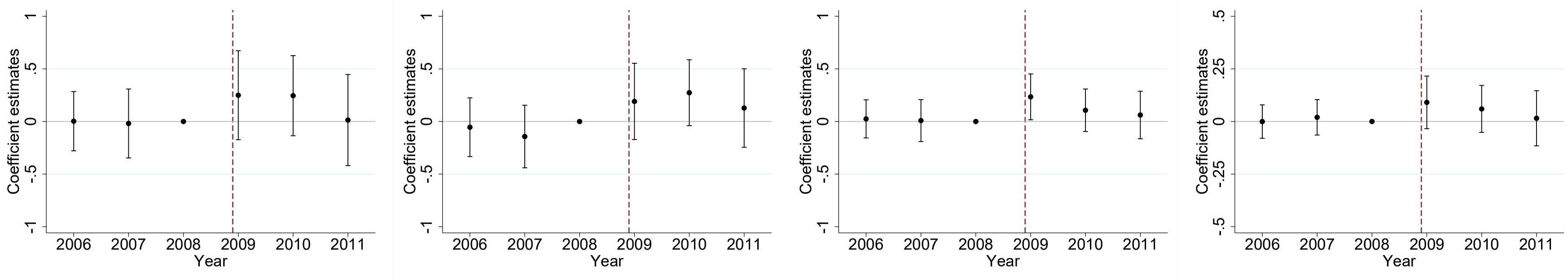


Notes: This figure shows the dynamic effect of the reform on government policies across four policy areas. The unit of observation is county. The sample period is 2005-2018. The sample includes 97 counties. The estimation equations are event study variants of the specifications in Table 5. The year 2008, one year before the reform in 2009, is omitted as the reference year. Panels (a)-(e) measure policies using a simple key words frequency method. Panels (f)-(j) measure policies using supervised machine learning methods (Random Forest and Support Vector Machine). The detailed procedures for constructing these measures are described in Appendix C. To alleviate multiple hypothesis testing issues, panels (a) and (f) report estimates using an standardized index by summarizing the four policy measures following Kling, Liebman and Katz (2007). Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure 7: Dynamic effect on bank credit

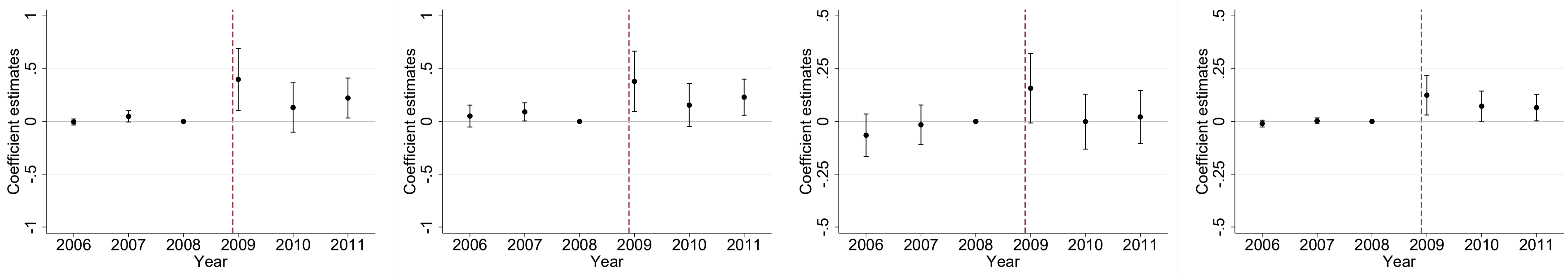
Difference-in-differences estimates

(a) IHS(Loans) (b) IHS(Loans to small firms) (c) IHS(# Firms granted bank loans) (d) IHS(# Branches with loan services)



Difference-in-difference-in-differences estimates

(e) IHS(Loans) (f) IHS(Loans to small firms) (g) IHS(# Firms granted bank loans) (h) IHS(# Branches with loan services)

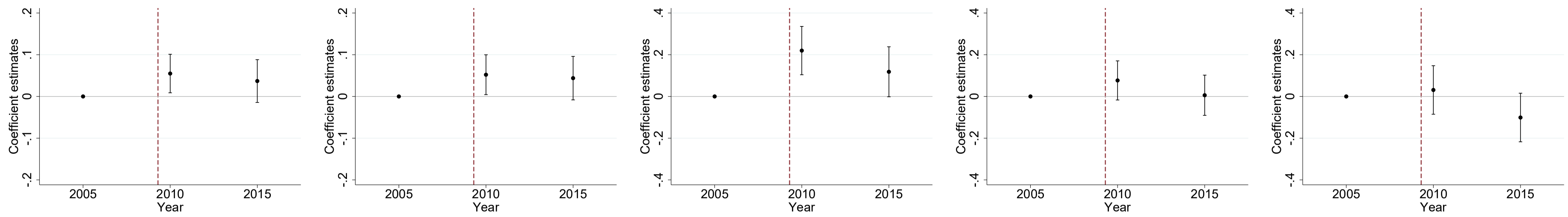


Notes: This figure shows the dynamic effect of the reform on bank credit. The unit of observation is county. The sample period is 2006-2011. The estimation equations are event study variants of the specifications in Table 6. The dependent variables are transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails of the skewed outcome distributions (Bellemare and Wichman, 2020). Panels (a)-(d) conduct a conventional difference-in-differences estimation. Panels (e)-(h) further conduct a difference-in-difference-in-differences estimation utilizing differential control of banks by local governments across counties. Government control over banks is measured as the standardized share of pre-reform loans from City Commercial Banks (CCBs) in a county, whose controlling shareholders are local governments. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure 8: Effect on firm entry

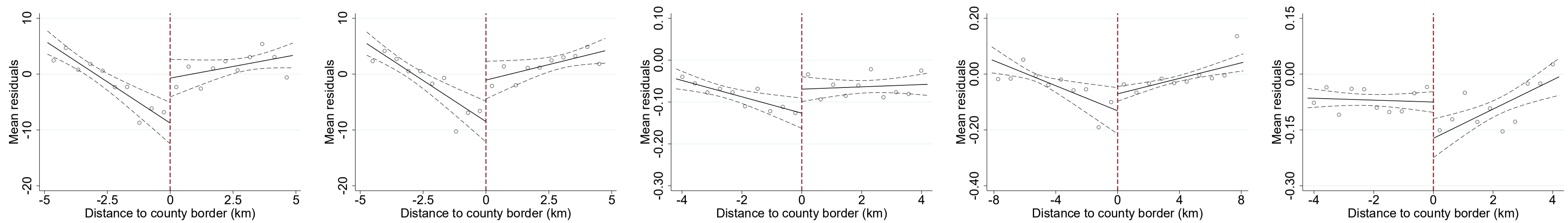
County-level DID estimates

(a) IHS(#All firms) (b) IHS(#Private firms) (c) IHS(#SOEs) (d) IHS(#Foreign firms) (e) IHS(#Collective firms)



Town-level RD estimates

(f) #All firms (g) #Private firms (h) #SOEs (i) #Foreign firms (j) #Collective firms



Notes: Panels (a)-(e) show the dynamic effect of the reform on firm entry at the county level using event study variants of the specifications in Table 7. The sample includes the years 2005, 2010, and 2015. The dependent variables are transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails of the skewed outcome distributions (Bellemare and Wichman, 2020). Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level. Panels (f)-(j) show the jumps in firm entry at the town level across county borders with different treatment statuses on each side. The sample includes only the post-reform years 2010 and 2015. Due to the small number of firm entry at the town level, here the IHS transformations are not adopted (Bellemare and Wichman, 2020). The circles denote the mean of town-level firm entry within a distance bin, after partialling out county border fixed effects. The solid lines are fitted separately for each side of the county border. The dashes are the associated 90% confidence intervals constructed using standard errors clustered at the county border level. The bandwidths are the optimal bandwidths proposed by Calonico, Cattaneo and Titiunik (2014).

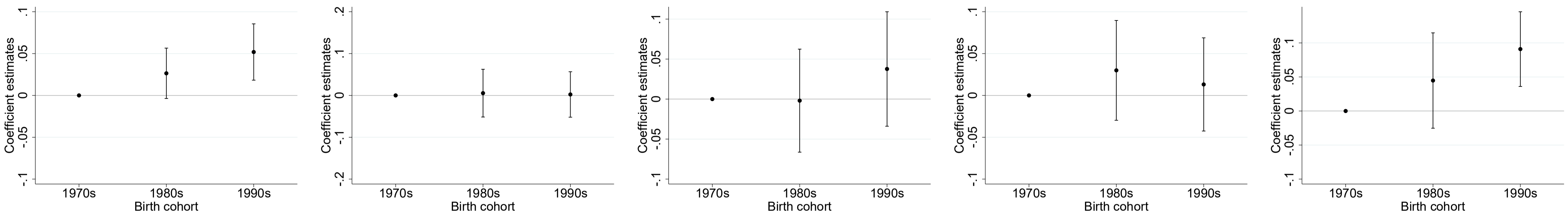
Figure 9: Effect on citizen attitudes towards local governments

Trust in local officials

(a) Trust in local officials (b) Placebo: trust in most people (c) Trust in local officials:2012 (d) Trust in local officials:2014 (e) Trust in local officials:2016

Evaluation of government performance

(f) Eval. of govt. performance (g) Placebo: Eval. of others' health (h) Eval. of govt. performance:2012 (i) Eval. of govt. performance:2014 (j) Eval. of govt. performance:2016



Notes: This figure shows the effect of the reform on citizens' attitudes towards local governments, and is created by visualizing the results in Table 8. The unit of observation is citizen (three waves of survey data from the China Family Panel Studies, 2012, 2014 and 2016). The estimation equation is a cohort difference-in-differences specification (equation 14) that utilizes two sources of variation: (1) treatment counties versus control counties; (2) most affected cohorts versus less unaffected cohorts within the same county. The latter source of variation is built on insights in the psychology and political science literature that citizens' political attitudes are most permeable during teenage years and keep stable since one's 30s (Wolfinger and Rosenstone, 1980; Krosnick and Alwin, 1989). As such, the 1990s cohort are defined as the most affected group as they would be younger than 30 during the reform period, and thus, their political attitudes would be most permeable. The 1980s cohort are defined as the less affected group because some of them would be older than 30 during the reform period. The omitted group is those born in the 1970s, as they would be older than 30 in the reform period and thus be unaffected by the reform. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Table 1: Balance test

	Treat		Control		Difference		
	Mean	SD	Mean	SD	T - C	SE	*p*-value
			Panel A: Demography				
log Population (2010)	13.00	0.70	12.50	0.81	0.50	0.04	0.00
Share urban (%, 2010)	33.65	12.30	34.64	14.31	-0.99	0.66	0.13
Share 15-64 (%, 2010)	72.53	4.56	72.44	4.53	0.10	0.22	0.67
Years of schooling (2010)	8.23	0.73	8.19	1.05	0.04	0.05	0.34
			Panel B: Economic development				
Unemployment rate (%, 2010)	2.08	1.39	2.16	1.52	-0.09	0.07	0.23
Share primary sectors (%, 2010)	64.97	18.20	64.30	17.93	0.68	0.88	0.44
Share secondary sectors (%, 2010)	16.10	12.79	15.49	11.54	0.61	0.59	0.30
log GDP (2004)	12.47	0.94	11.97	1.03	0.50	0.05	0.00
log GDP (2008)	12.96	0.97	12.48	1.06	0.48	0.05	0.00
GDP growth (%, 2002-2004 average)	11.28	6.26	11.57	6.53	-0.29	0.33	0.38
GDP growth (%, 2006-2008 average)	12.86	6.38	12.67	6.63	0.19	0.32	0.54
Light growth (%, 2002-2004 average)	18.24	12.27	18.82	16.28	-0.58	0.72	0.42
Light growth (%, 2006-2008 average)	6.86	10.31	7.67	11.93	-0.81	0.55	0.14
Distance to major roads (km, 2010)	69.18	96.10	75.21	80.62	-6.03	4.24	0.16
Distance to major railways (km, 2010)	70.16	102.37	74.22	90.45	-4.06	4.63	0.38
			Panel C: Geography				
County area (km^2)	3900	7453	4128	10222	-228	446	0.61
Precipitation (inches, 2004)	0.04	0.09	0.03	0.10	0.00	0.00	0.55
Temperature (degrees, 2004)	13.91	5.23	13.57	5.41	0.34	0.26	0.19
Precipitation (inches, 2008)	0.05	0.11	0.05	0.11	0.01	0.01	0.20
Temperature (degrees, 2008)	13.73	5.08	13.43	5.25	0.30	0.25	0.23
Distance to major rivers (km)	59.17	61.15	57.76	59.03	1.41	2.90	0.63
Distance to country border (km)	346.52	251.00	345.25	251.49	1.26	12.16	0.92
Distance to coastline (km)	616.84	612.04	640.02	568.79	-23.18	28.46	0.42
Distance to prefecture center (km)	60.14	41.78	62.66	46.66	-2.52	2.17	0.25

Notes: This table provides balance tests by comparing counties with the survey teams deployed in 2005 to those without. The *p*-values reported in the last column are from *t*-tests of mean equality between groups. Except for GDP, all variables in this table are from sources that the county has no control on. The reason for using the year 2010 for the demographic data, instead of years before the reform in 2009, is that the population census was only conducted in 2000 and 2010. Similarly, the transportation data is also in 2010 due to data limitation

Table 2: Effect on GDP growth manipulation

	(1)	(2)	(3)	(4)
Dep. var.:	Reported GDP growth (%)			
Treat × Post	-0.751**	-0.869***	-0.552***	-0.576***
	(0.316)	(0.331)	(0.162)	(0.161)
Light growth (%)	0.023***	0.022***	0.017***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)
County FE	X	X	X	X
Year FE	X	X	X	X
Demographic controls × Post		X	X	X
Economic controls × Post			X	X
Geographic controls × Post				X
Cluster level	County	County	County	County
Observations	23,360	22,580	20,343	20,273
R-squared	0.269	0.269	0.362	0.362
Mean dep. var.	10.97	10.97	10.84	10.84

Notes: This table shows the effect of the reform on GDP growth manipulation. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Estimating spillover effect

	(1)	(2)	(3)	(4)
Dep. var.:	Reported GDP growth (%)			
Treat × Post	-0.576***	-0.580***	-0.565***	-0.585***
	(0.161)	(0.164)	(0.164)	(0.162)
# Treat neighbors × Post		-0.010		
		(0.071)		
1(# Treat neighbors>0) × Post			0.108	
			(0.263)	
1(# Treat neighbors>Median=2) × Post				-0.133
				(0.178)
Light growth (%)	0.017***	0.017***	0.017***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)
County FE	X	X	X	X
Year FE	X	X	X	X
County controls × Post	X	X	X	X
Neighbor number FE × Post		X	X	X
Cluster level	County	County	County	County
Mean dep. var.	10.84	10.84	10.84	10.84
Mean number of neighbors	5.91	5.91	5.91	5.91
Mean number of treated neighbors	1.97	1.97	1.97	1.97

Notes: The tables shows the spillover effect of the reform on GDP growth manipulation. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. # Treat neighbors denote the number of treatment counties among a county's neighbors, where neighbors are defined as counties sharing a common boundary segment with a county. Standard errors clustered at the county level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Effect on GDP growth manipulation - IV estimates

	(1)	(2)	(3)	(4)
Panel A: First-stage estimates				
Dep. var.:	Treat × Post			
Treat^{1984} × Post	0.739***	0.724***	0.727***	0.729***
	(0.017)	(0.015)	(0.016)	(0.016)
Light growth (%)	-0.000***	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Effective F-statistic	1,945	2,210	2,044	2,080
Critical value for 5% worst case bias	37.4	37.4	37.4	37.4
Panel B: Second-stage estimates				
Dep. var.:	Reported GDP growth (%)			
Treat × Post	-0.711*	-0.839**	-0.589***	-0.590***
	(0.400)	(0.402)	(0.205)	(0.203)
Light growth (%)	0.021***	0.021***	0.016***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)
Mean dep. var.	10.98	10.97	10.84	10.84
County FE	X	X	X	X
Year FE	X	X	X	X
Province FE × Post	X	X	X	X
Demographic controls × Post		X	X	X
Economic controls × Post			X	X
Geographic controls × Post				X
Cluster level	County	County	County	County
Observations	22,998	22,580	20,343	20,273

Notes: This table shows the IV estimates on the effect of the reform on GDP growth manipulation. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with a survey team deployed in 2005. Treat^{1984} is a dummy variable indicating counties with a randomly assigned rural survey team in 1984. Post is a dummy variable indicating years after the reform in 2009. Panel A provides the first-stage estimates. Panel B provides the second-stage estimates with Treat^{1984} serving as an instrument for Treat. The effective F-statistics and corresponding critical values are constructed following Montiel Olea and Pflueger (2013). Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Effect on government policies

Dep. var.:	(1) Standardized index	(2) Business attraction	(3) Infrastructure investment	(4) Market reform	(5) Policy experimentation
Panel A: Key words frequency					
Treat × Post	0.474***	0.001***	-0.000	0.001***	0.000
	(0.167)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.482	0.513	0.481	0.606	0.615
Mean dep. var.	0	0.002	0.002	0.004	0.001
Panel B: Topic score predicted by Random Forest					
Treat × Post	0.452***	0.006***	-0.005	0.012**	0.004*
	(0.155)	(0.002)	(0.004)	(0.005)	(0.002)
R-squared	0.427	0.435	0.361	0.591	0.512
Mean dep. var.	0	0.013	0.040	0.068	0.016
Panel C: Topic score predicted by Support Vector Machine					
Treat × Post	0.508***	0.007***	-0.003	0.012***	0.003
	(0.163)	(0.002)	(0.005)	(0.004)	(0.002)
R-squared	0.474	0.427	0.435	0.613	0.587
Mean dep. var.	0	0.014	0.051	0.068	0.019
County FE	X	X	X	X	X
Year FE	X	X	X	X	X
County controls × Post	X	X	X	X	X
Cluster level	County	County	County	County	County
Observations	883	883	883	883	883

Notes: This table shows the effect of the reform on government policies across four policy areas. The unit of observation is county. The sample period is 2005-2018. The sample includes 97 counties. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Panel A measures policies using a simple key words frequency method. Panel B and C measure policies using supervised machine learning methods (Random Forest and Support Vector Machine). The detailed procedures for constructing these measures are described in Appendix C. To alleviate multiple hypothesis testing issues, column (1) reports estimates using an standardized index by summarizing the four policy measures following Kling, Liebman and Katz (2007). Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Effect on bank credit

Dep. var. is IHS of:	(1) Total amount of loans	(2) Loans to small firms	(3) # Firms granted loans	(4) # Branches granting loans
Panel A: Difference-in-differences				
Treat × Post	0.174	0.263**	0.123	0.049
	(0.123)	(0.116)	(0.076)	(0.036)
R-squared	0.329	0.424	0.520	0.482
Panel B: Difference-in-difference-in-differences				
Treat × Government control × Post	0.237**	0.208**	0.086	0.090***
	(0.103)	(0.097)	(0.062)	(0.033)
Treat × Post	0.027	0.062	0.024	-0.009
	(0.124)	(0.117)	(0.073)	(0.036)
Government control × Post	0.026	0.026	0.062	0.017
	(0.089)	(0.081)	(0.050)	(0.029)
R-squared	0.325	0.424	0.551	0.514
County controls × Post	X	X	X	X
County FE	X	X	X	X
Year FE	X	X	X	X
Observations	8,922	8,922	8,922	8,922

Notes: This table shows the effect of the reform on bank credit. The unit of observation is county. The sample period is 2006-2011. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Government control denotes the standardized share of pre-reform loans from City Commercial Banks (CCBs) in a county, whose controlling shareholders are local governments. The dependent variables are transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails of the skewed outcome distributions (Bellemare and Wichman, 2020). Panel A conducts a difference-in-differences estimation as usual. Panel B adopts a difference-in-difference-in-differences estimation utilizing differential control of banks by local governments across counties. The estimation equation is equation (12). Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Effect on firm entry

	(1)	(2)	(3)	(4)	(5)
Firm type:	All	Private	SOEs	Foreign	Collective
Panel A: Difference-in-differences estimates					
Dep. var.:	IHS(# Firm registrations)				
Treat × Post	0.046*	0.048*	0.169***	0.041	-0.035
	(0.026)	(0.027)	(0.062)	(0.051)	(0.062)
County FE	X	X	X	X	X
Year FE	X	X	X	X	X
County controls × Post	X	X	X	X	X
Cluster level	County	County	County	County	County
Observations	4,494	4,494	4,494	4,494	4,494
R-squared	0.943	0.944	0.535	0.726	0.571
Panel B: Regression discontinuity estimates at the town level					
Dep. var.:	# Firm registrations in post-reform period				
Treat	8.594**	8.241**	0.035	0.016	-0.084**
	(3.378)	(3.369)	(0.027)	(0.025)	(0.037)
County border FE	X	X	X	X	X
Cluster level	County border	County border	County border	County border	County border
RD polynomial	Linear	Linear	Linear	Linear	Linear
Observations	10,776	10,492	9,266	15,356	9,304
R-squared	0.334	0.335	0.229	0.372	0.249
Mean dep. var.	18.50	18.35	0.10	0.09	0.06
Bandwidth (in km)	4.88	4.74	4.20	8.14	4.21

Notes: This table shows the effect of the reform on firm entry. The unit of observation is county in panel A and town in panel B. The sample includes the years 2005, 2010, and 2015 in panel A and the years 2010 and 2015 (post-reform years) in panel B. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Panel A conducts a difference-in-differences estimation as usual. Panel B conducts a regression discontinuity estimation at the town level along county borders with different treatment statuses on each side. The estimation equation is a local linear specification (equation 13) using the optimal bandwidth proposed by Calonico, Cattaneo and Titiunik (2014). The dependent variables at the county level in panel A are transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails of the skewed outcome distributions (Bellemare and Wichman, 2020). For the dependent variables at the town level in panel B, such transformations are not adopted due to the small number of firm entry (Bellemare and Wichman, 2020). Standard errors clustered at the indicated (county or county border) level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Effect on citizens attitudes towards local governments

	(1)	(2)	(3)	(4)	(5)
Panel A: Trust in local officials					
Dep. var.:	Trust in local officials: 2012-2016	Placebo: trust in most people: 2012-2016	Trust in local officials: 2012	Trust in local officials: 2014	Trust in local officials: 2016
Treat × 1980s cohort	-0.051	0.010	-0.109	0.060	-0.031
	(0.129)	(0.021)	(0.168)	(0.221)	(0.264)
Treat × 1990s cohort	0.452***	-0.002	0.399	0.560**	0.319*
	(0.154)	(0.027)	(0.299)	(0.268)	(0.188)
Observations	17,163	17,166	6,338	5,651	5,174
R-squared	0.069	0.086	0.096	0.110	0.089
Number of counties	82	82	81	82	82
Mean dep. var.	4.84	0.57	4.82	4.90	4.80
Panel B: Evaluation of local government performance					
Dep. var.:	Eval. of govt. performance: 2012-2016	Placebo: eval. of others' health: 2012-2016	Eval. of govt. performance: 2012	Eval. of govt. performance: 2014	Eval. of govt. performance: 2016
Treat × 1980s cohort	0.026	0.004	-0.002	0.030	0.045
	(0.018)	(0.034)	(0.039)	(0.036)	(0.042)
Treat × 1990s cohort	0.052**	0.002	0.038	0.013	0.091***
	(0.020)	(0.033)	(0.043)	(0.033)	(0.033)
Observations	17,194	11,999	6,353	5,659	5,182
R-squared	0.060	0.108	0.097	0.096	0.081
Number of counties	82	82	81	82	82
Mean dep. var.	0.82	0.25	0.53	0.81	0.85
County FE	X	X	X	X	X
Cohort FE	X	X	X	X	X
County controls × Cohort FE	X	X	X	X	X
Survey wave FE	X	X	X	X	X
Citizen controls	X	X	X	X	X
Cluster level	County	County	County	County	County

Notes: This table shows the effect of the reform on citizens' attitudes towards local governments. The unit of observation is citizen (three waves of survey data from the China Family Panel Studies, 2012, 2014 and 2016). Treat is a dummy variable indicating counties with the survey teams deployed in 2005. 1980s cohort and 1990s cohort are dummy variables indicating citizens born in the 1980s and the 1990s, respectively. The estimation equation is a cohort difference-in-differences specification (equation 14) that utilizes two sources of variation: (1) treatment counties versus control counties; (2) most affected cohorts versus less unaffected cohorts within the same county. The latter source of variation is built on insights in the psychology and political science literature that citizens' political attitudes are most permeable during teenage years and keep stable since one's 30s (Wolfinger and Rosenstone, 1980; Krosnick and Alwin, 1989). As such, the 1990s cohort are defined as the most affected group as they would be younger than 30 during the reform period, and thus, their political attitudes would be most permeable. The 1980s cohort are defined as the less affected group because some of them would be older than 30 during the reform period. The omitted group is those born in the 1970s, as they would be older than 30 in the reform period and thus be unaffected by the reform. Standard errors clustered at the county level are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Online Appendix

Curbing Bureaucratic Information Manipulation

Yongwei Nian

A Theory appendix

This section shows how to solve for the equilibrium effort m^* (effort in manipulating GDP growth) and e^* (effort in stimulating the economy). To write leader i's maximization problem, note that there are four situations: (1) with probability $(1-p)^2$, both i and $-i$ are not detected for manipulation; (2) with probability $(1-p)p$, i is not detected for manipulation, but $-i$ is; (3) with probability $p(1-p)$, i is detected for manipulation, but $-i$ is not; (4) with probability p^2, both are detected for manipulation. Let q_1, q_2, q_3, q_4 denote i's promotion probability in each case, respectively:

$$\begin{aligned} q_1 &= Pr(G_i + \varepsilon_i > G_{-i} + \varepsilon_{-i}) \\ &= \frac{1}{2} + \phi[h(m_i) + g(e_i) - h(m_{-i}) - g(e_{-i})] \end{aligned}$$

$$\begin{aligned} q_2 &= Pr[G_i + \varepsilon_i > G_{-i} + \varepsilon_{-i} - \delta h(m_{-i})] \\ &= \frac{1}{2} + \phi[h(m_i) + g(e_i) - (1-\delta)h(m_{-i}) - g(e_{-i})] \end{aligned}$$

$$\begin{aligned} q_3 &= Pr[G_i + \varepsilon_i - \delta h(m_i) > G_{-i} + \varepsilon_{-i}] \\ &= \frac{1}{2} + \phi[(1-\delta)h(m_i) + g(e_i) - h(m_{-i}) - g(e_{-i})] \end{aligned}$$

$$\begin{aligned} q_4 &= Pr[G_i + \varepsilon_i - \delta h(m_i) > G_{-i} + \varepsilon_{-i} - \delta h(m_{-i})] \\ &= \frac{1}{2} + \phi[(1-\delta)h(m_i) + g(e_i) - (1-\delta)h(m_{-i}) - g(e_{-i})] \end{aligned}$$

Then one can write i's expected payoff Z as:

$$
\begin{aligned}
Z = & (1-p)^2 q_1 u(R) + (1-p)^2(1-q_1)u(r) \\
& + (1-p)pq_2 u(R) + (1-p)p(1-q_2)u(r) \\
& + p(1-p)q_3[u(R) - \lambda h(m_i)] + p(1-p)(1-q_3)[u(r) - \lambda h(m_i)] \\
& + p^2 q_4[u(R) - \lambda h(m_i)] + p^2(1-q_4)[u(r) - \lambda h(m_i)]
\end{aligned}
$$

Substituting q_1, q_2, q_3, q_4 into Z, i' problem can be write as:

$$
\max_{e_i, m_i, e_i + m_i <= \bar{C}} u(r) + [u(R) - u(r)]\phi g(e_i) + [(u(R) - u(r))\phi(1 - p\delta) - \lambda p]h(m_i)
$$

Assume $(u(R) - u(r))\phi(1 - p\delta) - \lambda p > 0$ and interior solution, one can write the first-order condition as:

$$
[u(R) - u(r)]\phi g'(e_i^*) = [(u(R) - u(r))\phi(1 - p\delta) - \lambda p]h'(m_i^*)
$$

Also note that at equilibrium $m_i^* + e_i^* = \bar{C}$. Let $K(\cdot)$=$[\frac{h'(\cdot)}{g'(c-\cdot)}]^{-1}$, then one could solve for m_i^* and e_i^*:

$$
\begin{aligned}
m^* &= K\left[\frac{V\phi}{V\phi(1 - p\delta) - \lambda p}\right] \\
e^* &= \bar{C} - m^*
\end{aligned}
$$

where $V = u(R) - u(r)$.

B Characterizing compliers in IV estimation

In this section, I follow Marbach and Hangartner (2020) to understand the characteristics of compliers in the presence of heterogeneous treatment effects in the IV estimation. Let $Treat$ denote the realized treatment status of all counties, which is equal to 1 if a county had survey team deployed in 2005, and 0 otherwise. Let $Treat^{1984}$ denote the instrument, which is equal to

1 if a county had a randomly assigned rural survey team in 1984, and 0 otherwise. Let $Treat(0)$ and $Treat(1)$ denote the potential treatment status of a county depending on the instrument. Under monotonicity, which is supported by Appendix Table A3, defier counties can be ruled out. Then under random assignment of the instrument, which is also true as I discuss in Section 4.2.3, the mean of covariate X among never-taker counties could be computed using their observed counterparts:

$$E[X|NeverTakers] = E[X|Treat = 0, Treat^{1984} = 1]$$

The covariate mean for always-taker counties could also be computed using their observed counterparts:

$$E[X|AlwaysTakers] = E[X|Treat = 1, Treat^{1984} = 0]$$

Under random assignment of the instrument, one could also use the realized values of $Treat$ to compute the fraction of never-taker counties and always-taker counties as follows, respectively:

$$Pr[NeverTaker] = Pr[Treat = 0|Treat^{1984} = 1] = 1 - E[Treat|Treat^{1984} = 1]$$

$$Pr[AlwaysTaker] = Pr[Treat = 1|Treat^{1984} = 0] = E[Treat|Treat^{1984} = 0]$$

Under monotonicity, the fraction of complier counties is equal to:

$$Pr[Complier] = 1 - Pr[NeverTaker] - Pr[AlwaysTaker]$$

Then according to the law of total expectation, the mean of covariate X can be written as:

$$\begin{aligned} E[X] = & E[X|Complier]Pr[Complier] \\ & + E[X|NeverTaker]Pr[NeverTaker] \\ & + E[X|AlwaysTaker]Pr[AlwaysTaker] \end{aligned}$$

Combining all these equations, one can back out the mean of covariate X among compliers, which takes the following form and can be empirically estimated:

$$
\begin{aligned}
E[X|Complier] = & \frac{E[X] - E[X|Treat = 0, Treat^{1984} = 1][1 - E[Treat|Treat^{1984} = 1]]}{E[Treat|Treat^{1984} = 1] - E[Treat|Treat^{1984} = 0]} \\
& - \frac{E[X|Treat = 1, Treat^{1984} = 0]E[Treat|Treat^{1984} = 0]}{E[Treat|Treat^{1984} = 1] - E[Treat|Treat^{1984} = 0]}
\end{aligned}
$$

C Textual analysis of government work reports

To construct textual measures of the government policies mentioned in Section 5.1, I first convert all county government work reports into 326,435 sentences, of which 25% are selected at random and manually labeled as belonging to one policy area or not. I then remove unnecessary characters (spaces, numbers, and punctuations) and stopwords from each sentence using a commonly used list of such characters and words.[1] As Chinese words are not space-delimited, I next adopt the open-source Chinese text segmentation library called *Jieba* to segment each sentence into words.[2] For the first textual measure, namely the keywords frequency measure, I simply count in a report the total number of mentions of the keywords corresponding to a certain policy (see Appendix Table A12 for the list of key words) and then divide this count by the total number of words in the report.

Then to construct the machine learning-based textual measures, I convert the tokenized sentences into numerical vectors using the term frequency-inverse document frequency (TF-IDF) method. This method weights each term (word) with its frequency in a document (sentence) and its inverse document (sentence) frequency.[3] So words with higher TF-IDF scores are of higher informativeness. I adopt two commonly used machine learning algorithms: random forest (RF) and support vector machine (SVM) (Gentzkow, Kelly and Taddy, 2019). The RF algorithm creates a forest of decision trees, with each tree trained on a random subset of the vectorized data points. To classify a new sentence, each tree in the forest can make a decision and the decision receiving the most votes is chosen as the final classification. The SVM algo-

[1] The list is from https://github.com/goto456/stopwords/blob/master/README.md

[2] The library can be found at https://github.com/fxsjy/jieba

[3] The IDF of a word in my case is defined as log(# sentences in sample / # sentences containing the word).

rithm operates by finding the optimal hyperplane that separates the two classess of vectorized data points. The optimization is acheived by maximizing the distance between the hyperplane and the cloest data points from either class. A new sentence is then vectorized and assigned to a class based on which side of the hyperplane it falls on. The outputs from these two algorithms are sentence-level binary policy scores and I take the sentence length weighted average of these scores to get a report-level policy score.

Finally, to compute textual similarity between two reports as used in Section 5.5, I use the aforementioned TF-IDF method to vectorize an entire report, and then calculate the pairwise cosine similarity between any two reports i and j following Kelly et al. (2021):

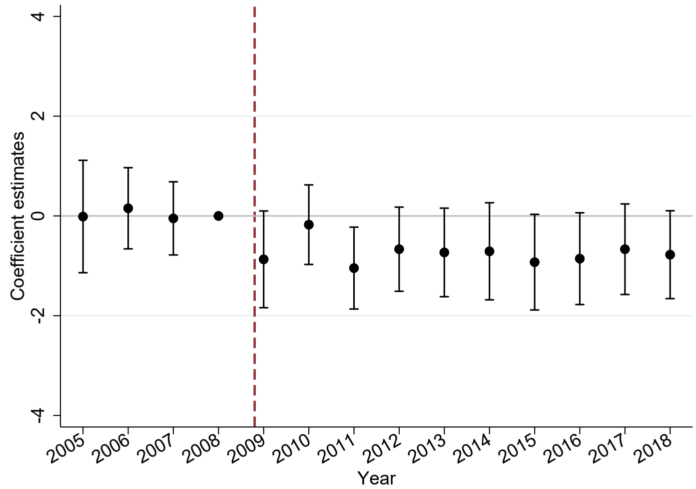
$$Similarity_{ij} = \frac{V_i \cdot V_j}{\| V_i \| \cdot \| V_j \|}$$

where V_i and V_j are vectors representig the two reports, and $V_i \cdot V_j$ is the dot product of these two vectors. $\| V_i \|$ and $\| V_j \|$ are norms of these two vectors. The output $Similarity_{ij}$ is then a measure between -1 and 1, with higher values denoting higher similarity.

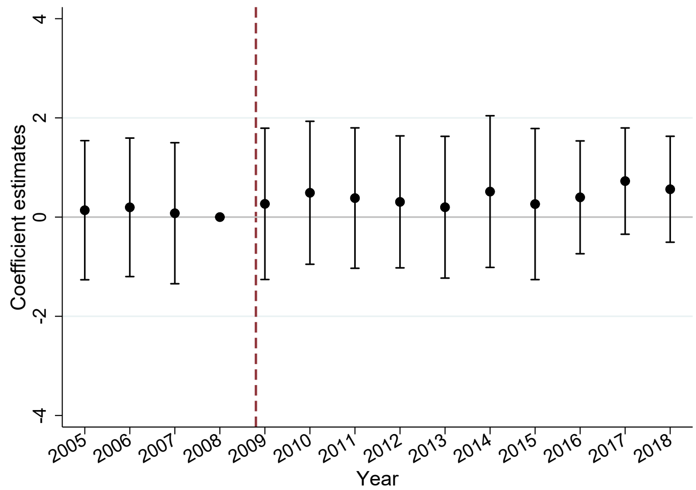
D Additional figures

Figure A1: Decomposing the effect of the reform

(a) Effect on reported GDP growth



(b) Effect on light growth

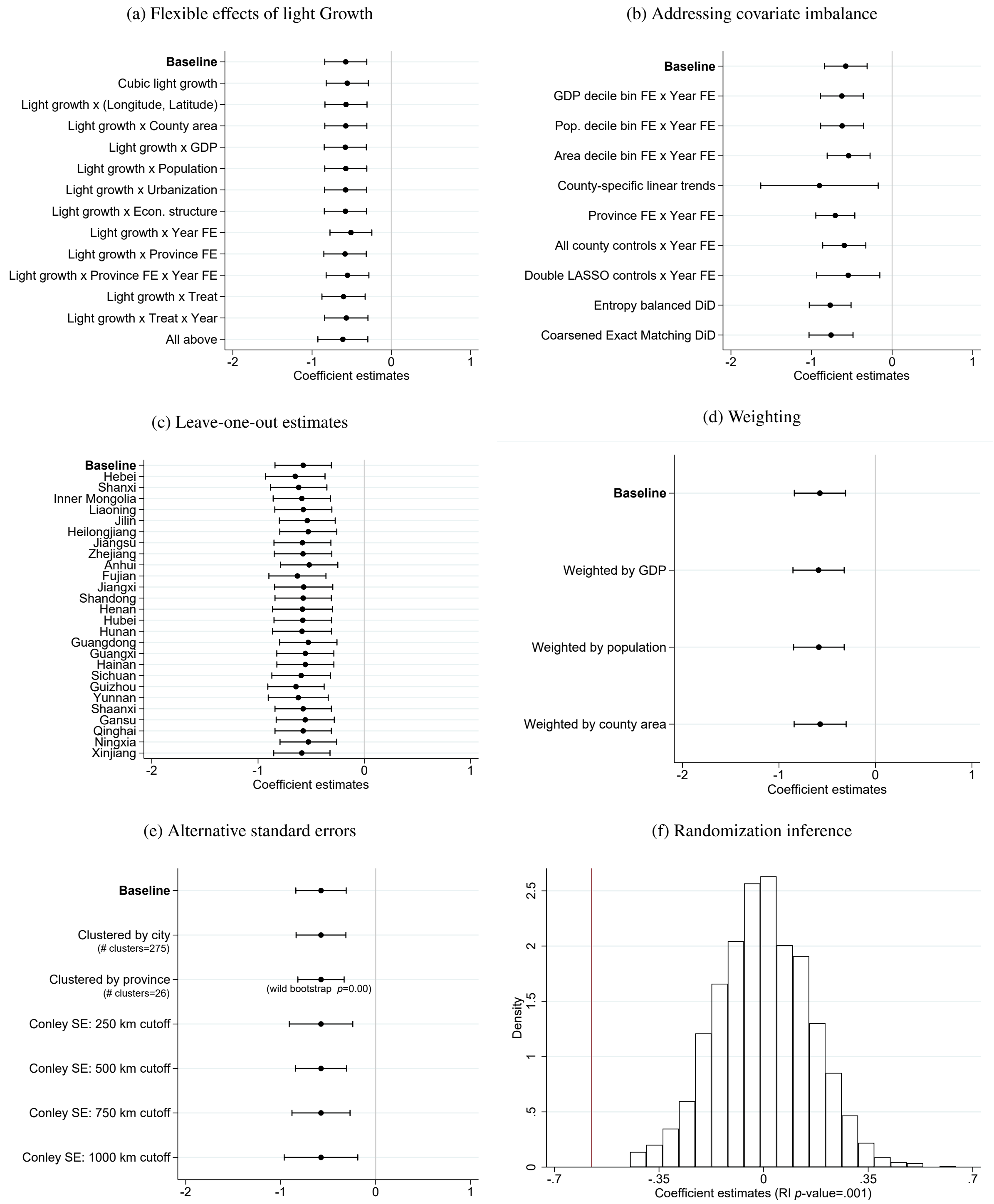


Notes: This figure shows the dynamic effect of the reform on reported GDP growth (panel a) and light growth (panel b) separately. Specifically, I estimate:

$$Y_{ct} = \sum_{j=2005,\ j\neq 2008}^{j=2018} \beta_j Treat_c \times 1_{\{t=j\}} + \delta_c + \lambda_t + \varepsilon_{ct}$$

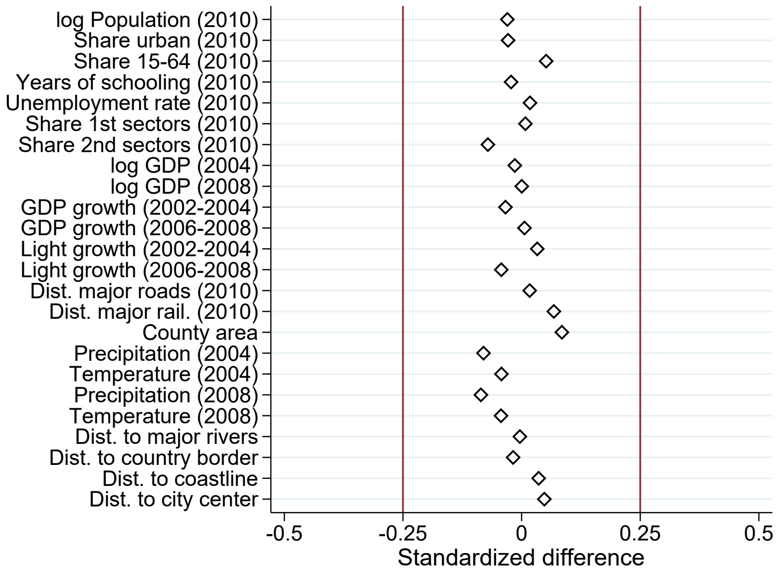
where Y_{ct} denotes either reported GDP growth (panel a) or light growth (panel b). Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure A2: Additional Robustness Checks on Baseline Estimates



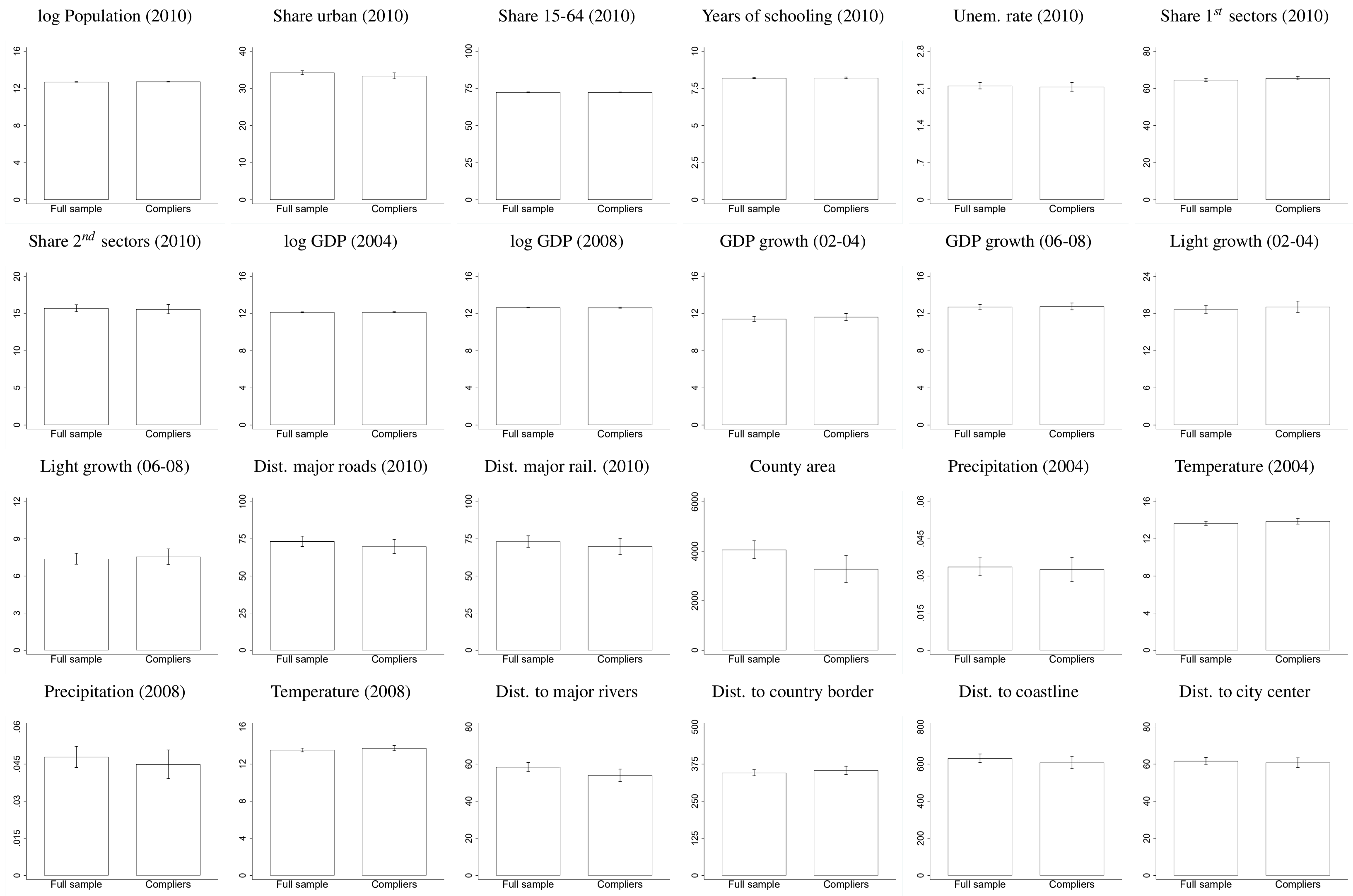
Notes: This figure provides additional robustness checks to the baseline estimates. All panels control for baseline county covariates interacted the post-reform dummy when appropriate. Panel (a) addresses the concern that the mapping between light growth and economic growth may not be uniform across counties or years, by allowing the mapping to vary flexibly. Panel (b) addresses the concern that the results may be confounded by covariate imbalance between treatment counties and control counties, by directly controlling for the sources of imbalance or achieving covariate balance through entropy balance and coarsened exact matching, among others. Panel (c) shows the estimates after excluding one province each time. Panel (d) shows the estimates weighted by baseline county size. Panel (e) shows the estimates with alternative clustering methods and spatial correlation correction. For province-level clustering with a small number of clusters of 26, I also report the wild bootstrap p-values with 2,000 replications (Roodman et al., 2019). For Conley standard errors (Conley, 1999), I account for serial correlation spanning all years and spatial correlation within distances of 250 km, 500 km, 750 km, and 1,000 km. Panel (f) shows the estimates using a randomization inference procedure with 2,000 permutations following Young (2019). The true estimate is denoted by the vertical line and the randomization inference p-value is reported below the figure. Standard errors used to construct the 90% confidence intervals, which are denoted by the spikes, are clustered at the county level when appropriate.

Figure A3: Standardized differences for the instrument



Notes: This graph shows the standardized differences between county groups defined by the instrument ($Treat^{1984} = 1$ vs $Treat^{1984} = 0$). To this end, I first compute the standardized differences between the two groups within each province, which are differences between the sample means normalized by the square root of the average of the sample variances. I then calculate a weighted average using the number of counties within each province as weights. The two vertical lines denote the 25% threshold recommended by Imbens and Rubin (2015) for covariate balance.

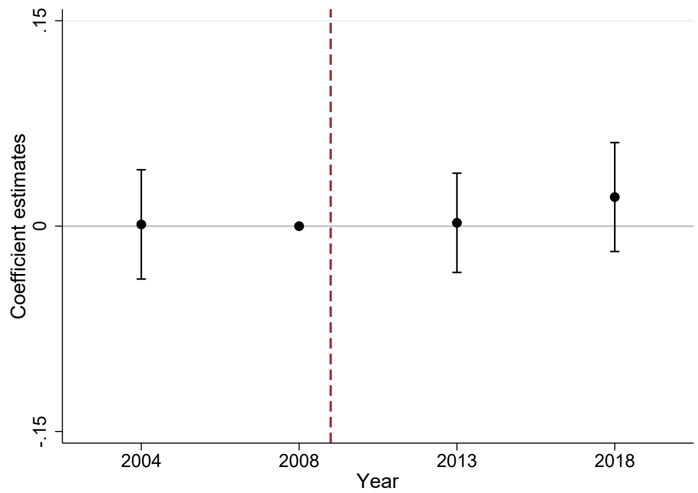
Figure A4: Complier characteristics



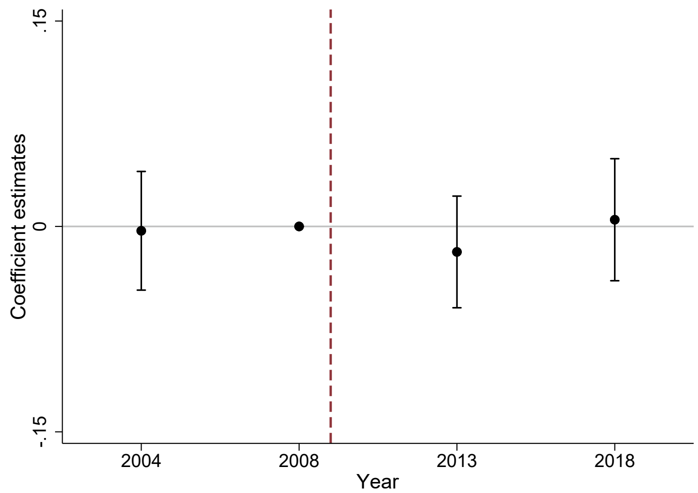
Notes: This figure plots the complier and full sample means, where the complier means are calculated following Marbach and Hangartner (2020) (see also Appendix B)

Figure A5: Dynamic effect on local statistical capacity

(a) Baseline specification



(b) Baseline + County covariates × Post

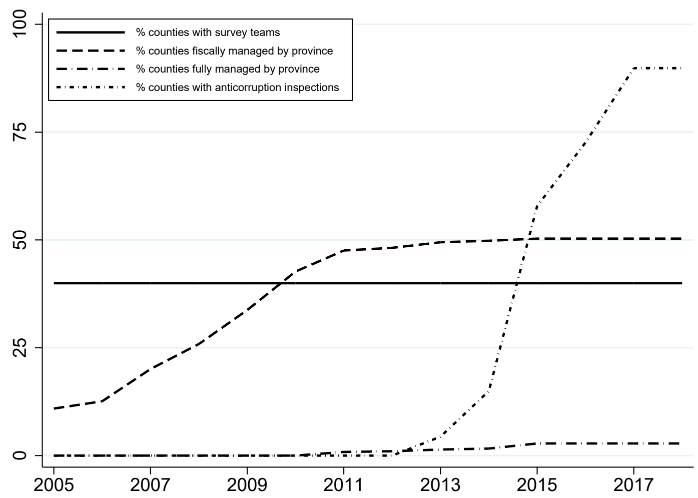


Notes: This figure shows the dynamic effect of the reform on local statistical capacity, and is estimated using the following specification:

$$Y_{ct} = \sum_{k=2004,2013,2018,k\neq 2008} \beta_k Treat_c \times 1_{\{t=k\}} + \delta_c + \lambda_g + \varepsilon_{ct} \tag{A1}$$

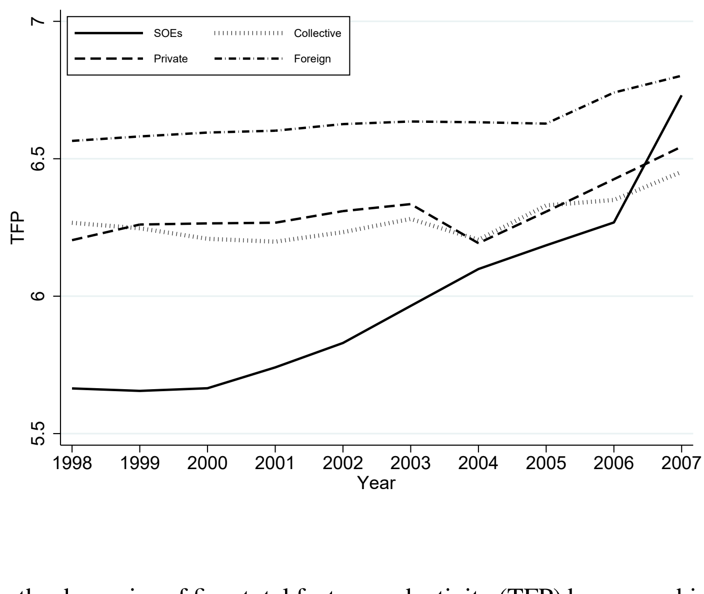
where Y_{ct} is a dummy variable equal to 1 if county c won an award for outstanding performance in conducting economic census in year t. The year 2008 is omitted as the reference group. This award is used as a proxy for local statistical capacity. The data on this award is only available for 2004, 2008, 2013, and 2018. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure A6: Rollout of concurrent reforms



Notes: This figure shows the rollout of various concurrent reforms. They include the fiscal province-managing-county (PMC) reform, the full province-managing-county (PMC) reform, and the anti-corruption campaign launched in 2013.

Figure A7: TFP by ownership

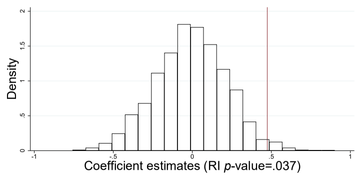


Notes: This figure shows the dynamics of firm total factor productivity (TFP) by ownership. It is created by taking the average TFP by ownership for manufacturing firms from 1998 to 2007 using the commonly used Annual Survery of Industrial Firms dataset. The calculation of the TFP uses the Levinsohn and Petrin method (Levinsohn and Petrin, 2003).

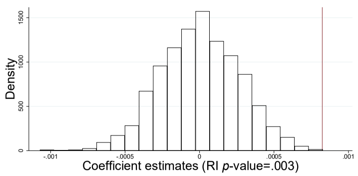
Figure A8: Effect on government policies - randomization inference

Key words frequency

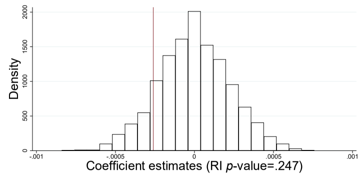
(a) Standardized index



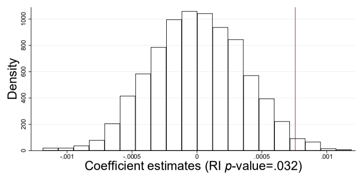
(b) Business attraction



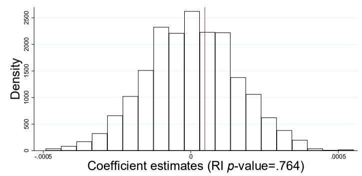
(c) Infrastructure investment



(d) Market reform

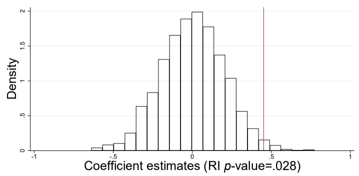


(e) Policy experimentation

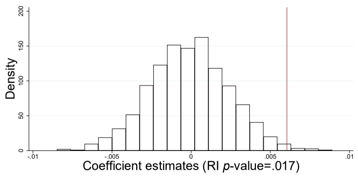


Topic score predicted by Random Forest

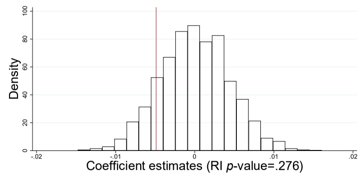
(f) Standardized index



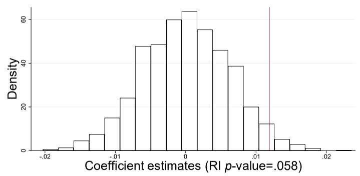
(g) Business attraction



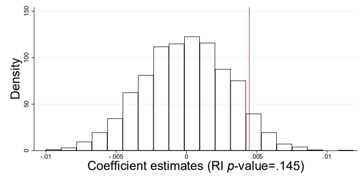
(h) Infrastructure investment



(i) Market reform

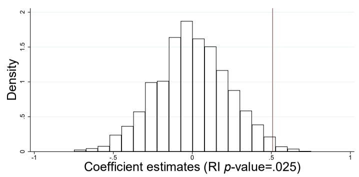


(j) Policy experimentation

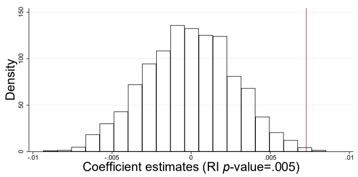


Topic score predicted by Support Vector Machine

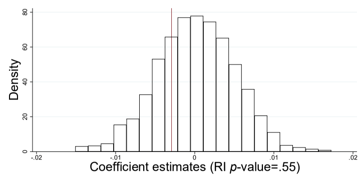
(k) Aggregate index



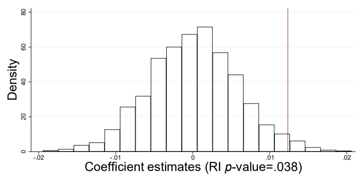
(l) Business attraction



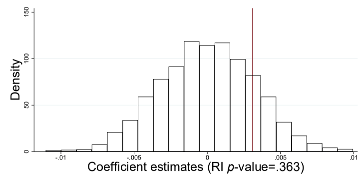
(m) Infrastructure investment



(n) Market reform



(o) Policy experimentation

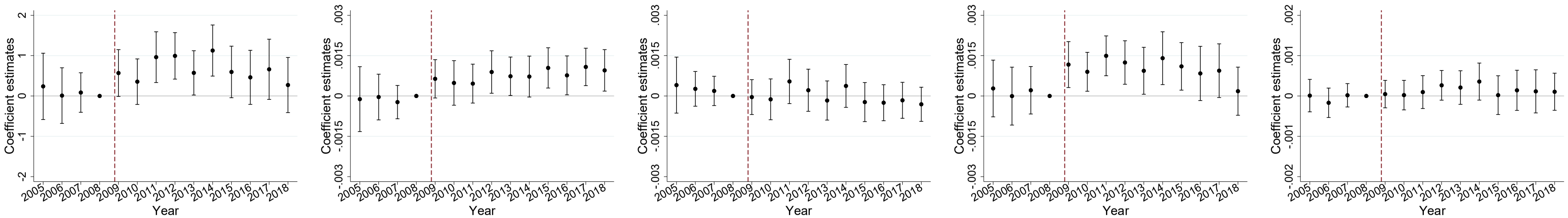


Notes: This figure adopts a randomization inference procedure with 2,000 permutations following Young (2019) to alleviate concerns about small sample size when estimating the effect of the reform on government policies. The true estimates are denoted by the vertical lines and the randomization inference p-values are reported below the figures. The unit of observation is county. The sample period is 2005-2018. The sample includes 97 counties. Panels (a)-(e) measure policies using a simple key words frequency method. Panels (f)-(o) measure policies using supervised machine learning methods (Random Forest and Support Vector Machine). The detailed procedures for constructing these measures are described in Appendix C. To alleviate multiple hypothesis testing issues, panels (a), (f), and (k) report estimates using a standardized index by summarizing the four policy measures following Kling, Liebman and Katz (2007).

Figure A9: Dynamic effect on government policies - IV estimates

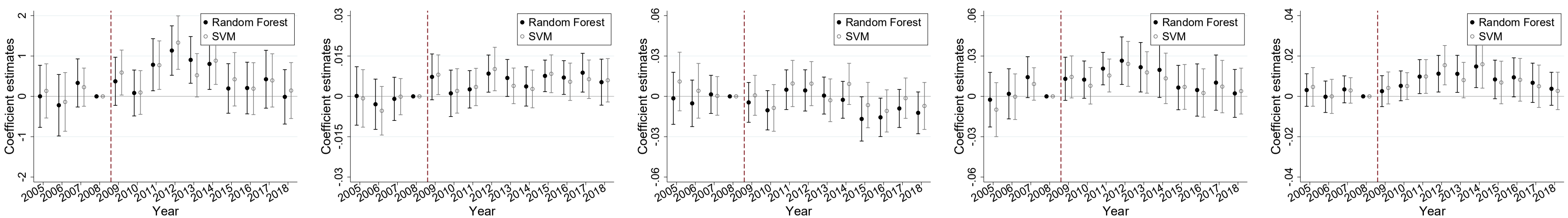
Key words frequency

(a) Standardized index (b) Business attraction (c) Infrastructure investment (d) Market reform (e) Policy experimentation



Machine learning-based topic score

(f) Standardized index (g) Business attraction (h) Infrastructure investment (i) Market reform (j) Policy experimentation

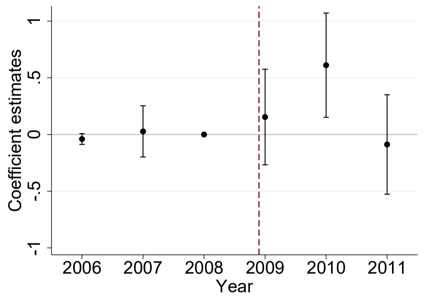


Notes: This figure shows the IV estimates on the dynamic effect of the reform on government policies across four policy areas. The unit of observation is county. The sample period is 2005-2018. The sample includes 97 counties. The estimation equations are event study variants of the specifications in Table 5 with the treatment instrumented by the randomly assigned rural survey teams in 1984. The year 2008, one year before the reform in 2009, is omitted as the reference year. Panels (a)-(e) measure policies using a simple key words frequency method. Panels (f)-(j) measure policies using supervised machine learning methods (Random Forest and Support Vector Machine). The detailed procedures for constructing these measures are described in Appendix C. To alleviate multiple hypothesis testing issues, panels (a) and (f) report estimates using a standardized index by summarizing the four policy measures following Kling, Liebman and Katz (2007). Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

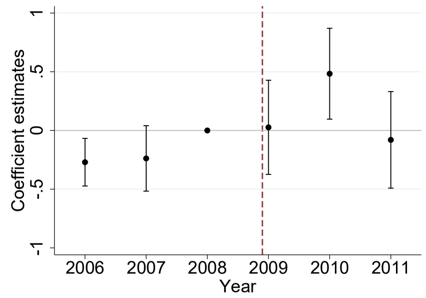
Figure A10: Dynamic effect on bank credit - IV estimates

Difference-in-differences estimates

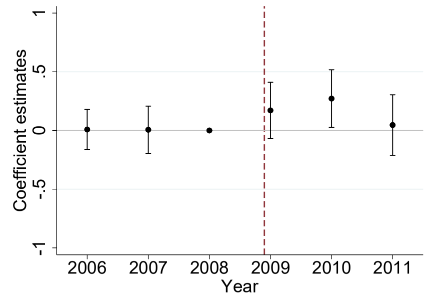
(a) IHS(Loans)



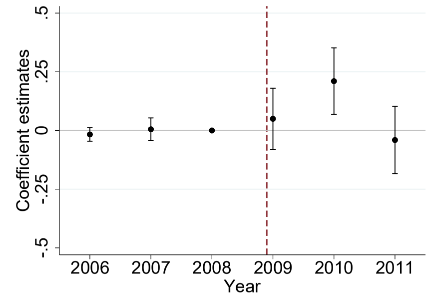
(b) IHS(Loans to small firms)



(c) IHS(# Firms granted bank loans)

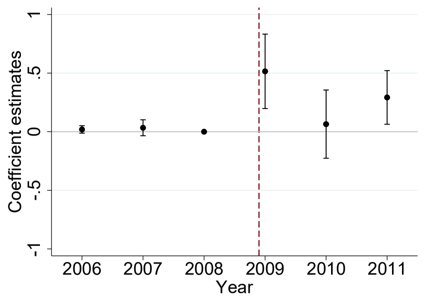


(d) IHS(# Branches with loan services)

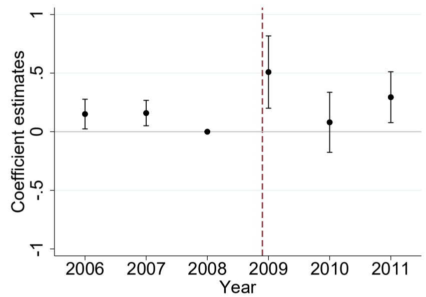


Difference-in-difference-in-differences estimates

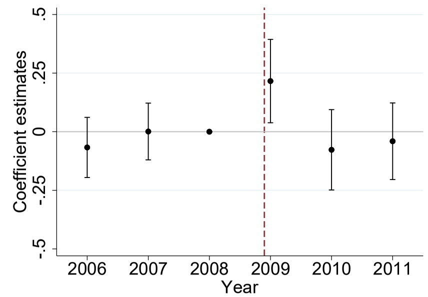
(e) IHS(Loans)



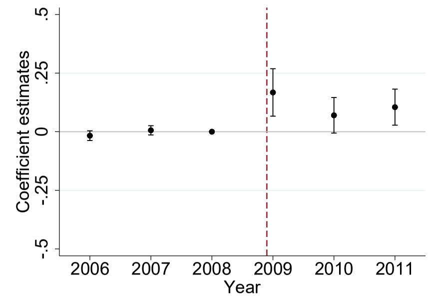
(f) IHS(Loans to small firms)



(g) IHS(# Firms granted bank loans)



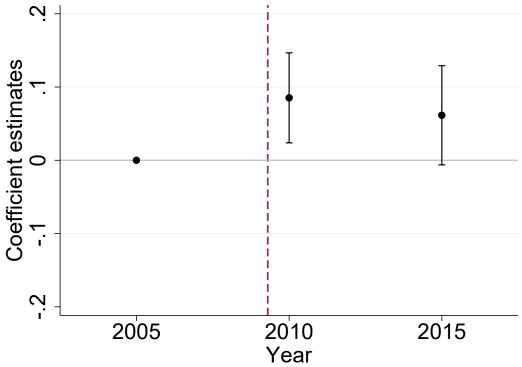
(h) IHS(# Branches with loan services)



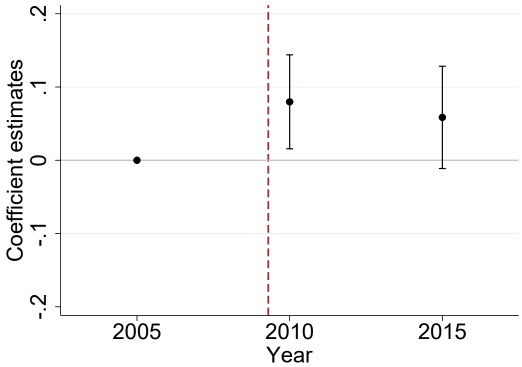
Notes: This figure shows IV estimates on the dynamic effect of the reform on bank credit. The unit of observation is county. The sample period is 2006-2011. The estimation equations are event study variants of the specifications in Table 6 with the treatment instrumented by the randomly assigned rural survey teams in 1984. The dependent variables are transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails of the skewed outcome distributions (Bellemare and Wichman, 2020). Panels (a)-(d) conduct a conventional difference-in-differences estimation. Panels (e)-(h) further conduct a difference-in-difference-in-differences estimation utilizing differential control of banks by local governments across counties. Government control over banks is measured as the standardized share of pre-reform loans from City Commercial Banks (CCBs) in a county, whose controlling shareholders are local governments. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure A11: Dynamic effect on firm entry - IV estimates

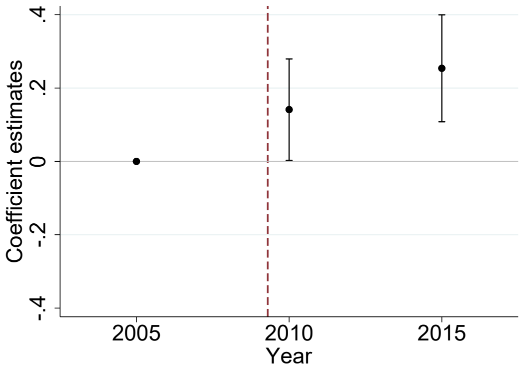
(a) IHS(# All firms)



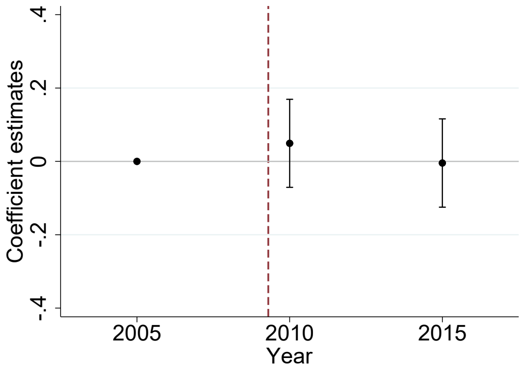
(b) IHS(# Private firms)



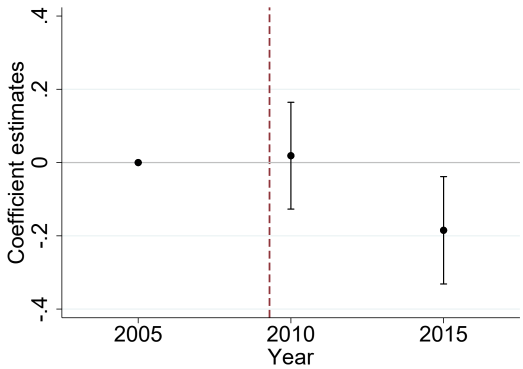
(c) IHS(# SOEs)



(d) IHS(# Foreign firms)



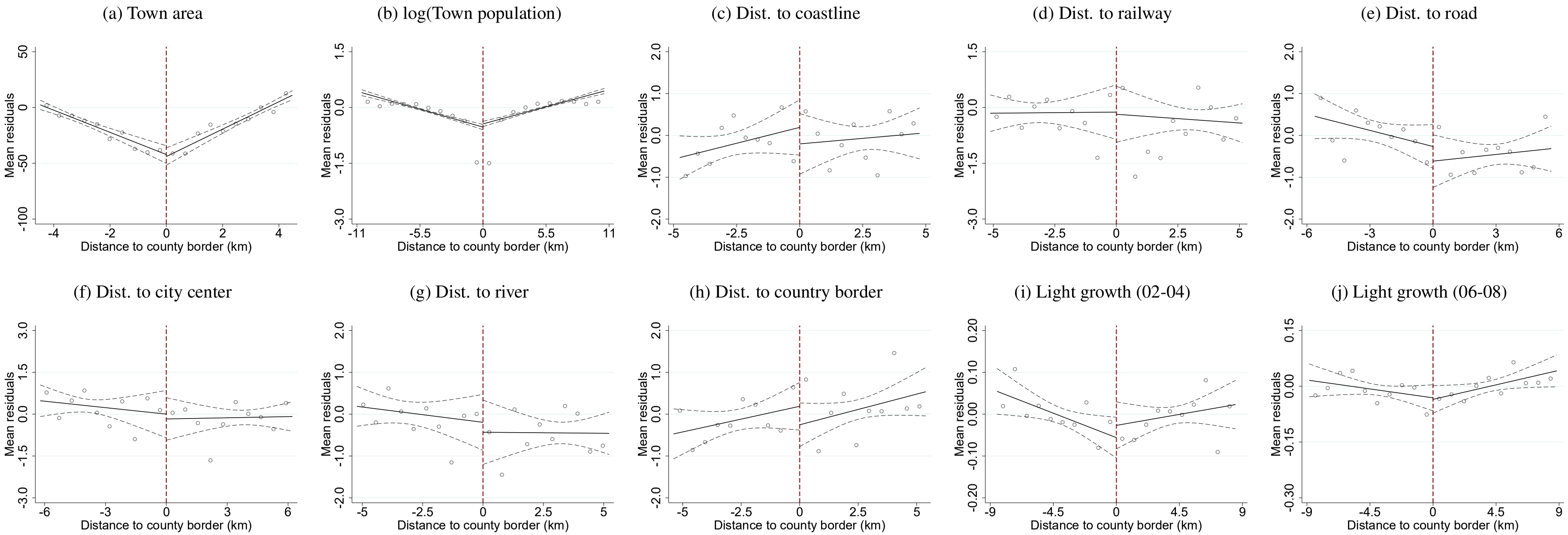
(e) IHS(# Collective firms)



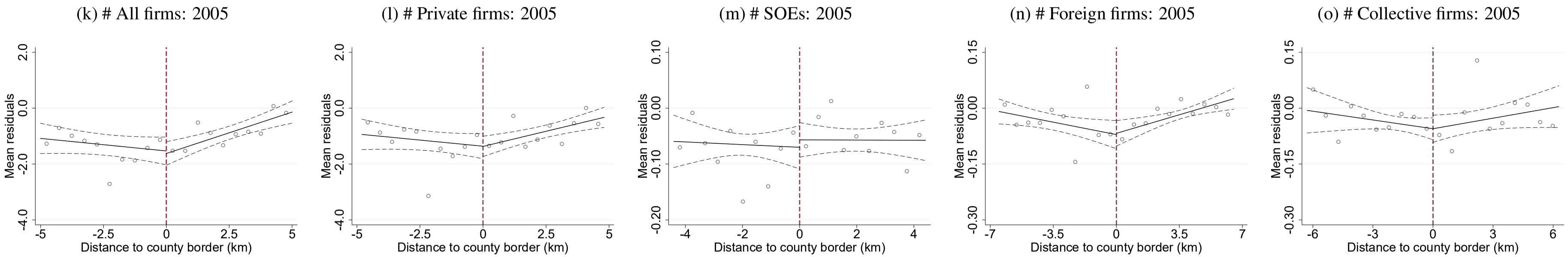
Notes: This figure shows the IV estimates on the dynamic effect of the reform on firm entry. The unit of observation is county. The sample includes the years 2005, 2010, and 2015. The estimation equations are event study variants of the specifications in Table 7 with the treatment instrumented by the randomly assigned rural survey teams in 1984. The dependent variables are transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails of the skewed outcome distributions (Bellemare and Wichman, 2020). Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure A12: Effect on firm entry - RD balance check and placebo test

Balance test using preexisting town covariates



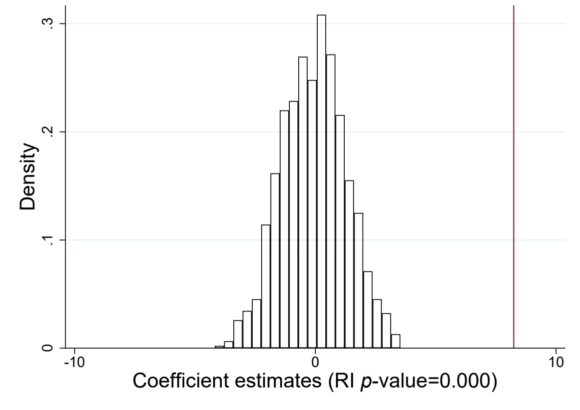
Placebo test using pre-reform firm entry data



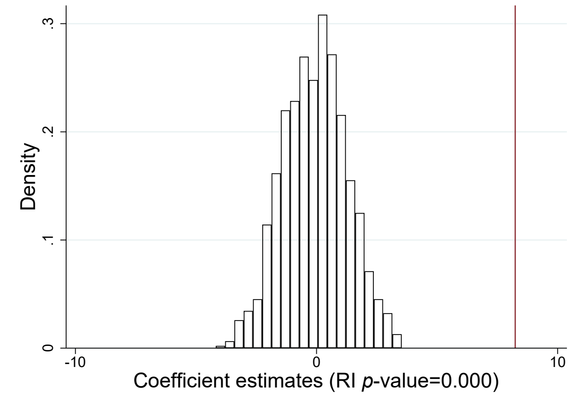
Notes: This figure conducts a balance check for the RD design using preexisting town covariates (panels a-j) and also a placebo test using pre-reform firm entry data (panels k-o). The unit of observation is town. The circles denote the mean of town covariates or firm entry within a distance bin, after partialling out county border fixed effects. The solid lines are fitted separately for each side of the county border. The dashes are the associated 90% confidence intervals constructed using standard errors clustered at the county border level. The bandwidths are the optimal bandwidths proposed by Calonico, Cattaneo and Titiunik (2014).

Figure A13: Effect on firm entry - Randomization inference for RD estimates

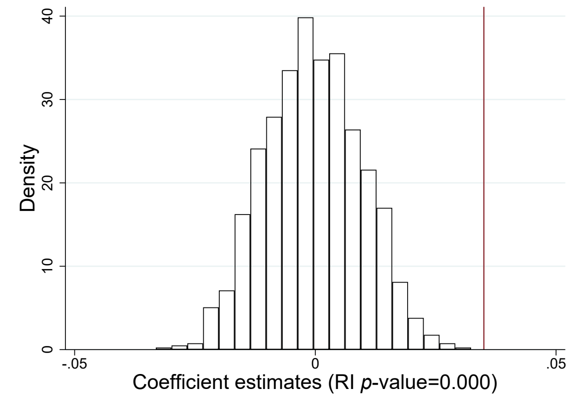
(a) # All firms



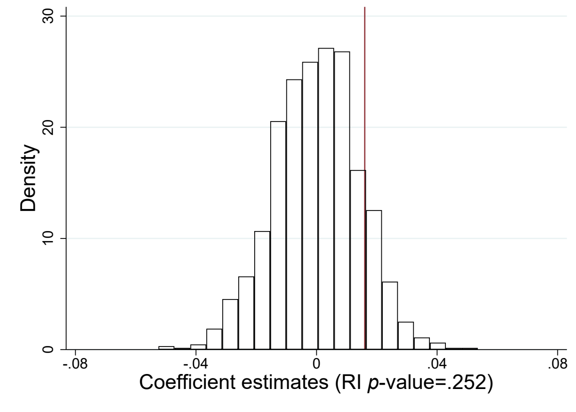
(b) # Private firms



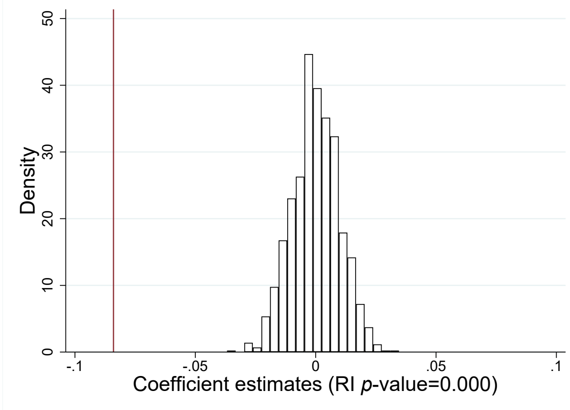
(c) # SOEs



(d) # Foreign firms

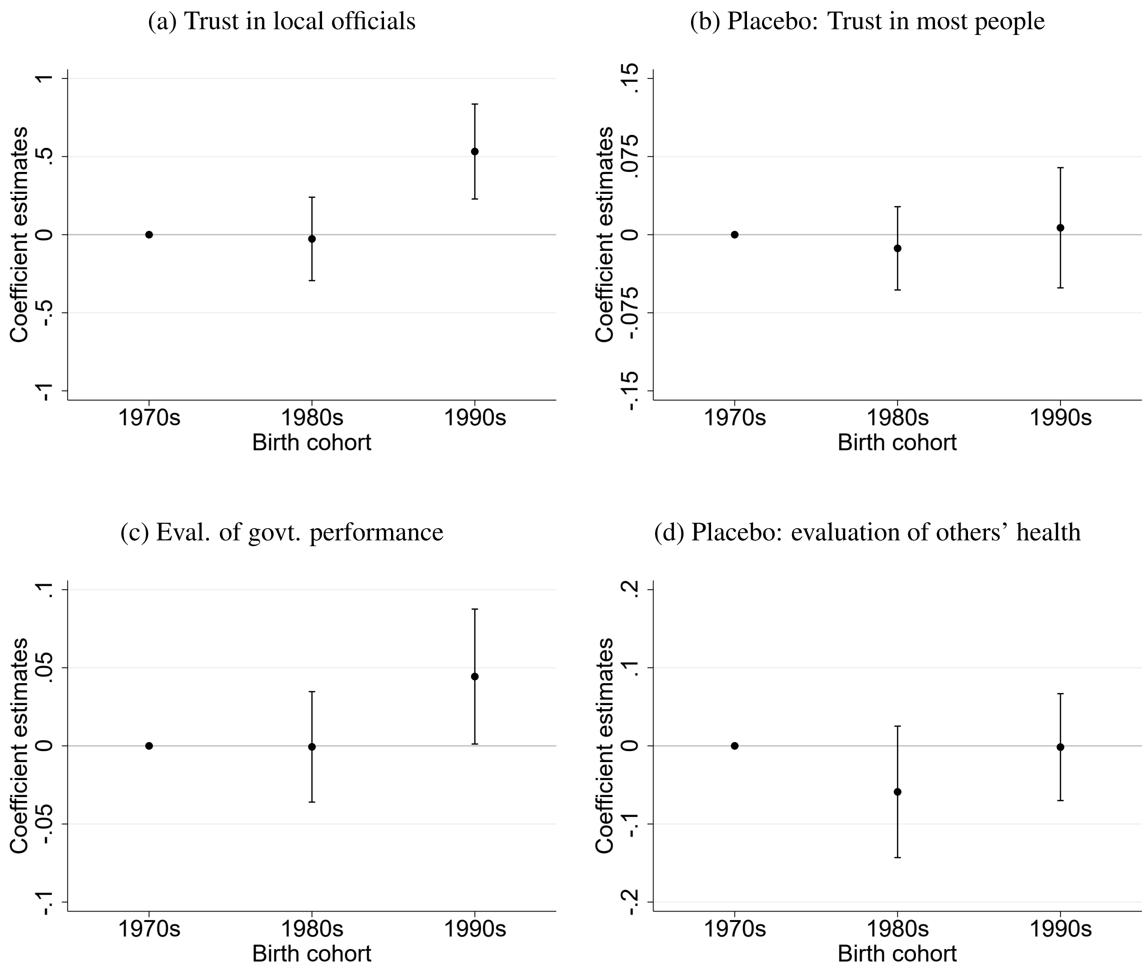


(e) # Collective firms



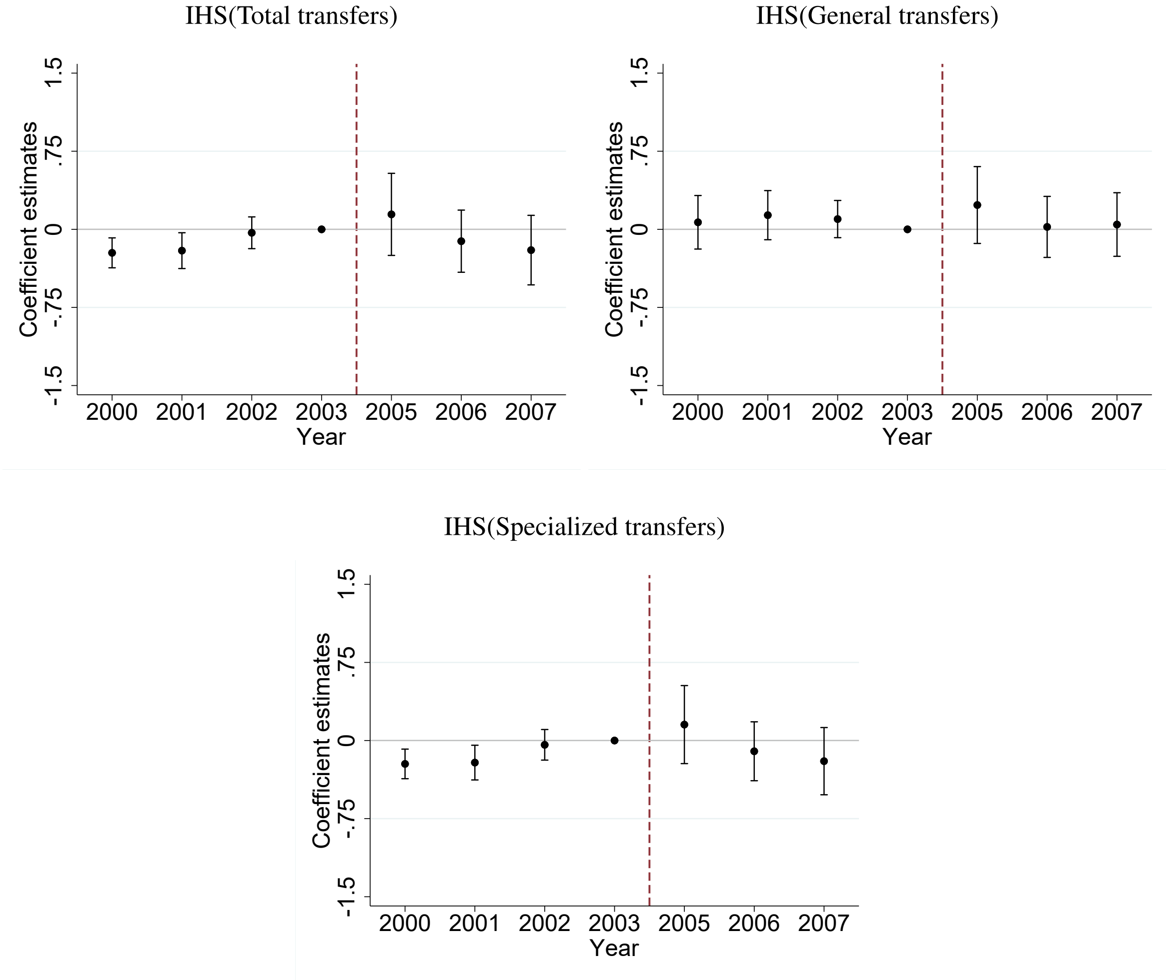
Notes: This figure adopts a randomization inference procedure with 2,000 permutations to show the robustness of the RD estimates following Ganong and Jäger (2018). Specifically, I keep only county borders across which there are no treatment variations, and then randomly create placebo treatment variation to each border. Based on these placebo borders, I then re-estimate the RD equation 13. This process is repeated for 2,000 times. The true estimates are denoted by the vertical lines and the randomization inference p-values are reported below the figures.

Figure A14: Effect on citizen attitudes - IV estimates



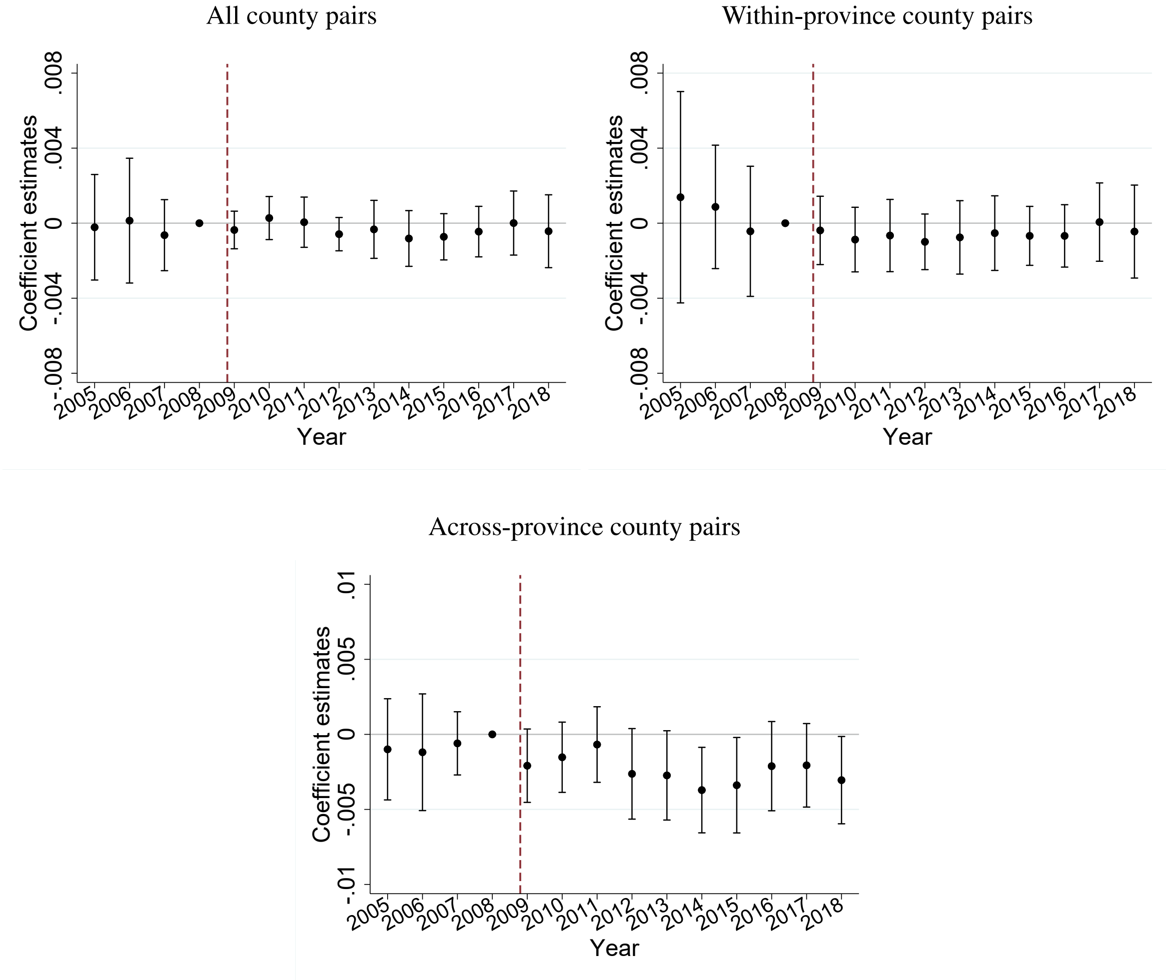
Notes: This figure shows the IV estimates on the effect of the reform on citizens' attitudes towards local governments, and is created by visualizing the results in Table A21. The unit of observation is citizen (three waves of survey data from the China Family Panel Studies, 2012, 2014 and 2016). The estimation equation is a cohort difference-in-differences specification (equation 14) that utilizes two sources of variation: (1) treatment counties versus control counties; (2) most affected cohorts versus less unaffected cohorts within the same county. The latter source of variation is built on insights in the psychology and political science literature that citizens' political attitudes are most permeable during teenage years and keep stable since one's 30s (Wolfinger and Rosenstone, 1980; Krosnick and Alwin, 1989). As such, the 1990s cohort are defined as the most affected group as they would be younger than 30 during the reform period, and thus, their political attitudes would be most permeable. The 1980s cohort are defined as the less affected group because some of them would be older than 30 during the reform period. The omitted group is those born in the 1970s, as they would be older than 30 in the reform period and thus be unaffected by the reform. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure A15: Dynamic effect on fiscal transfers around 2005



Notes: This table examines whether the launch of the survey teams would increase fiscal transfers from upper-level governments using an event study specification. The unit of observation is county. The sample is from 2000 to 2007. The dependent variables are transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails of the skewed outcome distributions (Bellemare and Wichman, 2020). Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

Figure A16: Dynamic effect on policy diffusion

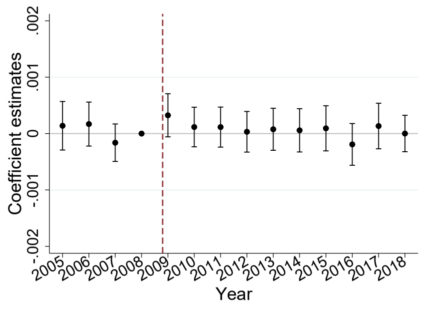


Notes: This figure shows the dynamic effect of the reform on policy diffusion across counties. The unit of observation is county pair. The dependent variable is the pairwise textual similarity of government work reports, which is calculated following Kelly et al. (2021) and described in Appendix C. Panel (a) considers all county pairs. Panel (b) considers county pairs within the same province. Panel (c) considers county pairs spanning different provinces. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are two-way clustered by both counties in a pair.

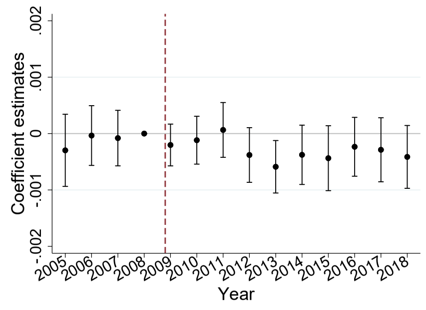
Figure A17: Effect on Government Policies - Social Welfare and Environmental Protection

Key words frequency

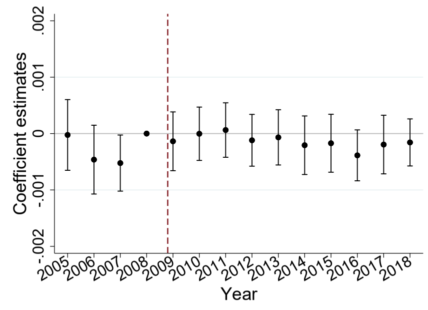
(a) Social welfare, OLS



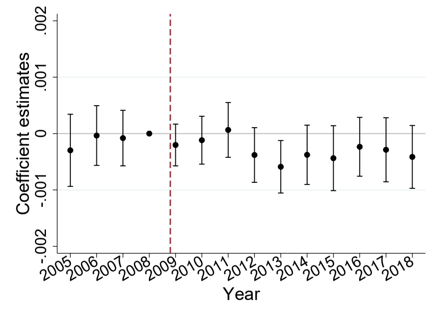
(b) Environmental protection, OLS



(c) Social welfare, IV

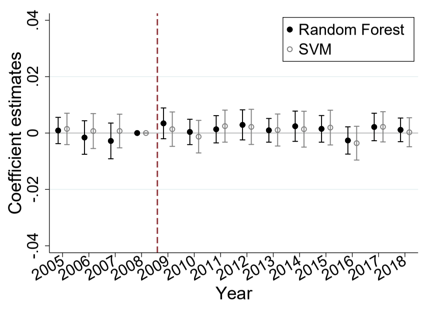


(d) Environmental protection, IV

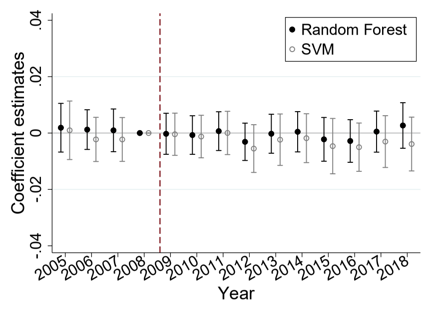


Machine learning-based topic score

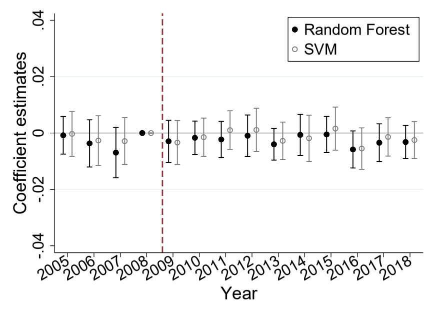
(e) Social welfare, OLS



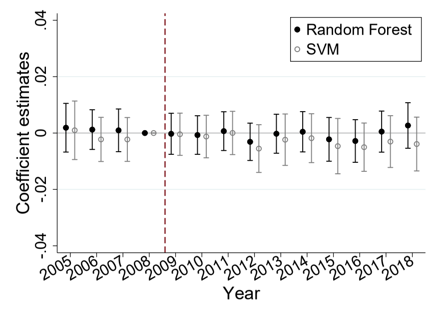
(f) Environmental protection, OLS



(g) Social welfare, IV



(h) Environmental protection, IV

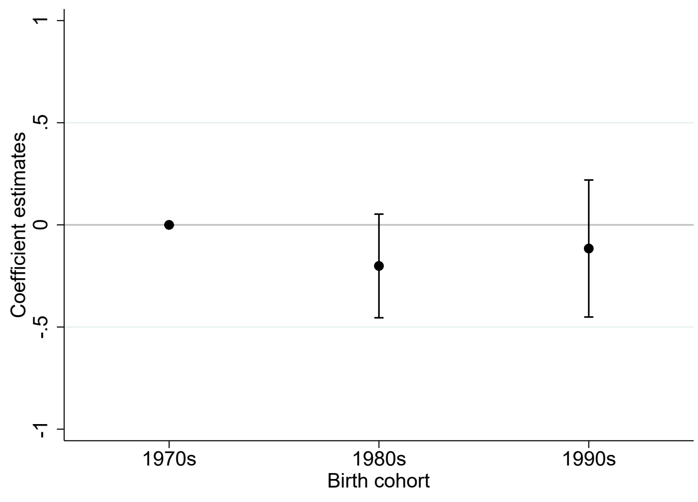


Notes: This figure shows the dynamic effect of the reform on government policies on social welfare and environmental protection. The unit of observation is county. The sample period is 2005-2018. The sample includes 97 counties. The policies are measured using either a simple key words frequency method or supervised machine learning methods (Random Forest and Support Vector Machine). The detailed procedures for constructing these measures are described in Appendix C. Standard errors used to construct the 90%

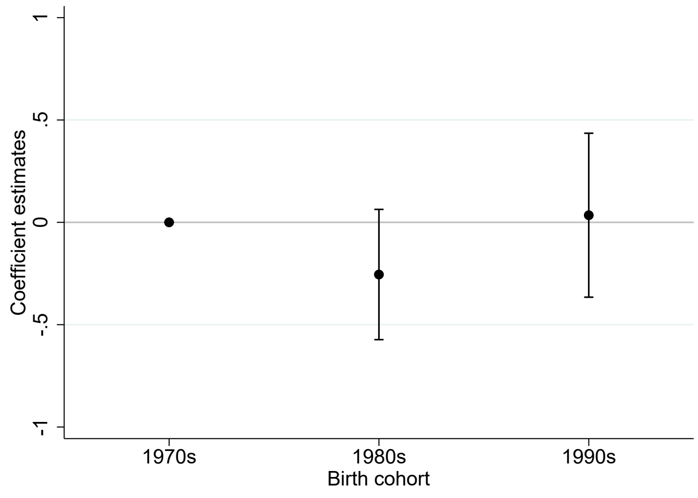
confidence intervals, denoted by the spikes, are clustered at the county level.

Figure A18: Effect on perceived corruption

(a) OLS estimates



(b) IV estimates



Notes: Notes: This figure shows the effect of the reform on citizens' perceived corruption about the government, and is created by visualizing the results in Table A28. The unit of observation is citizen (two waves of survey data from the China Family Panel Studies: 2014 and 2016). The estimation equation is a cohort difference-in-differences specification (equation 14) that utilizes two sources of variation: (1) treatment counties versus control counties; (2) most affected cohorts versus less unaffected cohorts within the same county. The latter source of variation is built on insights in the psychology and political science literature that citizens' political attitudes are most permeable during teenage years and keep stable since one's 30s (Wolfinger and Rosenstone, 1980; Krosnick and Alwin, 1989). As such, the 1990s cohort are defined as the most affected group as they would be younger than 30 during the reform period, and thus, their political attitudes would be most permeable. The 1980s cohort are defined as the less affected group because some of them would be older than 30 during the reform period.

The omitted group is those born in the 1970s, as they would be older than 30 in the reform period and thus be unaffected by the reform. Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level.

E Additional tables

Table A1: Decomposing the effect of the reform

	(1)	(2)	(3)	(4)
	OLS	OLS	IV	IV
Dep. var.:	Reported GDP growth (%)	Light growth (%)	Reported GDP growth (%)	Light growth (%)
Treat × Post	-0.576***	0.254	-0.592***	0.162
	(0.161)	(0.294)	(0.203)	(0.355)
County FE	X	X	X	X
Year FE	X	X	X	X
Province FE × Post			X	X
Demographic controls × Post	X	X	X	X
Economic controls × Post	X	X	X	X
Geographic controls × Post	X	X	X	X
Cluster level	County	County	County	County
Observations	20,276	20,969	20,276	20,969
R-squared	0.361	0.364	0.100	0.020
Mean dep. var.	10.84	9.01	10.84	9.01
Effective F-statistic			2,278	2,287

Notes: This table shows the effect of the reform on reported GDP growth and light growth separately. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Columns (1)-(2) present OLS estimates. Columns (3)-(4) present IV estimates, with Treat instrumented by the randomly assigned rural survey teams in 1984. The effective F-statistics are constructed following Montiel Olea and Pflueger (2013). Standard errors clustered at the county level are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A2: Estimating spillover effect - robustness

	(1)	(2)	(3)	(4)
Dep. var.:	Reported GDP growth (%)			
Treat × Post	-0.610***	-0.559***	-0.581***	-0.576***
	(0.162)	(0.164)	(0.164)	(0.164)
# Treat neighbors within 50km × Post	-0.094			
	(0.091)			
# Treat neighbors within 100km × Post		0.027		
		(0.036)		
# Treat neighbors (GDP weighted) × Post			-0.011	
			(0.068)	
# Treat neighbors (population weighted) × Post				0.003
				(0.069)
Light growth (%)	0.017***	0.017***	0.017***	0.017***
	(0.005)	(0.005)	(0.005)	(0.005)
County FE	X	X	X	X
Year FE	X	X	X	X
County controls × Post	X	X	X	X
Neighbor number FE × Post	X	X	X	X
Cluster level	County	County	County	County
Mean dep. var.	10.84	10.84	10.84	10.84
Mean number of neighbors	2.40	11.45	5.91	5.91
Mean number of treat neighbors	0.88	4.47	1.97	1.97

Notes: The tables shows the robustness of the spillover effect of the reform on GDP growth manipulation. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. In columns (1) and (2), neighbors are defined as counties within a certain distance of a county, while in columns (3) and (4) neighbors are defined as counties sharing a common boundary segment with a county and are weighted by their sizes (GDP or population). Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Testing for monotonicity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep.var:	Treat $\times$ Post							
Group var.:	log Pop. (2010)		Share urban (%, 2010)		Share 15-64 (%, 2010)		Years of schooling (2010)	
	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)
Treat^{1984} $\times$ Post	0.747***	0.743***	0.768***	0.717***	0.762***	0.715***	0.748***	0.729***
	(0.024)	(0.021)	(0.022)	(0.025)	(0.023)	(0.025)	(0.024)	(0.024)
Observations	12,245	12,264	12,248	12,261	12,217	12,292	12,021	12,488
R-squared	0.905	0.900	0.900	0.881	0.897	0.881	0.892	0.885
Group var.:	Unem. rate (%, 2010)		Share 1st sectors (%, 2010)		Share 2nd sectors (%, 2010)		log GDP (2004)	
	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)
Treat^{1984} $\times$ Post	0.775***	0.703***	0.724***	0.751***	0.741***	0.737***	0.750***	0.748***
	(0.022)	(0.026)	(0.025)	(0.023)	(0.024)	(0.024)	(0.025)	(0.021)
Observations	12,049	12,460	12,250	12,259	12,231	12,278	11,251	13,258
R-squared	0.902	0.876	0.887	0.895	0.891	0.890	0.902	0.893
Group var.:	log GDP (2008)		GDP growth (%, 2002-2004)		GDP growth (%, 2006-2008)		Light growth (%, 2002-2004)	
	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)
Treat^{1984} $\times$ Post	0.755***	0.738***	0.725***	0.750***	0.725***	0.751***	0.736***	0.744***
	(0.023)	(0.022)	(0.025)	(0.023)	(0.025)	(0.024)	(0.024)	(0.024)
Observations	12,371	12,138	10,945	13,564	12,246	12,263	12,167	12,342
R-squared	0.905	0.893	0.888	0.890	0.883	0.894	0.887	0.892
Group var.:	Light growth (%, 2006-2008)		Dist. major roads (km, 2010)		Dist. major rail. (km, 2010)		County area (km2)	
	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)
Treat^{1984} $\times$ Post	0.731***	0.747***	0.765***	0.716***	0.778***	0.711***	0.779***	0.701***
	(0.024)	(0.024)	(0.023)	(0.025)	(0.023)	(0.025)	(0.023)	(0.025)
Observations	12,210	12,299	11,997	12,512	12,068	12,441	12,180	12,329
R-squared	0.886	0.892	0.898	0.885	0.897	0.881	0.903	0.881
Group var.:	Precipitation (inches, 2004)		Temperature (degrees, 2004)		Precipitation (inches, 2008)		Temperature (degrees, 2008)	
	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)
Treat^{1984} $\times$ Post	0.739***	0.739***	0.730***	0.743***	0.726***	0.752***	0.729***	0.746***
	(0.024)	(0.024)	(0.025)	(0.023)	(0.025)	(0.023)	(0.025)	(0.023)
Observations	12,301	12,208	12,273	12,236	12,301	12,208	12,273	12,236
R-squared	0.891	0.893	0.887	0.891	0.887	0.898	0.886	0.891
Group var.:	Dist. to major rivers (km)		Dist. to country border (km)		Dist. to coastline (km)		Dist. to city center (km)	
	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)	<p(50)	>=p(50)
Treat^{1984} $\times$ Post	0.776***	0.702***	0.706***	0.778***	0.740***	0.733***	0.753***	0.734***
	(0.022)	(0.026)	(0.025)	(0.022)	(0.024)	(0.025)	(0.023)	(0.024)
Observations	12,249	12,260	12,264	12,245	12,180	12,329	12,213	12,296
R-squared	0.903	0.875	0.879	0.899	0.893	0.883	0.891	0.888

Notes: This table provides evidence to support the monotonicity assumption required for IV estimation in the presence of heterogeneous treatment effects. It does so by reporting the first-stage results for subsamples divided by the medians of baseline county covariates. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Treat^{1984} is a dummy variable indicating counties with a randomly assigned rural survey team in 1984. Post is a dummy variable indicating years after the reform in 2009. The estimation equation is equation (9). County fixed effects, year fixed effects, province fixed effects interacted the post-reform dummy, and a control for light growth are included but not shown due to space limitation. Standard errors clustered at the county level are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A4: Reweighting OLS to match IV

	(1)	(2)	(3)	(4)	(5)
Dep. var.:			Reported GDP growth (%)		
Treat × Post	-0.576***	-0.565***	-0.568***	-0.567***	-0.535***
	(0.161)	(0.161)	(0.161)	(0.162)	(0.162)
Light growth (%)	0.017***	0.016***	0.016***	0.016***	0.016***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
County FE	X	X	X	X	X
Year FE	X	X	X	X	X
County controls × Post	X	X	X	X	X
Reweighting		X	X	X	X
Cluster level	County	County	County	County	County
Observations	20,273	20,273	20,273	20,273	20,273
R-squared	0.362	0.361	0.361	0.361	0.360
Number of subgroups	-	10	20	30	2x2x2x2x2
Mean dep. var.	10.84	10.84	10.84	10.84	10.84

Notes: This table adopts a reweighting method to adjust the OLS estimates to match the sample of compliers. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Column (1) reproduces the baseline OLS estimates for comparison. Columns (2)-(4) divides the sample into 10-30 groups of equal size, based on quantiles of the first principal component of baseline county covariates, and then reweight the raw OLS estimation using the complier share in each group as weights. Column (5) divides the sample into 32 groups using the medians of the first five principal components of baseline county covariates, and then reweights the OLS using the complier share in each group as weights. Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Effect in a trimmed sample without personnel changes

	(1)	(2)	(3)
Dep. var.:		Reported GDP growth (%)	
Sample:	Baseline	Terms straddling 2009	Terms covering 2007-2011
Treat × Post	-0.576***	-0.617***	-0.647***
	(0.161)	(0.207)	(0.207)
Light growth rate (%)	0.017***	0.017***	0.016***
	(0.005)	(0.005)	(0.005)
County FE	X	X	X
Year FE	X	X	X
County controls × Post	X	X	X
Cluster level	County	County	County
Observations	20,273	16,722	16,384
R-squared	0.362	0.370	0.367
Mean dep. var.	10.84	10.98	10.94

Notes: This table examines whether the reduction in manipulation stemmed from personnel changes among local officials, by utilizing a trimmed sample in which there were no personnel changes in the treatment counties. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Standard errors clustered at the county level are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A6: Effects on leader traits and turnovers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. var:	Years in office	Term length	Age	1(Local)	Schooling	Connection	1(Turnover)
Panel A: Magistrates							
Treat × Post	0.046	0.104	0.048	-0.005	-0.147	0.052*	0.001
	(0.160)	(0.071)	(0.230)	(0.017)	(0.112)	(0.028)	(0.007)
Observations	22,628	22,628	9,876	7,626	7,038	8,211	22,628
R-squared	0.881	0.975	0.866	0.873	0.858	0.504	0.173
Mean dep. var.	2.49	4.53	45.13	0.19	16.02	0.24	0.06
Panel B: Party secretaries							
Treat × Post	0.113	0.029	-0.075	-0.009	-0.157	-0.029	-0.001
	(0.163)	(0.071)	(0.294)	(0.016)	(0.104)	(0.028)	(0.006)
Observations	22,695	22,695	10,674	8,243	7,994	9,027	22,695
R-squared	0.879	0.976	0.769	0.857	0.829	0.475	0.151
Mean dep. var.	2.73	4.86	47.68	0.16	16.40	0.24	0.05
County FE	X	X	X	X	X	X	X
Year FE	X	X	X	X	X	X	X
Cluster level	County	County	County	County	County	County	County

Notes: This table test the effects of the reform on personnel turnovers and personnel traits. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Standard errors clustered at the county level are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A7: Estimating promotion incentives

	(1)	(2)	(3)	(4)
Dep. var.:	1(Promotion)			
Official type:	Party secretary		Magistrate	
Model:	LPM	Probit	LPM	Probit
Start age	-0.033***	-0.234***	-0.017**	-0.176***
	(0.009)	(0.056)	(0.007)	(0.065)
Connection	0.022**	0.142***	0.019**	0.139**
	(0.009)	(0.053)	(0.008)	(0.066)
Education	0.008	0.035	0.017**	0.152*
	(0.009)	(0.062)	(0.007)	(0.087)
Start age × Connection	0.012	0.107**	0.005	0.056
	(0.009)	(0.049)	(0.007)	(0.055)
Start age × Education	-0.002	-0.012	0.001	0.018
	(0.010)	(0.059)	(0.008)	(0.077)
Connection × Education	0.020**	0.105**	0.017**	0.117*
	(0.009)	(0.053)	(0.008)	(0.071)
Observations	1,093	1,093	1,018	1,018

Notes: This table estimates the promotion incentives of county leaders. The unit of observation is local leader, either the party secretary or the magistrate in a county. The sample period is 2005-2018. The promotion incentives denote local leaders' ex ante likelihood of promotion based on their start ages, years of schooling, and political connections with upper-level leaders following Avis, Ferraz and Finan (2018) and Wang, Zhang and Zhou (2020). Odd columns adopt linear probability models and even columns adopt Probit models. Note that other performance variables, such as GDP growth, are intentionally excluded from this regression, so the estimated probabilities capture the ex ante likelihood of promotion (Wang, Zhang and Zhou, 2020). Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Testing the promotional discipline effect

	(1)	(2)	(3)
Dep. var.:	Reported GDP growth (%)		
Treat × Post	-0.576***	-0.658***	-0.562**
	(0.161)	(0.243)	(0.269)
Light growth rate (%)	0.017***	0.019***	0.011
	(0.005)	(0.007)	(0.008)
Treat × Post × 1(Age>52)		0.253	
		(1.164)	
1(Age>52)		-0.931	
		(0.687)	
Treat × 1(Age>52)		-0.081	
		(1.129)	
Post × 1(Age>52)		0.535	
		(0.709)	
Treat × Post × Promotion incentive			0.127
			(0.217)
Promotion incentive			-0.112
			(0.138)
Treat × Promotion incentive			-0.027
			(0.210)
Post × Promotion incentive			0.042
			(0.146)
County FE	X	X	X
Year FE	X	X	X
Cluster level	County	County	County
Observations	20,273	17,980	11,989
R-squared	0.362	0.355	0.370
Mean dep. var.	10.84	10.67	10.77

Notes: This table tests the promotional discipline effect, by checking whether the reform effect is larger for local officials with greater promotion incentives. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Promotion incentive denotes local officials' ex ante likelihood of promotion estimated based on their start ages, years of schooling, and political connections with upper-level leaders (Avis, Ferraz and Finan, 2018; Wang, Zhang and Zhou, 2020), and are estimated separately for the party secretary and the magistrate. Column (1) reproduces the baseline estimate. In columns (2)-(3), I adopt a slightly different specification to allow the two leaders in a county (the party secretary and the magistrate) to separately affect GDP growth manipulation. This means that for each county-year, I generate two parallel observations that are identical except for one distinction: one includes only the party secretary, while the other includes only the magistrate. Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Testing the soft information channel

	(1)	(2)	(3)
Dep. var.:		Reported GDP growth (%)	
Treat × Post	-0.576***	-0.575***	-0.549*
	(0.161)	(0.163)	(0.287)
Light growth (%)	0.017***	0.017***	0.015**
	(0.005)	(0.005)	(0.007)
Treat × Post × Distance to upper-level govt.		0.009	
		(0.175)	
Treat × Post × Connection			-0.057
			(0.069)
Connection			0.058
			(0.046)
Treat × Connection			-0.073
			(0.064)
Post × Connection			-0.030
			(0.050)
County FE	X	X	X
Year FE	X	X	X
County controls × Post	X	X	X
Cluster level	County	County	County
Observations	20,273	20,273	10,802
R-squared	0.362	0.362	0.365
Mean of dep. var.	10.84	10.84	10.64

Notes: This table tests whether the reduction in manipulation was driven by soft information provided by the survey teams. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Distance to upper-level govt. denotes the shortest distance from a county to its overseeing upper-level government. Connection is a dummy variable that equals 1 if the party secretary or magistrate in a county shares the same hometown or educational background with upper-level leaders, and 0 otherwise. Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Effect on statistical capacity

	(1)	(2)
Dep. var.:	1(Award for outstanding performance in economic census)	
Treat × Post	0.011	-0.005
	(0.017)	(0.020)
County FE	X	X
Year FE	X	X
County controls × Post		X
Cluster level	County	County
Observations	7,116	5,992
R-squared	0.391	0.420
Mean of dep. var.	0.13	0.13

Notes: This table shows the effect of the reform on local statistical capacity. The unit of observation is county. The sample period is 2004, 2008, 2013, and 2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. The dependent variable is a dummy denoting whether a county won an award for outstanding performance in conducting economic census. Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Controlling for concurrent reforms

	(1)	(2)	(3)	(4)
Dep. var.:	Reported GDP growth (%)			
Treat × Post	-0.593***	-0.579***	-0.576***	-0.596***
	(0.162)	(0.161)	(0.161)	(0.162)
Light growth (%)	0.016***	0.017***	0.017***	0.016***
	(0.005)	(0.005)	(0.005)	(0.005)
Fiscal PMC	X			X
Full PMC		X		X
Anticorruption inspection			X	X
County FE	X	X	X	X
Year FE	X	X	X	X
County controls × Post	X	X	X	X
Cluster level	County	County	County	County
Observations	20,273	20,273	20,273	20,273
R-squared	0.362	0.362	0.362	0.363
Mean of dep. var.	10.84	10.84	10.84	10.84

Notes: This table shows the robustness of te baseline results after controlling for several concurrent reforms that may also strengthen the monitoring of local officials. The unit of observation is county. The sample period is 2005-2018. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Fiscal PMC is a dummy for the fiscal province-managing-county (PMC) reform. Full PMC is a dummy for the full province-managing-county (PMC) reform. Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Keywords in each policy

Policy	Chinese keywords	English translation
business attraction	招商引资	attract businesses
	外商直接投资	foreign direct investment
	外资	foreign capital
	对外开放	open up
infrastructure	基础设施	infracstrcture
	基建	shorthand term for infractrcture
	工程建设	project construction
	建设项目	construction project
market reform	改革	reform
	非公有制	non-public ownership
	民营企业	private firms
	私营企业	private firms [variant]
	民企	shorthand term for private firms
	私企	shorthand term for private firms [variant]
policy experimentation	试点	experimental places
	试验区	experimental zones
social welfare	社会保险	social insurance
	社保	shorthand term for social insurance
	养老保险	endowment insurance
	医疗保险	medical insurance
	养老金	pension
	社会保障	social security
Environmental protection	环境保护	environmental protection
	环保	shorthand term for environmental protection
	污染治理	pollution control
	减排	emission reduction

Notes: This table lists the keywords in each policy, which is used to create keywords frequency measures of local officials' emphasis on each policy. The first four policies are used in Section 5.1. The last two policies are examined in Section 5.5.

Table A13: Effect on government policies - IV estimates

Dep. var.:	(1) Standardized index	(2) Business attraction	(3) Infrastructure investment	(4) Market reform	(5) Policy experimentation
Panel A: Key words frequency					
Treat × Post	0.664***	0.001***	-0.000	0.001***	0.000
	(0.201)	(0.000)	(0.000)	(0.000)	(0.000)
R-squared	0.114	0.139	0.105	0.096	0.083
Mean dep. var.	0	0.002	0.002	0.004	0.001
Panel B: Topic score predicted by Random Forest					
Treat × Post	0.539***	0.007***	-0.005	0.011*	0.007**
	(0.188)	(0.002)	(0.005)	(0.006)	(0.003)
R-squared	0.063	0.091	0.047	0.047	0.059
Mean dep. var.	0	0.013	0.040	0.068	0.016
Panel C: Topic score predicted by Support Vector Machine					
Treat × Post	0.566***	0.007**	-0.004	0.011**	0.007**
	(0.195)	(0.003)	(0.005)	(0.005)	(0.003)
R-squared	0.088	0.118	0.063	0.055	0.084
Mean dep. var.	0	0.014	0.051	0.068	0.019
County FE	X	X	X	X	X
Year FE	X	X	X	X	X
Province FE × Post	X	X	X	X	X
County controls × Post	X	X	X	X	X
Cluster level	County	County	County	County	County
Observations	883	883	883	883	883
Effective F-statistic	74.06	74.06	74.06	74.06	74.06

Notes: This table shows the IV estimates on the effect of the reform on government policies across four policy areas. The unit of observation is county. The sample period is 2005-2018. The sample includes 97 counties. Treat is a dummy variable indicating counties with the survey teams deployed in 2005, and is instrumented by the randomly assigned rural survey teams in 1984. Post is a dummy variable indicating years after the reform in 2009. Panel A measures policies using a simple key words frequency method. Panel B and C measure policies using supervised machine learning methods (Random Forest and Support Vector Machine). The detailed procedures for constructing these measures are described in Appendix C. To alleviate multiple hypothesis testing issues, column (1) reports estimates using an standardized index by summarizing the four policy measures following Kling, Liebman and Katz (2007). The effective F-statistics are constructed following Montiel Olea and Pflueger (2013). Standard errors clustered at the county level are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A14: Effect on bank credit - IV estimates

Dep. var. is IHS of:	(1) Total amount of loans	(2) Loans to small firms	(3) # Firms granted loans	(4) # Branches granting loans
Panel A: Difference-in-differences				
Treat × Post	0.190	0.270*	0.138	0.090*
	(0.158)	(0.150)	(0.095)	(0.046)
Effective F-statistic	2,058	2,058	2,058	2,058
Panel B: Difference-in-difference-in-differences				
Treat × Government control × Post	0.273**	0.191*	0.055	0.118***
	(0.118)	(0.103)	(0.065)	(0.037)
Treat × Post	0.223	0.309**	0.160*	0.073
	(0.155)	(0.152)	(0.096)	(0.046)
Government control × Post	0.191**	0.172**	0.141***	0.059**
	(0.086)	(0.076)	(0.046)	(0.028)
Effective F-statistic	1,022	1,022	1,022	1,022
County FE	X	X	X	X
Year FE	X	X	X	X
Province FE × Post	X	X	X	X
County controls × Post	X	X	X	X
Observations	8,922	8,922	8,922	8,922

Notes: This table shows the IV estimates on the effect of the reform on bank credit. The unit of observation is county. The sample period is 2006-2011. Treat is a dummy variable indicating counties with the survey teams deployed in 2005, and is instrumented by the randomly assigned rural survey teams in 1984. Post is a dummy variable indicating years after the reform in 2009. Government control denotes the standardized share of pre-reform loans from City Commercial Banks (CCBs) in a county, whose controlling shareholders are local governments. The dependent variables are highly skewed and are thus transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails (Bellemare and Wichman, 2020). Panel A adopts a difference-in-differences estimation as usual. Panel B adopts a difference-in-difference-in-differences estimation utilizing differential control of banks by local governments across counties. The estimation equation is based on equation (12). The effective F-statistics are constructed following Montiel Olea and Pflueger (2013). Standard errors clustered at the county level are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A15: Effect on bank credit - untransformed variables

Dep. var.:	(1) Total amount of loans	(2) Loans to small firms	(3) # Firms granted loans	(4) # Branches granting loans
Panel A: Difference-in-differences				
Treat × Post	24,960.061**	8,911.357**	31.685	1.012*
	(10,783.378)	(3,741.146)	(31.821)	(0.558)
R-squared	0.881	0.765	0.611	0.859
Panel B: Difference-in-difference-in-differences				
Treat × Government control × Post	84,666.050*	19,681.718	-608.650	1.311*
	(50,782.638)	(13,658.901)	(540.779)	(0.691)
Treat × Post	129,932.700***	25,907.130***	-376.102	-0.248
	(33,496.942)	(9,745.098)	(323.677)	(0.585)
Government control × Post	67,981.316***	18,991.578***	683.584	0.655*
	(23,239.581)	(6,560.406)	(535.012)	(0.358)
R-squared	0.877	0.747	0.188	0.891
County controls × Post	X	X	X	X
County FE	X	X	X	X
Year FE	X	X	X	X
Observations	8,922	8,922	8,922	8,922
Mean dep. var.	368,082	78,087	525	41.11

Notes: This table shows the effect of the reform on bank credit, without inverse hyperbolic sine transformation of the outcomes to alleviate concerns raised by Chen and Roth (2023). The unit of observation is county. The sample period is 2006-2011. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Government control denotes the standardized share of pre-reform loans from City Commercial Banks (CCBs) in a county, whose controlling shareholders are local governments. Panel A adopts a difference-in-differences estimation as usual. Panel B adopts a difference-in-difference-in-differences estimation utilizing differential control of banks by local governments across counties. The estimation equation is equation (12). Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A16: Effect on firm entry - untransformed variables

	(1)	(2)	(3)	(4)	(5)
Dep. var.:	# Firm registrations				
Firm type:	All	Private	SOEs	Foreign	Collective
Treat × Post	23.635**	23.173**	0.743***	0.144	-0.426
	(10.976)	(10.933)	(0.264)	(0.213)	(0.305)
County FE	X	X	X	X	X
Year FE	X	X	X	X	X
County controls × Post	X	X	X	X	X
Cluster level	County	County	County	County	County
Observations	4,494	4,494	4,494	4,494	4,494
R-squared	0.834	0.833	0.496	0.811	0.505
Mean dep. var.	305.6	298.1	3.358	2.305	1.831

Notes: This table shows the effect of the reform on firm entry, without inverse hyperbolic sine transformation of the outcomes to alleviate concerns raised by Chen and Roth (2023). The unit of observation is county. The sample includes the years 2005, 2010, and 2015. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A17: Effect on firm entry - IV estimates

	(1)	(2)	(3)	(4)	(5)
Dep. var.:	IHS(# Firm registrations)				
Firm type:	All	Private	SOEs	Foreign	Collective
Treat × Post	0.073**	0.069**	0.197***	0.022	-0.083
	(0.034)	(0.035)	(0.073)	(0.064)	(0.078)
County FE	X	X	X	X	X
Year FE	X	X	X	X	X
Province FE × Post	X	X	X	X	X
County controls × Post	X	X	X	X	X
Cluster level	County	County	County	County	County
Observations	4,494	4,494	4,494	4,494	4,494
R-squared	0.057	0.068	0.019	0.017	0.033
Effective F-statistic	2,260	2,260	2,260	2,260	2,260

Notes: This table shows the IV estimates on the effect of the reform on firm entry. The unit of observation is county. The sample includes the years 2005, 2010, and 2015. Treat is a dummy variable indicating counties with the survey teams deployed in 2005, and is instrumented by the randomly assigned rural survey teams in 1984. Post is a dummy variable indicating years after the reform in 2009. The effective F-statistics are constructed following Montiel Olea and Pflueger (2013). Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A18: Effect on firm entry - RD robustness

	(1)	(2)	(3)	(4)	(5)
Dep. var.:			# Firm registrations		
Firm type:	All	Private	SOEs	Foreign	Collective
Panel A: IK optimal bandwidth					
Treat	7.515*	7.953**	0.045	0.055*	-0.049
	(3.954)	(3.830)	(0.027)	(0.033)	(0.040)
RD kernel	Uniform	Uniform	Uniform	Uniform	Uniform
RD polynomial	1	1	1	1	1
Rd bandwidth	IK	IK	IK	IK	IK
Observations	8,716	8,954	10,480	11,092	6,938
R-squared	0.354	0.349	0.205	0.427	0.254
Mean dep. var.	18.49	18.24	0.11	0.08	0.05
Bandwidth	3.97	4.07	4.74	5.05	3.30
Panel B: Quadratic RD polynomial					
Treat	8.448*	8.225*	0.084**	0.146	-0.090**
	(4.446)	(4.432)	(0.034)	(0.123)	(0.038)
RD kernel	Uniform	Uniform	Uniform	Uniform	Uniform
RD polynomial	2	2	2	2	2
Rd bandwidth	CCT	CCT	CCT	CCT	CCT
Observations	13,800	13,686	15,096	16,966	16,698
R-squared	0.266	0.266	0.164	0.501	0.212
Mean dep. var.	20.15	19.88	0.14	0.11	0.09
Bandwidth	6.71	6.62	7.87	10.06	9.71
Panel C: Triangular RD polynomial					
Treat	8.897**	8.877**	0.042*	0.059**	-0.073**
	(3.660)	(3.661)	(0.024)	(0.029)	(0.031)
RD kernel	Triangular	Triangular	Triangular	Triangular	Triangular
RD polynomial	1	1	1	1	1
Rd bandwidth	CCT	CCT	CCT	CCT	CCT
Observations	11,126	10,936	12,548	14,820	12,408
R-squared	0.341	0.341	0.218	0.375	0.214
Mean dep. var.	18.78	18.26	0.12	0.09	0.08
Bandwidth	5.07	4.98	5.83	7.60	5.74
County border FE	X	X	X	X	X
Cluster level	County border	County border	County border	County border	County border

Notes: The table shows the baseline RD estimates are robust to alternative bandwidth (panel A), quadratic RD polynomial (panel B), and triangular kernel (panel C). The unit of observation is town. The sample includes the post-reform years 2010 and 2015. The regression discontinuity estimation is conducted at the town level along county borders with different treatment statuses on each side. The bandwidths in panel A are selected following Imbens and Kalyanaraman (2012) while the bandwidths in panel B and C are selected following Calonico, Cattaneo and Titiunik (2014). Treat is a dummy variable indicating towns located in counties with the survey teams deployed in 2005. Standard errors clustered at the county border level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A19: Effect on citizen attitudes - Alternative cohort groups

Dep. var.:	(1) Trust in local officials	(2) Trust in most people	(3) Eval. of govt. performance	(4) Eval. of others' health
Treat × 1990s cohort	0.308**	-0.008	0.060***	0.000
	(0.136)	(0.025)	(0.018)	(0.034)
County FE	X	X	X	X
Cohort FE	X	X	X	X
County controls × Cohort FE	X	X	X	X
Survey wave FE	X	X	X	X
Citizen controls	X	X	X	X
Cluster level	County	County	County	County
Observations	41,385	41,476	41,665	29,215
R-squared	0.067	0.066	0.057	0.084
Mean dep. var.	5.16	0.54	0.80	0.25

Notes: This table shows the effect of the reform on citizens' attitudes towards local governments, using an alternative definition of affected and unaffected cohorts. The unit of observation is citizen (three waves of survey data from the China Family Panel Studies, 2012, 2014 and 2016). Treat is a dummy variable indicating counties with the survey teams deployed in 2005. 1990s cohort is a dummy variable indicating citizens born in the 1990s. The estimation equation is a cohort difference-in-differences specification that utilizes two sources of variation: (1) treatment counties versus control counties; (2) affected cohorts versus unaffected cohorts within the same county. The latter source of variation is built on insights in the psychology and political science literature that citizens' political attitudes are most permeable during teenage years and keep stable since one's 30s (Wolfinger and Rosenstone, 1980; Krosnick and Alwin, 1989). As such, the 1990s cohort are defined as the affected group as they would be younger than 30 during the reform period, and thus, their political attitudes would be most permeable. Those born in or before the 1970s are the unaffected cohorts, as they would be older than 30 in the reform period and thus be unaffected by the reform. Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A20: Effect on citizen attitudes - Controlling for media access

Dep. var.:	(1) Trust in local officials	(2) Trust in most people	(3) Eval. of govt. performance	(4) Eval. of others' health
Treat × 1980s cohort	-0.052	0.010	0.027	0.010
	(0.133)	(0.021)	(0.018)	(0.034)
Treat × 1990s cohort	0.447***	-0.001	0.054**	0.004
	(0.158)	(0.028)	(0.020)	(0.033)
County FE	X	X	X	X
Cohort FE	X	X	X	X
County controls × Cohort FE	X	X	X	X
Survey wave FE	X	X	X	X
Citizen controls	X	X	X	X
Citizen media access	X	X	X	X
Cluster level	County	County	County	County
Observations	17,163	17,166	17,194	11,999
R-squared	0.070	0.086	0.061	0.109
Mean dep. var.	4.84	0.57	0.82	0.25

Notes: This table shows the effect of the reform on citizens' attitudes towards local governments, controlling for citizens' media access. The unit of observation is citizen (three waves of survey data from the China Family Panel Studies, 2012, 2014 and 2016). Treat is a dummy variable indicating counties with the survey teams deployed in 2005. 1980s cohort and 1990s cohort are dummy variables indicating citizens born in the 1980s and the 1990s, respectively. The estimation equation is a cohort difference-in-differences specification (equation 14) that utilizes two sources of variation: (1) treatment counties versus control counties; (2) most affected cohorts versus less unaffected cohorts within the same county. The latter source of variation is built on insights in the psychology and political science literature that citizens' political attitudes are most permeable during teenage years and keep stable since one's 30s (Wolfinger and Rosenstone, 1980; Krosnick and Alwin, 1989). As such, the 1990s cohort are defined as the most affected group as they would be younger than 30 during the reform period, and thus, their political attitudes would be most permeable. The 1980s cohort are defined as the less affected group because some of them would be older than 30 during the reform period. The omitted group is those born in the 1970s, as they would be older than 30 in the reform period and thus be unaffected by the reform. To account for citizens' media access, I include three variables constructed from the survey regarding media access: (1) the number of days political news was accessed via television in the last week; (2) the number of days political news was accessed via Internet in the last week; (3) whether you have posted comments related to political issues and major national events on Internet in the past 12 months. Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A21: Effect on citizen attitudes - IV estimates

Dep. var.:	(1) Trust in local officials	(2) Trust in most people	(3) Eval. of govt. performance	(4) Eval. of others' health
Treat × 1980s cohort	-0.027	-0.013	-0.001	-0.059
	(0.160)	(0.024)	(0.021)	(0.051)
Treat × 1990s cohort	0.532***	0.007	0.044*	-0.002
	(0.183)	(0.035)	(0.026)	(0.041)
County FE	X	X	X	X
Cohort FE	X	X	X	X
Province FE × Cohort FE	X	X	X	X
County controls x Cohort FE	X	X	X	X
Survey wave FE	X	X	X	X
Citizen controls	X	X	X	X
Cluster level	County	County	County	County
Observations	17,163	17,166	17,194	11,999
Effective F-statistic	44.46	44.24	44.43	45.49
Mean dep. var.	4.84	0.57	0.82	0.25

Notes: This table shows the IV estimates on the effect of the reform on citizens' attitudes towards local governments. The unit of observation is citizen (three waves of survey data from the China Family Panel Studies, 2012, 2014 and 2016). Treat is a dummy variable indicating counties with the survey teams deployed in 2005, and is instrumented by the randomly assigned rural survey teams in 1984. 1980s cohort and 1990s cohort are dummy variables indicating citizens born in the 1980s and the 1990s, respectively. The estimation equation is a cohort difference-in-differences specification (equation 14) that utilizes two sources of variation: (1) treatment counties versus control counties; (2) most affected cohorts versus less unaffected cohorts within the same county. The latter source of variation is built on insights in the psychology and political science literature that citizens' political attitudes are most permeable during teenage years and keep stable since one's 30s (Wolfinger and Rosenstone, 1980; Krosnick and Alwin, 1989). As such, the 1990s cohort are defined as the most affected group as they would be younger than 30 during the reform period, and thus, their political attitudes would be most permeable. The 1980s cohort are defined as the less affected group because some of them would be older than 30 during the reform period. The omitted group is those born in the 1970s, as they would be older than 30 in the reform period and thus be unaffected by the reform. The effective F-statistics are constructed following Montiel Olea and Pflueger (2013). Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A22: Correcting for multiple hypothesis testing

Dep. var.:	(1) Standardized index of govt. policies	(2) IHS(Total loans)	(3) IHS(Firm entry)	(4) Trust in local officials	(5) Eval. of govt. performance
Treat × Post	0.474***	0.174	0.046*		
	(0.167)	(0.123)	(0.026)		
Raw p-value	[0.006]	[0.158]	[0.078]		
Sharpened q-value	{0.022}	{0.118}	{0.085}		
Treat × 1980s cohort				-0.051	0.026
				(0.129)	(0.018)
Raw p-value				[0.695]	[0.148]
Sharpened q-value				{0.226}	{0.118}
Treat × 1990s cohort				0.452***	0.052**
				(0.154)	(0.020)
				[0.004]	[0.012]
				{0.022}	{0.022}

Notes: This table adjusts for multiple hypothesis testing for the estimates on several aggregate outcomes used to measure local officials' development effort. To this end, I reproduce the results for several aggregate outcomes in this table (see the description of these results in Section 5), and report the sharpened q-values proposed by Anderson (2008) in braces to correct for multiple hypothesis testing. The raw p-values are reported in brackets for comparison.

Table A23: Fiscal transfers

	(1)	(2)	(3)
Dep. var.:		IHS(Fiscal Transfers)	
Transfer type:	All transfers	General transfers	Specialized transfer
Treat × Post2005	0.059	0.025	0.070
	(0.151)	(0.159)	(0.145)
County FE	X	X	X
Year FE	X	X	X
Cluster level	County	County	County
Observations	12,453	12,453	12,453
R-squared	0.529	0.513	0.538

Notes: This table examines whether the launch of the survey teams would increase fiscal transfers from upper-level governments. The unit of observation is county. The sample is from 2000 to 2007. Treat is a dummy variable indicating counties with the survey teams deployed in 2005. Post2005 is a dummy variable for years after 2005. The dependent variables are transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails of the skewed outcome distributions (Bellemare and Wichman, 2020). Standard errors clustered at the county level are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A24: Effect on policy diffusion

	(1)	(2)	(3)
Dep. var.:	Similarity of government work reports between county i and county j		
County pair type:	All types	Within province	Across province
Treat × Post	-0.000	-0.001	-0.002*
	(0.001)	(0.001)	(0.001)
County (i) × Year FE	X	X	X
County (j) × Year FE	X	X	X
County pair FE	X	X	X
Cluster level	two-way (i,j)	two-way (i,j)	two-way (i,j)
Observations	49,506	31,845	17,561
R-squared	0.977	0.979	0.987
Mean dep. var.	0.418	0.403	0.447

Notes: This table shows the effect of the reform on policy diffusion across counties. The unit of observation is county pair ij (with ij equivalent to ji). Treat is a dummy variable indicating county pairs in which both counties had a survey team deployed in 2005. Post is a dummy variable indicating years after the reform in 2009. The dependent variable is the pairwise textual similarity of government work reports, which is calculated following Kelly et al. (2021) and described in Appendix C. Column (1) considers all county pairs. Column (2) considers county pairs within the same province. Column (3) considers county pairs spanning different provinces. Standard errors that are two-way clustered by both counties in a pair are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A25: Effect on government policies - social welfare and environmental protection

	(1)	(2)	(3)	(4)	(5)	(6)
Method:	Key words frequency		Random Forest score		SVM score	
Dep. var.:	Social welfare	Enviro. protection	Social welfare	Enviro. protection	Social welfare	Enviro. protection
Panel A: OLS estimates						
Treat × Post	0.000	-0.000	0.002	-0.001	0.000	-0.002
	(0.000)	(0.000)	(0.002)	(0.001)	(0.002)	(0.002)
R-squared	0.378	0.381	0.304	0.369	0.334	0.397
Mean dep. var.	0.002	0.001	0.015	0.018	0.018	0.022
Panel B: IV estimates						
Treat × Post	0.000	-0.000*	0.000	-0.003	-0.000	-0.003
	(0.000)	(0.000)	(0.002)	(0.002)	(0.002)	(0.003)
Effective F-statistic	74.06	74.06	74.06	74.06	74.06	74.06
Mean dep. var.	0.002	0.001	0.015	0.018	0.018	0.022
County FE	X	X	X	X	X	X
Year FE	X	X	X	X	X	X
County controls × Post	X	X	X	X	X	X
Cluster level	County	County	County	County	County	County
Observations	883	883	883	883	883	883

Notes: This table shows the effect of the reform on government policies on social welfare and environmental protection. The unit of observation is county. The sample period is 2005-2018. The sample includes 97 counties. Treat is a dummy variable indicating counties with the survey teams deployed in 2005, and is instrumented by the randomly assigned rural survey teams in 1984 in panel B. Post is a dummy variable indicating years after the reform in 2009. Columns (1)-(2) measure policies using a simple key words frequency method. Columns (3)-(6) measure policies using supervised machine learning methods (Random Forest and Support Vector Machine). The detailed procedures for constructing these measures are described in Appendix C. The effective F-statistics in panel B are constructed following Montiel Olea and Pflueger (2013). Standard errors clustered at the county level are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A26: Effect on local government debt

	(1)	(2)	(3)
Dep. var.:		IHS(Bond issuance)	
Treat	0.101	-0.019	-0.173
	(0.278)	(0.258)	(0.254)
County controls		X	X
Province FE			X
Observations	1,752	1,498	1,498
R-squared	0.002	0.244	0.183
Effective F-statistic	1,942	2,041	2,214

Notes: This table shows the IV estimates on effect of the reform on local government debt. The unit of observation is county. Treat is a dummy variable indicating counties with a survey team deployed in 2005, and is instrumented using the randomly assigned rural survey teams in 1984 as the instrument. The dependent variable is the total amount of bond issuance by local government financing vehicles (LGFVs), which serves as a proxy for local government debt. The earliest issuance at the county level was in 2009. The dependent variables are highly skewed and are thus transformed by inverse hyperbolic sine (IHS) to reduce influences from the tails (Bellemare and Wichman, 2020). The effective F-statistics are constructed following Montiel Olea and Pflueger (2013). Robust standard errors are reported in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A27: Effect on corruption convictions

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.:	# Corruption convictions					
Time period:	2012-2016			2015-2016		
Type of corruption:	All types	Bribery & Appropriation	Other types	All types	Bribery & Appropriation	Other types
Treat	0.343	0.242	0.101	0.188	0.072	0.117
	(0.301)	(0.231)	(0.142)	(0.221)	(0.165)	(0.120)
# Anti-corruption inspections	-0.067	-0.060	-0.007	-0.029	-0.037	0.008
	(0.060)	(0.045)	(0.029)	(0.044)	(0.032)	(0.025)
County controls	X	X	X	X	X	X
Province FE	X	X	X	X	X	X
Cluster level	County	County	County	County	County	County
Observations	1,498	1,498	1,498	1,498	1,498	1,498
R-squared	0.099	0.109	0.027	0.068	0.073	0.020
Effective F-statistic	2,204	2,204	2,204	2,204	2,204	2,204
Mean dep. var.	4.648	3.303	1.344	2.933	1.997	0.935

Notes: This table shows the IV estimates on the effect of the reform on corruption convictions. The unit of observation is county. Treat is a dummy variable indicating counties with the survey teams deployed in 2005, and is instrumented using the randomly assigned rural survey teams in 1984 as the instrument.. The dependent variable denotes the number of corruption convictions by type and period. This data contains 10,797 corruption convictions from 2005 to 2016, with a vast majority (10,788) happening after 2012 when China's anti-corruption campaigns began. The few convictions (9) before 2012 were likely caused by the lack of enforcement instead of less corruption, and are dropped from my analysis. The number of anti-corruption inspections are included to address concerns about differential enforcement. The effective F-statistics are constructed following Montiel Olea and Pflueger (2013). Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A28: Effect on corruption perception

	(1)	(2)
	OLS	IV
Dev. var.:	Corruption perception	
Treat × 1980s cohort	-0.201	-0.255
	(0.152)	(0.191)
Treat × 1990s cohort	-0.116	0.035
	(0.202)	(0.241)
County FE	X	X
Cohort FE	X	X
Province FE × Cohort FE		X
County controls × Cohort FE	X	X
Survey wave FE	X	X
Citizen controls	X	X
Cluster level	County	County
Observations	10,747	10,747
R-squared	0.103	0.038
Effective F-statistic		42.18
Mean dep. var.	4.85	4.85

Notes: This table shows the effect of the reform on citizens' perceived corruption about the government. The unit of observation is citizen (two waves of survey data from the China Family Panel Studies: 2014 and 2016). Treat is a dummy variable indicating counties with the survey teams deployed in 2005, and is instrumented by the randomly assigned rural survey teams in 1984 in column (2). 1980s cohort and 1990s cohort are dummy variables indicating citizens born in the 1980s and the 1990s, respectively. The estimation equation is a cohort difference-in-differences specification (equation 14) that utilizes two sources of variation: (1) treatment counties versus control counties; (2) most affected cohorts versus less unaffected cohorts within the same county. The latter source of variation is built on insights in the psychology and political science literature that citizens' political attitudes are most permeable during teenage years and keep stable since one's 30s (Wolfinger and Rosenstone, 1980; Krosnick and Alwin, 1989). As such, the 1990s cohort are defined as the most affected group as they would be younger than 30 during the reform period, and thus, their political attitudes would be most permeable. The 1980s cohort are defined as the less affected group because some of them would be older than 30 during the reform period. The omitted group is those born in the 1970s, as they would be older than 30 in the reform period and thus be unaffected by the reform. The effective F-statistics are constructed following Montiel Olea and Pflueger (2013). Standard errors used to construct the 90% confidence intervals, denoted by the spikes, are clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Chapter 2

Incentives, Penalties, and Rural Air Pollution: Evidence from Satellite Data

Yongwei Nian *

Abstract

I test the role of economic incentives and command and control in reducing agricultural fires, a major source of air pollution in most rural regions across the world caused by burning crop residues after harvest. To tackle data shortage, I use high-resolution satellite data to construct a fine measure of agricultural fires as well as other geographic characteristics at 1 km $\times$ 1 km resolution for China. Using the staggered arrival of biomass power plants, which purchase crop residues as production inputs from nearby areas, as a shock to economic incentives, I find a more than 30% drop in agricultural fires in the vicinity of a plant after its opening relative to areas farther away. Such drop cannot be explained by structural transformation, migration, or enhanced regulation near the plant, and is consistent with an incentive-based explanation. I then examine the effectiveness of a command and control policy that bans agricultural fires within 15 km of airports. Using a spatial regression discontinuity design, I find no evidence that such policy works.

JEL classification: Q53, Q10, O18, P26
Keywords: Agricultural fires; Environmental policies; China

1 Introduction

Burning crop residues after harvest is a long-standing practice across the world, especially in developing countries.[1] It is perceived as a convenient way to prepare land for subsequent cultivation. At the same time, the open burning releases a number of pollutants, which could travel a long distance and generate a wide range of harms to human health (Rangel and Vogl, 2019; Graff Zivin et al., 2020; He, Liu and Zhou, 2020). Despite such negative externalities, how to effectively reduce agricultural fires is largely unexplored.

In practice, there are two widely used instruments to fight pollution: one is based on economic incentives and the other is based on command and control (Oates, Portney and McGartland, 1989).[2] For the former, a typical example is to provide subsidies to polluters for pollution abatement; for the latter, a typical example is to punish polluters through regulations and laws. In the context of fighting agricultural fires, which policy could be effective is, however, unclear for two reasons. The first reason is conceptual. Subsidies to farmers could be lower than their opportunity cost of not burning, either due to weak capacity for taxation or corruption (Greenstone and Jack, 2015). Even if subsidies are large enough, the effects are still indeterminate, since farmers could use the subsidies to expand farming, which may lead to more agricultural fires. The effectiveness of command and control also hinges on a number of factors such as the dissemination of the laws among farmers, the capacity to detect and punish burning, and the rule-abiding norms. In sum, disentangling the effects of these policies entails empirical investigation.

The second reason is that empirical tests typically suffer from at least three challenges. First, there exists an inherent selection bias associated with the rollout of a policy. For example, a policy may be launched in certain area following an outburst of agricultural fires. Then the subsequent reduction in agricultural fires could reflect reversion to the mean. Second, many policies have no clear or very aggregated target areas such as cities or provinces, making the

[1] A general description of agricultural fires worldwide can be found in a report by World Bank at https://openknowledge.worldbank.org/handle/10986/29504.

[2] In terms of curbing fires, technological innovation, such as cutting crop residues through specialized harvesters and returning them into the soil, could be a third alternative. However, this would induce a nontrivial cost. Due to the lack of economies of scale in agriculture and farmers' low marginal willingness to pay for clean air, such technology and related machines are not well developed and generalized in many developing countries.

results easily contaminated by other concurrent policies or omitted variables. Third, existing measures of fires, which are usually extracted directly from satellite observations, do not differentiate between fires from agricultural activities and from non-agricultural activities. Therefore, studies based on such measures may suffer from measurement errors.

I address these challenges in several steps. First, I use two unique quasi-experiments to control for selection bias. The first is the staggered adoption of biomass power plants in China that purchase crop residues from farmers as inputs, creating an economic incentive for farmers to reduce burning. Crucially, the main purpose of these plants is to expand energy sources and reduce the reliance on oil, instead of curbing agricultural fires. This experiment enables me to adopt a spatial difference-in-differences (DID) strategy to test the effects of economic incentives on agricultural fires. The second is the drawing of no burning zones, which target solely agricultural fires within 15 km of airports in China. The zone boundaries are purely artificially drawn and do not overlap with administrative boundaries or other policy thresholds. This experiment enables me to conduct a spatial regression discontinuity (RD) design to test the effects of command and control on agricultural fires. Second, I rely on micro data at the 1 km $\times$ 1 km resolution and include granular fixed effects to control for potential confounding factors such as local economic conditions. Third, I identify agricultural fires by matching satellite fire observations with detailed land use data to tackle the measurement issue. The land use data enable me to differentiate between agricultural fires and other fires.

I justify the validity of the DID strategy in various ways. First, I compare areas very close to a plant with areas that are a bit farther away. Due to spatial adjacency, these areas should be somewhat comparable ex ante, alleviating the concern that areas with and without a plant may be quite different in unobservable aspects. Specifically, I define two treatment groups based on qualitative and quantitative evidence: the most treated group consisting of areas within 5 km of a plant, and the less treated group consisting of areas within 5-10 km of a plant. The control group includes areas within 10-15 km of a plant. Second, I include highly disaggregated fixed effects, such as grid cell, plant $\times$ year, and city $\times$ year fixed effects, to control for a number of confounding factors. I also include a set of flexible trends specific to each grid cell to rule out the possibility that the DID estimates are just reflecting differential trends across areas. Third, I use an event study specification to show that there is no diverging pre-trends between treatment groups and control groups. Fourth, I conduct two placebo tests using non-agricultural fires and randomly generated plant openings as placebos, respectively. Finally, I directly control for some factors that affect location selection and correct for selection on unobservables by assuming the selection on observables is informative about selection on unobservables.

Using data from 2001 to 2016 on agricultural fires and the openings of 190 biomass power plants, I find a significant reduction in agricultural fires for both the most treated group and the less treated group following the arrival of a plant. In my preferred specification that controls for grid cell fixed effects, plant $\times$ year fixed effects, and differential trends across grid cells, the drop in agricultural fires amounts to 57.7% and 30.9% of the sample mean for the most treated group and the less treated group, respectively. The results are robust to correcting for the bias associated with staggered treatment timing. I further verify such decreasing effect with distance using alternative specifications in which I replace the two binary treatment indicators with continuous distance or topography-based distance to a plant. These findings are consistent with the explanation that biomass power plants create economic incentives for farmers to reduce burning, and the economic incentives become weaker when it becomes more costly to transport the crop residues to a plant. I corroborate such explanation using qualitative evidence and also quantitative evidence that shows weaker effects in steeper, more rugged, and more fragmented areas, where it is more difficult to collect crop residues. I also show that farmers' income increases in response to the plant openings.

I conduct additional tests to rule out some alternative explanations. First, the findings are not caused by structural transformation induced by a plant. Following the arrival of a plant, I find no change in agricultural land or other dimensions of agricultural production such as crop types and agricultural productivity. I also find no change in the industrial sector, including the very small firms, as measured by the number and size of industrial firms. Second, the findings are not caused by migration due to fear of possible pollution from a plant. I focus on areas downwind of a plant that are perceived to be affected more by the plant and do not find a larger effect. I also show that the effects in areas with access to transport networks, where it is easier to move out, are similar to other areas. Third, the findings are not caused by enhanced inspection by local officials around a plant, as the arrival of a plant in an area could signal local officials' attention paid to that area. I do not find a larger effect in areas with greater penetration of local officials. I also focus on areas where local officials have stronger promotion incentives, as pollution control is considered in promotion. I do not find a larger effect.

For the analysis of no burning zones, I justify the validity of the RD strategy in the following ways. First, I check the relevant documents and show that the drawing of no burning zones is a uniform and top-down decision made by the central government, which is uncorrelated with local economic and environmental conditions. Second, I check the preexisting covariates and find no discernible jumps at the zone boundaries. Third, I test possible selective sorting of farmers, firms and airports. The results either suggest no selective sorting or are against the main findings (i.e., some types of sorting would lead to a significant reduction in fires, whereas I find a small and insignificant reduction). Therefore, the main findings are unlikely to be driven by various forms of sorting. Fourth, I conduct a difference-in-discontinuities analysis exploiting within grid cell variation to rule out confounding differences in areas across the zone boundaries. Finally, I complement the RD design with a DID strategy that considers areas far away from the boundaries to strengthen external validity.

Using data on 205 airports, I find little evidence that no burning zones could reduce agricultural fires. In the preferred specification that uses a linear RD polynomial in distance to the boundaries and controls for airport $\times$ boundary segment fixed effects, the estimated coefficients only account for 1% of the sample mean, and are statistically insignificant. The results are robust to alternative bandwidths, fixed effects, samples, polynomial orders, and inclusion of covariates. I also use a two-dimensional RD specification using the longitude and latitude of each grid cell as polynomials, and the results barely change. I next check whether such insignificant correlation between no burning zones and agricultural fires masks some heterogeneity, by conducting RD estimation separately for each year and each airport. The effects are still small and statistically insignificant for most years and airports.

I further test whether the effects are significant in areas that are supposed to have more stringent regulatory enforcement. I look at three types of areas. The first type is areas upwind of the airport. Due to safety concerns, local officials should pay greater attention to such areas.

The second type is areas in provincial cities. These cities are more developed and have higher political status. Hence, local officials should have stronger incentive to control pollution for both economic and political concerns. Third, areas where the city party secretary has stronger promotion incentive. As pollution control is considered as part of the criteria for promotion, regulatory enforcement should also be stronger in such areas. I still cannot find significant effects in all these three types of areas. In addition, I also check whether the zones shift the burning from daytime to nighttime, which may still be meaningful due to more flights in the daytime. I do not find any supporting evidence.

This paper is primarily related to a nascent literature showing the causal effects of environmental policies on agricultural fires, a typical type of pollution in the rural areas (He, Liu and Zhou, 2020). As shown by prior work, agricultural fires could deteriorate air quality (Guo, 2020), affect infant and adult health (Rangel and Vogl, 2019; He, Liu and Zhou, 2020), and harm cognitive performance (Graff Zivin et al., 2020). Moreover, agricultural fires could generate large spillover effects in the sense that the pollutants could travel a long distance of about several hundreds of kilometers (Guo, 2020). Apart from such direct effects on human health, smokes from intensive burning could also threaten traffic safety, creating indirect effects on human health (Sager, 2019). Given such substantial health costs, finding the most effective way to reduce agricultural fires has significant welfare implications. In a closely related study, He, Liu and Zhou (2020) show that straw-recycling subsidy at the provincial level can reduce agricultural fires. My paper differs in two aspects. First, the straw-recycling subsidy studied in He, Liu and Zhou (2020) is similar to a traditional conditional cash transfer, which may entail a nontrivial enforcement cost; the plants studied in my paper, however, work more like a market device to automatically incentivize farmers to reduce burning. Second, in addition to the incentive-based policy, I study another commonly used command and control policy, which is unexamined in He, Liu and Zhou (2020). Using micro data and hyperlocal variation in policy exposure, I am able to test the casual effects of these policies and show that in contrast to command and control, resorting to economic incentives could significantly reduce agricultural fires.

This paper is also broadly related to a growing literature studying the effectiveness of environmental policies in developing countries. A common finding in this strand of literature is that command and control policies in these countries tend to fail due to weak regulatory enforcement (Davis, 2008; Duflo et al., 2013, 2018; Oliva, 2015; Souza-Rodrigues, 2019). However, a few studies do find a significant effect of command and control policies. For example, Greenstone and Hanna (2014) demonstrate the effectiveness of the regulation in India targeting air quality. Harrison et al. (2015) focus on firm-level outcomes in India and find both the incentive-based policy and the command and control policy could work for pollution abatement. Some studies based on China also find significant reductions in pollution from command and control policies (Tanaka, 2015; Viard and Fu, 2015; He, Wang and Zhang, 2020). Apart from the mixed findings, a common feature of these studies is that they focus on pollution in the urban areas, which differ from rural areas in many aspects such as social norms, enforcement costs, and marginal willingness to pay for environmental quality. Therefore, whether the existing results apply to rural areas is unclear. In terms of the rural setting, my paper is closely related to Souza-Rodrigues (2019), who shows through a structural estimation that economic incentives, rather than command and control, can reduce deforestation in the Amazon. However, as deforestation mainly happens in areas with a shortage of agricultural land and weak state capacity, it is still unclear whether such results could apply to my setting.

The rest of this paper proceeds as follows. Section 2 discusses the institutional background. Section 3 describes the data and presents descriptive statistics. Section 4 and 5 examine the effects of biomass power plants and no burning zones on agricultural fires, respectively. Section 6 discusses some additional issues and Section 7 concludes.

2 Institutional Background

2.1 Agricultural Fires across the World and in China

Across various farming systems, burning is one of the most popular ways to remove the cumbersome crop residues after harvest. Farmers view burning as not only a convenient way to clean the land for subsequent cultivation but also a way to nourish soils and kill pests, although burning is demonstrated to be harmful to soil quality (Villar et al., 2004). The pollutants from intense burning in a short period could release various pollutants (Chen et al., 2017). The most typical pollutant is particulate matter (e.g., $PM_{2.5}$), which could stay in the air for a long time and cause respiratory and cardiovascular syndromes, among others (Dominici et al., 2006). Besides, the fine particles could also cause haze and reduce visibility (Cheng et al., 2014). Other pollutants include CO, VOCs, NOx, SO_2, and other compound toxic substances. In sum, burning crop residues not only affects human health but also generates environmental damages.

Despite the rapid economic growth in recent years, China remains a largely rural country and is currently the largest burner of crop residues in aggregate terms. There are several features specific to China that could explain the burning. First, there is a very short window for removing crop residues between one harvest and the next sowing, typically one or two weeks. Given such short period and lack of labor force due to migration of young people,[3] burning seems to be the most efficient and economic way to prepare for sowing. Second, crop residues were traditionally used as fuel for cooking and were collected and used gradually. The recent spread of natural gas among the rural areas makes the crop residues quite useless, leading to intense field burning in a short time. Third, in China each farmer only owns a small piece of land. The wide dispersion of farming makes it hard for local officials to monitor burning.

2.2 Biomass Power Plants

Biomass power plants use agricultural waste or other organic waste to generate energy, either directly through burning or indirectly through gasification. After the 1973 oil crisis, Denmark first proposed this technology as a safeguard against potential oil supply crisis, as it relied heavily on oil. China was quite late in adopting this technology and it was not until the early 2000s that biomass power plants started to appear. In the Renewable Energy Law enacted in 2005,[4] biomass energy was listed as a promising energy source. The National Development and Reform Commission then launched various policies to subsidize biomass power plants to boost their development. As said in the law, the main purpose of promoting this type of energy is to expand energy sources and decrease the reliance on fossil energy. Since then, biomass

[3]Migration here refers to short-term change of residence, as permanent or long-term change is hard for farmers due to the *hukou* system.

[4]See http://www.gov.cn/ziliao/flfg/2005-06/21/content_8275.htm for the relevant Chinese document.

power plants have grown quite rapidly.

Compared to open burning of crop residues, biomass power plants are much cleaner in the sense that they could use dedusting, desulfurization and denitrification equipment to absorb most of the pollutants. Apart from burning, these plants could also transform the residues into gas, which generates fewer pollutants after burning. In this paper, I focus on the plants that mainly use agricultural waste as inputs.[5] For these plants, they typically use crop residues from surrounding farms.[6] Although farmers are paid for the crop residues, they need to collect and package the residues from the farm, which is a laborious process. The price of the crop residues is quite low, making some farmers view it unworthy to collect and sell the crop residues to plants far away. That being said, the revenue from crop residues is still nontrivial and could be attractive to some farmers, as shown in Appendix F.1.

2.3 No Burning Zones

In China, airports are typically located in the suburbs and surrounded mostly by farms. The smoke from agricultural fires, which is much more intense and dangerous than naturally formed fog, could severely affect the take-off and landing of airplanes.[7] On May 13th, 1998, due to intense agricultural fires, Shuangliu Airport in southwest China was forced to close for three times in one day, leading to the delay, return, or diversion of dozens of planes.[8] In an effort to protect air quality and secure airplane safety, the Ministry of Environmental Protection proposed a ban in 1999 on burning crop residues within 15 km of airports. [9] To make the ban public, local officials were required to print out copies of the ban and disseminate them to farmers. Moreover, local officials should patrol these zones routinely to discipline or prosecute the violators.

Several features make such zones unique and suitable to address identification challenges.

[5] Another type of plant uses mainly organic waste from the urban areas as inputs, and is typically located in the urban areas.

[6] A typical biomass power plant usually has a storage centre that stores residues collected during the harvest season to smooth the supply of residues over the year. See http://www.nantong.gov.cn/jcms/jcms_files/jcms1/web14/site/attach/0/160429135950571.pdf.

[7] See http://news.sina.com.cn/c/2002-10-09/0853759792.html for a survey with officers from an airport in China who expressed concerns about the threats from agricultural fires.

[8] See https://cn.govopendata.com/renminribao/1998/5/19/4/#1120306.

[9] See http://www.mee.gov.cn/gkml/zj/wj/200910/t20091022_171920.htm.

First, they are formulated uniformly from the top and therefore are unlikely to be correlated with local economic and environmental conditions. Second, they are drawn artificially along the 15 km-radius circle surrounding each airport and do not overlap with other important boundaries, avoiding the compound treatments issue faced by many geographic regression discontinuity designs utilizing administrative boundaries (Keele and Titiunik, 2015). Third, due to the household registration system (*hukou*), it is quite hard for farmers to migrate, largely alleviating the selective sorting issue in many regression discontinuity designs (McCrary, 2008).

In addition to the no burning zones around airports, the 1999 regulation also banned fires in areas within 1 or 2 km of major transportation networks. Local governments were also entitled to set up additional zones. It would also be interesting to examine the impact of these additional zones. However, the zones along major transportation networks may be too narrow to fit a smooth function in RD design. Furthermore, areas closer to major transportation networks could differ from areas farther away ex ante, and such difference may be discontinuous at the zone boundaries given that the zones are too narrow. Finally, even if all factors are ex ante continuous, it is hard to separate the effect of the zones from other ex post changes created by proximity to transportation networks. For the local government-specified zones, they are also not utilized in my analysis for several reasons. First, there is little exact delineation of these zones. In very few cases where there is a well-specified zone, the border of such zone usually coincides with major roads or administrative boundaries, making it difficult to separate the role of no burning zones from the role of roads or other boundaries. Second, as these zones are mostly in the urban areas, there may not be enough agricultural land around the zone boundaries. Therefore, one may find null effects simply due to low statistical power. Third, as these zones are determined at the discretion of local governments, they have strong political incentives to place the zones in areas with lower burnings. In sum, it is difficult to draw credible causal inference from these additional zones.

3 Data and Summary Statistics

The data in this paper are aggregated at the 1 km $\times$ 1 km grid cell level. To this end, I create a fixed grid with cells of 1 km $\times$ 1 km and then match all the data to this grid. I use this level of aggregation for several reasons. First, both biomass power plants and no burning zones could affect areas spanning multiple administrative units, making it hard to determine the treatment unit and control unit. For example, if a plant is located near the border of a city, then it is unreasonable to assign only this city as the treatment city. Second, as the boundary of each no burning zone is a 15 km-radius circle around the airport, then using such grid cells allows me to generate enough observations around the treatment threshold to conduct RD estimation. Third, using such disaggregated analysis allows me to better measure the geographic characteristics and control for confounding factors using grid cell fixed effects. More importantly, I could also control for the selective arrival of a plant in an area by including plant $\times$ year and city $\times$ year fixed effects.

3.1 Main Data

Agricultural fires data Data on fires are generated by the NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the Terra and Aqua satellites, available from 2001 onwards. I use the data from 2001 to 2016. The two satellites pass over the equator for four times each day (10:30 am, 10:30 pm, 1:30 pm, and 1:30 am) and observe almost every area on the earth for at least 4 times due to overlapping orbits. The MODIS sensor detects fires using a contextual algorithm utilizing the strong emission of mid-infrared radiation from fires and reports the longitude and latitude of a fire. (Giglio et al., 2003). Depending on the observing conditions, the minimum detectable fire size ranges from 50 to 1000 m^2. As the average size of cropland in China is much larger (typically about several thousand m^2), the sensor should be able to detect most agricultural fires. As the focus of this paper is agricultural fires, I next match the fire data with land cover data from the Copernicus Land Monitoring Service (CLMS), to determine fires occurring in cropland. The CLMS classifies global land surface into several classes, including cropland and settlement land. The data are available annually from 1992 onwards. I define a fire as agricultural fire if the location of the fire is inside cropland.

As is common in most satellite observations data, the fire observations data are not immune to potential measurement errors. The first type of measurement error comes from varying observing conditions. That is, the detection could be affected by factors like cloud cover, land surface temperature, and amount of smoke. This type of error is plausibly random and therefore can be viewed as classical measurement error. The results are still unbiased, but the precision will be lower. The second type of measurement error comes from the limited number of observations each day, leading to an under-detection of fires. For example, farmers in certain area may choose to burn when the satellites are not passing over that area, either due to behavioral changes or fear of being detected by relevant authorities. However, as I include the plant$\times$year, city$\times$year, or airport$\times$year fixed effects, depending on the specifications, such under-detection should be largely controlled by these fixed effects.

Biomass power plants data Data on biomass power plants are obtained from a Chinese consulting firm. The firm collected the information on all the active plants in 2016 and verified the information through multiple ways including phone calls. The information includes plant name, detailed address, year of opening, and main business. I use the APIs from Baidu Maps, the Chinese equivalent of Google Maps, to geocode the plants to get the longitudes and latitudes.[10] There are mainly two types of plants: one uses crop residues as inputs and the other uses organic waste from the urban areas as inputs. I keep only those plants that use crop residues as inputs. There are 253 such plants. I validate such number using an official report and discuss the potential bias in Appendix E.1. These plants are basically located in the rural areas and on average over 70% of the surrounding land is cropland. To match the fire data that are available only from 2001, I drop plants established before 2002 to have at least one year of fire data before the opening of any plant. Similarly, I also drop plants established after 2015 to have at least one year of fire data after the opening of any plant.[11] Then I am left with 249 plants. Using each of the remaining plants as a center, I draw a circle with a 15 km radius. I drop those plants with overlapping circles, as in this case it would be difficult to define the treatment and control areas. This step excludes about 24% plants and I am left with 190 plants. Appendix E.3 shows that such sample selection has little impact on my findings. The spatial

[10] The longitudes and latitudes obtained directly from the *geocoding* API from Baidu Maps use the GCJ02 coordinate system, which is censored as it adds some offsets about several hundred meters to the widely used WGS84 coordinates. To match with other geographic data that use the WGS84 coordinates, I rely on the *coordination conversion* API to iteratively convert the GCJ02 coordinates into WGS84 coordinates. The procedure is described in detail in Appendix E.2.

[11] This step drops very few plants and results are similar if I do not drop these plants.

and temporal distribution of these plants are shown in Figure 2. The plants are concentrated in the major grain-producing provinces such as Shandong, Henan, and Heilongjiang, and increase sharply after the introduction of the Renewable Energy Law in 2005.

No burning zones data To construct the no burning zones, I first get the list of civic airports from the Great Circle Mapper.[12] For confidentiality reasons, data on military airports are not available. I use the 205 active civic airports at the end of 2016. For each airport, I have the longitude and latitude and the year of opening. I then draw a circle of 30 km radius around each airport. Because these 30 km-radius circles do not overlap with each other, all the airports are kept. All the areas inside the circle constitute the sample for RD design. The treatment threshold is a circle of 15 km radius.[13] Inside this circle, agricultural fires are banned, and outside, not. The spatial and temporal distribution of these airports are shown in Figure 2. As can be seen, some airports are located in provinces where agriculture is not well developed, such as Qinghai and Xizhang province in southwest China. Therefore, areas around these airports may not be suitable for studying agricultural fires. I address such concern by eliminating airports around which very little cropland exists in robustness checks.

3.2 Other Data

Below I list some data that are used either for mechanism tests or heterogeneity analysis in the main text. For some other data used in the appendix, I will describe them where they first appear.

Annual survey of industrial firms (ASIF) data I use the ASIF data from 1998 to 2007, collected by the National Bureau of Statistics of China, to construct measures reflecting the development of the industrial sector in each grid cell. The data include private industrial firms with annual sales over 5 million *yuan* (about 0.78 million USD) and all state-owned industri-

[12] See www.gcmap.com for details.

[13] When drawing the 15 km-radius circles, an airport is treated as a point using its location shown on the map, instead of a polygon. This is consistent with the procedure used by the Ministry of Environmental Protection of China to determine whether a fire is within the no burning zones (Fang, Zhang and Xu, 2006). To alleviate concerns that airports may expand and slightly change their locations on the map, I show in Appendix Table A23 that the results are robust when using a "donut hole" RD specification to exclude areas close to the 15 km-radius circles.

al enterprises (SOEs). The data also have detailed address information for each firm, which enables me to geocode the location. As the data contain some outliers, I follow standard procedures used in the literature to clean the data (Brandt, Van Biesebroeck and Zhang, 2012).

Geography and transportation data Data on land cover from the CLMS are used to construct the share of cropland and settlement land in each grid cell. Data on topography from the NASA's Shuttle Radar Topography Mission (SRTM) are used to construct the slope, elevation, and ruggedness of each grid cell, following the method used by Nunn and Puga (2012). Data on river networks from the National Geomatics Center of China are used to construct the number of rivers in each grid cell and the distance from the centroid of each grid cell to the nearest rivers. Data on transportation networks come from the Geographic Data Platform maintained by Peking University, including railways, national roads, and expressways. The data are available only for 2000.

Wind direction data I obtain wind direction data from the National Oceanic and Atmospheric Administration (NOAA) to gauge how frequent a grid cell is downwind of a plant (or upwind of an airport). The wind direction is observed at three-hour frequencies by ground-based weather stations and then averaged to the daily level using the vector decomposition method.[14] I assign the daily wind direction to a plant (or an airport) using the observations from the nearest weather station and drop plants (or airports) with no stations within 100 km. I then calculate the percentage of days in a year that a grid cell spends downwind of a plant (or upwind of an airport), where downwind (or upwind) is defined following the literature (Rangel and Vogl, 2019; Graff Zivin et al., 2020; He, Liu and Zhou, 2020) and also illustrated in Appendix Figure A1.

City leader data Data on the top 2 leaders in a city (the party secretary and mayor) collected from multiple sources are used to construct measures of political incentives. The sources include the provincial and city statistical yearbooks, Baidu Baike (China's Wikipedia) and Wikipedia. The data contain detailed information about the resume of each leader from 2001 to

[14] See http://www.webmet.com/met_monitoring/622.html.

2016, and are cross-validated whenever possible.

3.3 Descriptive Statistics

Appendix Table A1 contains descriptive statistics on the final sample used to analyze the effects of biomass power plants. That is, areas within 15 km of a plant. Appendix Table A2 contains descriptive statistics on the final sample used to analyze the effects of no burning zones based on a 4 km bandwidth, which is roughly the average bandwidth across different variables. Most of the time-varying variables are measured from 2001 to 2016. For grid cells around a plant, the average number of agricultural fires is 0.022, and the probability of observing at least one agricultural fire is 0.017. For grid cells around the no burning zone boundaries, the average number of agricultural fires is 0.011, and the probability of observing at least one agricultural fire is 0.009.

Figure 3 gives a description of the evolution of agricultural fires around biomass power plants and no burning zone boundaries. At the plant level, the number of agricultural fires is steadily increasing (Panel A). Panel C gives the average number of agricultural fires at the grid cell level, conditional on treatment status. Overall, there is an increasing trend across all three groups, although such trend seems to be reversed after 2013.[15] Between groups, there is no clear difference. Panel B and D show the evolution of agricultural fires around the no burning zone boundaries. The pattern is similar. Appendix Figure A2 gives a comparison of agricultural fires before and after a plant opening for different treatment groups. In Panel A, the most treated areas seem to have the largest number of agricultural fires before the opening (although the difference across groups is small), and such pattern is reversed after the opening. A somewhat similar pattern is also found in Panel B where I examine the probability of observing at least one agricultural fire. Such pattern sheds some light on the effects of biomass power plants on reducing burning, although it is just correlation.

[15]This reversal is consistent with China's significant efforts on combating air pollution since 2013 (Greenstone et al., 2021).

4 The Effects of Biomass Power Plants

4.1 Model Specification

I evaluate the effects of biomass power plants on agricultural fires using a geographic difference-in-differences (DID) strategy (Currie and Walker, 2011; Currie et al., 2015). As illustrated in Figure 1, the treatment and control groups are defined based on the proximity of each area to the nearest plant. The specification is as follows:

$$Y_{ijt} = \delta_i + \lambda_{jt} + \beta_1 Distance_{ij}^{<5km} \times Post_{jt} + \beta_2 Distance_{ij}^{5-10km} \times Post_{jt} + \varepsilon_{ijt} \tag{1}$$

where the unit of observation is 1 km $\times$ 1 km grid cells and the sample includes grid cells within 15 km of a plant. i, j, and t denotes grid cell, plant, and year, respectively. Y_{ijt} denotes measures of agricultural fires in grid cell i around plant j at year t, either in terms of incidence (presence of at least one fire) or in terms of intensity (number of fires). δ_i denotes grid cell fixed effects, controlling for time-invariant regional characteristics at the grid cell level such as topography. λ_{jt} denotes plant$\times$year fixed effects, controlling for all factors, including those time-varying, affecting all the grid cells around a plant such as weather conditions, plant performance, and local economic prosperity. $Distance_{ij}^{<5km}$ is an indicator equal to 1 if the distance from the centroid of grid cell i to plant j is smaller than 5 km and 0 otherwise; $Distance_{ij}^{5-10km}$ is an indicator equal to 1 if the distance from the centroid of grid cell i to plant j is between 5 km and 10 km and 0 otherwise. $Post_{jt}$ is an indicator equal to 1 for years after the opening of plant j and 0 otherwise. As I do not observe the exact month of opening, in Appendix E.4 I show that the results are robust when excluding years around a plant opening or assigning an expected treatment value between 0 and 1 to $Post_{jt}$ in the year of a plant opening following Currie et al. (2015).

The two indicators $Distance_{ij}^{<5km}$ and $Distance_{ij}^{5-10km}$ denote the most treated areas and the less treated areas, respectively. Choosing such areas as treatment groups is motivated by the fact that crop residues bear high transportation costs and low prices. Anecdotal evidence shows that farmers typically are only willing to send them to plants several kilometers away. Therefore, I choose the areas within 10 km of a plant as the treatment groups and the areas between 10 km and 15 km as the control group. I further bolster such choice by estimating a flexible variant of the baseline equation with many 5-km distance bins in Appendix A. To circumvent choosing treatment thresholds, I also use continuous treatment intensity with the two binary indicators replaced by continuous distance to a plant in robustness checks, and use least cost distance calculated based on topography in Appendix C.[16] The coefficients of interest are β_1 and β_2, which are standard DID estimators. Specifically, β_1 measures the change in agricultural fires in areas within 5 km of a plant after its opening, relative to areas 10 km away; β_2 measures the change in agricultural fires in areas within 5-10 km of a plant after its opening, relative to areas 10 km away.

In the baseline specification, I cluster the standard errors at the plant level to account for temporal correlation that is severe in DID estimation (Bertrand, Duflo and Mullainathan, 2004; Cameron and Miller, 2015), as well as spatial correlation across grid cells around the same plant. Given the nature of the data, especially the high spatial resolution, controlling for spatial correlation is important. I establish robustness of the results to adjusting the standard errors in two more conservative ways: (1) clustering at the 2° longitude $\times$ 2° latitude level;[17] (2) allowing for arbitrary correlation across spatially adjacent observations (Conley, 1999; Hsiang, 2010).

Recent econometric literature shows that the estimates from staggered DID designs are not easily interpretable (e.g., Callaway and Sant'Anna, 2020; de Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021). In Appendix B, I show that this is not a major concern in my setting; furthermore, the results are similar when I account for such issue using a more interpretable estimator following Callaway and Sant'Anna (2020).

4.2 Identification Concerns

The causal interpretation of the effects found in equation (1) relies on the common trends assumption. That is, agricultural fires in the treatment group and the control group should grow

[16]Additionally, one may also want to use travel distance from a grid cell to the plant. However, data on roads in the rural area are essentially unavailable. Moreover, such travel distance would be endogenous and may also capture the confounding effect of transportation connection.

[17] For anywhere on the earth, 1° latitude $\approx$ 111 km. For China, 1° longitude $\approx$ 70-100 km (larger when one moves southward).

in parallel in the absence of the plant. The major threat to this assumption is that the arrival of a plant in certain area is not random and may be correlated with certain local conditions. The grid cell fixed effects and plant $\times$ year fixed effects already take into consideration time-invariant factors at the grid cell level and time-varying factors around the plant that could drive the arrival of the plant. For most plants, both the treatment areas and control areas are also nested in the city where the plant is located. Therefore, the plant $\times$ year fixed effects could largely control for city-level factors that could lead to selection issue. I also control for both plant $\times$ year fixed effects and city $\times$ year fixed effects in robustness checks.

The remaining concerns are some time-varying factors at the grid cell level that may affect the arrival of a plant and also fires. One may argue that a plant could be placed in areas with growing agricultural fires. Conceptually, as the plants are mainly used to expand energy source instead of reducing agricultural fires, this issue is less severe. Nevertheless, I adopt a number of strategies to address such concerns.

First, I use a more demanding specification to allow for both differential linear trends in each grid cell and arbitrary trends in each grid cell depending on some geographic characteristics. The specification is as follows:

$$
\begin{aligned}
Y_{ijt} = \;& \delta_i + \lambda_{jt} + \delta_i \times Year_t + \mathbf{X}_i \times \theta_t + \beta_1 Distance_{ij}^{<5km} \times Post_{jt} \\
& + \beta_2 Distance_{ij}^{5-10km} \times Post_{jt} + \varepsilon_{ijt}
\end{aligned} \tag{2}
$$

where $Year_t$ denotes the number of years elapsed since the initial year. The term $\delta_i \times Year_t$ thus allows for a differential linear time trend in each grid cell. Appendix Table A4 shows similar results for higher-order time trends. $\mathbf{X}_i$ is a set of geographic characteristics including longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. By interacting these characteristics with year fixed effects θ_t, I allow the grid cells to evolve differentially depending on these characteristics. Conditional on the previous fixed effects and such flexible trends, the common-trends assumption is more likely to hold.

Second, I adopt an event study specification to directly test whether there exist any diverging pre-trends:

$$\begin{aligned}
Y_{ijt} = \; & \delta_i + \lambda_{jt} + \delta_i \times Year_t + \mathbf{X}_i \times \theta_t + \sum_k \beta_1^k Distance_{ij}^{<5km} \times Post_{jt}^k \\
& + \sum_k \beta_2^k Distance_{ij}^{5-10km} \times Post_{jt}^k + \varepsilon_{ijt} \\
& s.t. \;\; k \in \{-5, -4, -3, -2, 0, 1, 2, 3, 4, 5+\}
\end{aligned} \tag{3}$$

where I replace the $Post_{jt}$ indicator in the previous equation with a set of indicators denoting different years relative to the year of a plant opening. Specifically, $Post_{jt}^k$ equals to 1 if year t is the kth year relative to the opening of plant j. Negative ks denote years prior to the opening. $k = 5+$ denotes the fifth year after a plant opening and all subsequent years. $k = -5$ denotes the fifth year before a plant opening and all prior years. The year prior to a plant opening is omitted as the reference year. Hence, β_1^k measures the change in agricultural fires in areas within 5 km of a plant between the kth year relative to its opening and the year prior to its opening, compared to areas 10 km away; β_2^k measures the change in agricultural fires in areas within 5-10 km of a plant between the kth year relative to its opening and the year prior to its opening, compared to areas 10 km away. If the β^ks are small and statistically insignificant for $k \leq 0$, then the common trends assumption is likely to hold.

Third, I conduct two placebo tests as follows

$$\begin{aligned}
Y_{ijt}^{placebo} = \; & \delta_i + \lambda_{jt} + \delta_i \times Year_t + \mathbf{X}_i \times \theta_t + \beta_1 Distance_{ij}^{<5km} \times Post_{jt} \\
& + \beta_2 Distance_{ij}^{5-10km} \times Post_{jt} + \varepsilon_{ijt}
\end{aligned} \tag{4}$$

$$\begin{aligned}
Y_{ijt} = \; & \delta_i + \lambda_{jt} + \delta_i \times Year_t + \mathbf{X}_i \times \theta_t + \beta_1 Distance_{ij}^{<5km} \times Post_{jt}^{placebo} \\
& + \beta_2 Distance_{ij}^{5-10km} \times Post_{jt}^{placebo} + \varepsilon_{ijt}
\end{aligned} \tag{5}$$

First, in equation (4), I replace the dependent variables in previous equations with measures of non-agricultural fires $Y_{ijt}^{placebo}$, which are unlikely to be affected by biomass power plants. If the common trends assumption holds, then by construction one should not find a significant effect. Second, I randomly permutate the timing of the plant openings to generate a large number of placebo treatments reflected by the indicator $Post_{jt}^{placebo}$ in equation (5). Then if the results are simply driven by differential trends, then one should see similar effects among many placebo estimations.

I address a number of other selection issues in Appendix D. Specifically, I show that the results hold if I directly control for possible selection into areas with high amount of crop residues or concentrated land ownership that may correlate with fires; I also show that the results are not sensitive to correcting for selection on unobservables using the method proposed by Oster (2019), which assumes that selection on observables is informative about selection on unobservables.

4.3 Baseline Results

Average effect Table 1 shows the DID estimation of the effects of biomass power plants on agricultural fires. Columns (1)-(4) examine the number of fires. Column (1) only includes plant $\times$ year fixed effects. The coefficients on the two treatment indicators $<$ *5 km of the plant* and *5-10 km of the plant* are small and statistically insignificant, meaning that the average number of fires are similar across the most treated areas, the less treated areas and the control areas before the opening of a plant. The coefficients on $<$ *5 km of the plant* $\times$ *Post* and *5-10 km of the plant* $\times$ *Post* are -0.0054 and -0.0027, respectively, and both are statistically significant. Compared to the sample mean (0.022), these estimates suggest that the opening of a plant accounts for a 24.5% drop and a 12.3% drop in the number of fires in the most treated areas and the less treated areas, respectively.[18]

Column (2) adds grid cell fixed effects to control for time-invariant characteristics at a highly disaggregated level. The results show a larger and statistically significant reduction in the number of fires in both the most treated areas and the less treated areas. Column (3) adds a set of geographic controls interacted with time fixed effects to allow for different areas to evolve differently according to their geographic characteristics. The results barely change. Column

[18] There is a concern that the smaller coefficient for the less treated areas may be driven by the initial difference in fire activity across different areas. To alleviate this concern, I use the logarithm of one plus the number of fires as the outcome variable. Appendix Table A5 shows that the coefficient (and economic magnitude) for the less treated areas is still much smaller.

(4) further allows different areas to have different trends by including grid cell-specific linear trends. This specification together with that in column (3) could alleviate the concerns that the results are capturing the differential trends across areas. Compared to the sample mean, this most restrictive specification in column (4) implies a 57.7% drop and a 30.9% drop in the number of fires in the most treated areas and in the less treated areas, respectively. Columns (5)-(8) replace the number of fires with a binary indicator denoting observing at least one fire and get similar patterns to those in columns (1)-(4). Take the most restrictive specification in column (8) as an example, compared to the sample mean, the opening of a plant leads to a 37.4% drop and a 15.5% drop in the probability that fires appear in the most treated areas and the less treated areas, respectively.

Event study To check whether there exist differential pre-trends, I use the event study specification, equation (3), in which I interact the treatment indicators with a set of indicators denoting years before or after a plant opening. The year prior to the opening is omitted as the reference year. The results are visualized in Figure 4. Panel A and B only include grid cell fixed effects and plant $\times$ year fixed effects. The difference in agricultural fires between the treatment areas and the control areas are small and statistically insignificant in years before the opening, and are larger and statistically significant in years after the opening. Panel C and D further add geographic controls interacted with year fixed effects and grid cell-specific linear trends to control for possible differential trends across areas. The patterns are similar to those shown in Panel A and B. Overall, these patterns confirm that the most treated areas, the less treated areas, and the control areas share similar trends in agricultural fires before the plant opening, and that these areas start to diverge only after the plant opening.

4.4 Robustness

Continuous treatment intensity Results are similar when I replace the binary treatment indicators with continuous treatment intensity, which is the distance from each area to the plant. This specification could alleviate concerns that a plant could also affect areas 10 km away. I estimate the following equation:

$$Y_{ijt} = \delta_i + \lambda_{jt} + \delta_i \times Year_t + \mathbf{X}_i \times \theta_t + \beta Distance_{ij} \times Post_{jt} + \varepsilon_{ijt} \tag{6}$$

where $Distance_{ij}$ denotes the distance from the centroid of grid cell i to plant j. The coefficient of interest is β, which is expected to be positive. Columns (1)-(2) of Table 2 confirm such conjecture and show that moving 1 km closer to the plant reduces the number of fires by 0.0013 and the probability that fires occur by 0.0007, accounting for 6% and 4% of the sample means, respectively. Figure 5 further conducts an event study plot and shows no diverging pre-trends across different areas.

Spillover effects Results are robust to controlling for spatial spillovers. The baseline results focus on areas within 15 km of a plant. However, a plant may also collect crop residues farther away. If so, the treatment indicators could be contaminated in the sense that they could also capture the effects of other plants. As shown in columns (3)-(4) of Table 2, after adding the interaction between the distance from the centroid of each grid cell to the nearest neighbor plant and year fixed effects, the results barely change, suggesting that the spillover effects from other plants are negligible.

Confounding policies Results are not driven by confounding policies. In the sample, for about 48% of the plants, the 15 km buffers span more than one city. The treatment indicators could therefore capture differential city-level policies or economic conditions. To assess to what extent the results are driven by such factors, I add city $\times$ year fixed effects in columns (5)-(6) of Table 2 to flexibly control for all city-level factors. The results are not affected. I also add the interaction between the distance from the centroid of each grid cell to the city center and a post-2013 indicator to capture the large-scale air quality monitoring program launched in 2013. The program mainly aims at improving air quality in the urban areas. If the distance to a plant is correlated with the distance to the city center, then the results could be contaminated by such program. The results are unaffected as shown in columns (7)-(8) of Table 2. In Appendix G, I further examine five other major regulations in earlier years and the results stay virtually unchanged.

Persistence of fires Results are stable when I include a one-period lag of fires to control for the persistence of fires over time. Current burning could depend on past behavior as burning crop residues is a long-lasting tradition in China. Although including lagged dependent variable suffers from the well-known Nickell bias (Nickell, 1981), the results are almost unaffected as shown in columns (1)-(2) of Table 3.

Spatial correlation Results are not affected by controlling for spatial correlation. Given that both agricultural fires and biomass power plants are clustered in grain-producing areas, it is important to account for such spatial correlation. I first cluster the standard errors at the $2° \times 2°$ level in columns (3)-(4) of Table 3. As one degree of longitude or latitude equals about 111 km at the equator, the clustering unit is large enough to capture most of the correlation. The standard errors show little change. I then allow the standard errors to be arbitrarily correlated across spatially adjacent observations using the Conley standard errors. Specifically, I assume spatial correlation for observations within a 200 km radius and temporal correlation across 5 years. The standard errors in columns (5)-(6) of Table 3 barely change.

Placebo tests The first placebo test is to replace the dependent variable with non-agricultural fires in columns (7)-(8) of Table 3. As expected, the coefficients are insignificant. I also conduct an event study plot for non-agricultural fires. The results are visualized in Figure 6. Reassuringly, there is neither diverging trends before the opening nor a trend break after the opening. The second placebo test is to randomly permutate the timing of plant openings, holding the distribution of plant openings across years the same as the true distribution. I repeat this exercise 1000 times and plot the distribution of coefficient estimates in Appendix Figure A3. The distribution of placebo coefficients is not centered at zero since the placebo openings could to some extent coincide with the true openings. However, the true coefficients, denoted by the red lines, are far smaller than the placebo coefficients. These two placebo tests further justify the validity of the DID strategy.

4.5 Mechanisms

In this section, I first show that the previous findings are most likely consistent with an economic incentive-based explanation and then rule out several alternative explanations, namely, structural transformation, migration, and enhanced inspection.[19]

Economic Incentives A natural explanation of the findings is that the arrival of a plant increases the economic incentives to preserve crop residues, or equivalently increases the opportunity cost of burning crop residues. Specifically, before the arrival of a plant, the opportunity cost of burning is close to zero, as crop residues have basically no use for farmers. The opportunity cost increases sharply after the arrival of a plant since farmers are paid for their crop residues if sold to the plant. In Appendix F.1, I provide some qualitative evidence that corroborates such mechanism. In particular, I show that collecting crop residues are generally doable for farmers and the revenue from this also accounts for a non-trivial proportion of total value of agricultural production. Furthermore, farmers are also aware of the benefits of selling crop residues and view the revenue as attractive.

I test such mechanism more formally by looking at areas where economic incentives are lower. In such areas, one should see a smaller effect. In columns (1)-(4) of Table 4, I use the slope and ruggedness of each grid cell to measure economic incentives, as in areas with large slope or ruggedness it would be costly to collect crop residues. I do see a smaller effect in grid cells with larger slope or ruggedness, although it is only statistically significant for the most treated areas. In columns (5)-(6), I further measure the difficulty to collect crop residues using the number of fragments created by rivers in a grid cell. The intuition is simple: if a grid cell is cut into a larger number of disconnected areas, then collecting the crop residues in that grid cell becomes more costly as one need to cross the rivers for more times. This idea is similar to Cutler and Glaeser (1997) and Bai and Jia (2016) who use rivers to construct measures of administrative costs in an area.[20] I find a smaller effect for areas that are more fragmented, and the effect is statistically significant for both the most treated areas and the less treated ar-

[19] Appendix F.6 further rules out an additional mechanism: potential fire bans under transmission towers.

[20] Slightly different from Cutler and Glaeser (1997) and Bai and Jia (2016), I do not directly use the number of rivers as in this setting the number of rivers might affect agricultural production. I instead include the number of rivers (interacted with year fixed effects) as control variables.

eas. Therefore, the results in this table are supportive of an economic incentive mechanism. I further exploit household survey data, although less granular, in Appendix F.2 to show that farmers' income indeed increases after a plant opening. Taken together, both the qualitative and quantitative evidence point to a mechanism of economic incentives behind the reduction in fires.

Structural Transformation The reduction in agricultural fires after a plant opening could be simply caused by a reduction in farming. This could happen if the plant directly employs farmers nearby, or if the plant stimulates the rise of other firms, which indirectly affect the economic structure (Dell and Olken, 2020).

To test such conjecture, I first check the effects of plant openings on land usage in columns (1)-(2) of Table 5. The coefficients are small and statistically insignificant, indicating little change in cropland or settlement land. I then check the effects of plant openings on industrial dynamics in the remaining columns of Table 5. In columns (3)-(4), I test the effect on the number of industrial firms and employees in each grid cell based on the ASIF data. Similarly, in columns (5)-(6), I test the effect on real capital and output.[21] Again, the coefficients are small and statistically insignificant, implying that there is little change in the industrial sector.[22] As the ASIF data do not capture the dynamics of small firms, I supplement such finding using additional data that cover the entire population of industrial firms in Appendix F.3 and the results are similar. Taken together, these results suggest that my baseline findings are unlikely to be driven by changes in farming caused by plant openings.

Relatedly, even if there is no substantial change in land usage, cropland and agricultural production near the plants may be different due to negative externalities, such as air pollution and noise, imposed by the the plants. In Appendix F.4, I show through a case study that a typical

[21] Output is measured by value added. Capital is the capital stock calculated following the procedure in Brandt, Van Biesebroeck and Zhang (2012). Both variables are then converted into real values using industry-level price indices.

[22] Several reasons could explain why the plants have little impact on local economic activity. First, the electricity generated is tapped into distribution grids and electrification rate in China was already close to 100% many years ago (see https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=CN), implying that the arrival of the plants have minimal effect on local economic activity through electrification. Second, given that most plants utilize technologies and machines imported from abroad, it is unlikely that the plants will significantly stimulate related industries locally. From the perspective of economics of density, agglomeration effect of the plants, if any, is also limited by the sparse distribution of industrial activities in the rural areas. Finally, the plants hire relative few formal employees, which is unlikely to generate significant economic impact through affecting the local labor market.

plant imposes negligible negative environmental externalities on nearby cropland, and that the results are robust to dropping areas close to a plant. Furthermore, farmers may also adapt to the plant by adjusting along various dimensions of agricultural production, such as crop choice, harvest techniques, labor usage, timing of harvest/planting or even productivity. These changes may not only affect crop residues and fires, but also be interesting in themselves. In Appendix F.5, I show that there is little change in crop choice and agricultural productivity. For harvest techniques, labor usage and timing of harvest/planting, I cannot test them directly due to lack of granular data. However, these changes should be largely reflected in the change of the value of agricultural production, which is not the case as shown in Appendix F.5.

Migration Another explanation for the findings is that farmers migrate away from a plant to avoid possible air pollution caused by the plant (Chen, Oliva and Zhang, 2017; Barwick et al., 2021). In this case, the cropland may be less utilized and generate fewer agricultural fires. To clarify, migration typically refers to long-term or permanent movement of residence. This form of migration is largely restricted by the *hukou* system in China. However, short-term change of residence is still possible because farmers may migrate temporarily to seek for seasonal work in big cities. Conceptually, pollution-induced migration is unlikely to happen since the biomass power plants use a cleaner technology to process the crop residues than farmers. Moreover, as pointed out by Chen, Oliva and Zhang (2017), migration usually happens at a longer horizon such as a five-year period. As I focus on annual changes, the results are unlikely to be affected by migration.

Empirically, I test such conjecture in the following ways. I first test whether the effects are larger in areas downwind of the plant, since such areas are proved to suffer more from possible pollution (Rangel and Vogl, 2019; Anderson, 2020). The procedure to determine downwind areas can be found in Section 3.2 and is also illustrated in Appendix Figure A1. In columns (1)-(2) of Table 6, I use a continuous variable that is the percentage of days downwind of the plant in a year to measure a grid cell's exposure to pollution. In columns (3)-(4), I further create an indicator for grid cells where the percentage of days downwind of the plant in a year is greater than the median percentage for that plant and year. The coefficients on the triple interaction terms in columns (1)-(4) are either statistically insignificant or have a positive sign, implying that the findings are unlikely to be driven by pollution-induced migration. In columns (5)-(6), I test whether the effects are larger in grid cells where it would be easier to migrate. I define such grid cells as those whose centroids are within 1 km of major transportation networks. I use the major roads and railways in 2000 due to data limitation. I cannot find a larger effect in areas with greater ease of migration. Therefore, migration cannot explain the findings.

Enhanced Inspection The arrival of a plant could signal local officials' attention paid to agricultural fires. Therefore, the findings could be driven by enhanced inspection by local officials that occurs concurrently with the plant opening.

In columns (1)-(2) of Table 7, I test whether the effects are larger in grid cells closer to the city center, as monitoring costs are lower in such areas (Huang et al., 2017). I measure the proximity to city center using the distance from the centroid of each grid cell to the center of the city where the plant is located. I cannot find a larger effect in such areas. In columns (3)-(4), I test whether the effects are weaker in grid cells that are located in a city different from the city where the plant is located, as local officials have little incentive to control pollution outside their jurisdictions (Lipscomb and Mobarak, 2016). I fail to find any differential effects. In columns (5)-(6), I test whether the effects are stronger in grid cells that are located in a city where the leader has stronger promotion incentives, as pollution control is considered by upper-level governments for promotion. I define high promotion incentives if the city party secretary is younger than 57, as the chance for promotion is little for secretaries older than 57 (He, Wang and Zhang, 2020).[23] I cannot find a larger effect for areas where the secretaries have stronger promotion incentives. Hence, the drop in agricultural fires cannot be explained by local officials' actions.

[23] Appendix Table A3 shows that the results are robust if I define promotion incentives using the mayor's age.

5 The Effects of No Burning Zones

5.1 Model Specification

Since exposure to a no burning zone is a deterministic and discontinuous function of the distance from each area to the 15 km-radius circle around the airport, I use a spatial regression discontinuity design to estimate the effects of no burning zones. Following Hahn, Todd and Van der Klaauw (2001), I specify a local linear regression:

$$\begin{aligned} Y_{isjt} =& \delta_{sj} + \beta_1 Inside_{ij} + \beta_2 Distance_{ij} + \beta_3 Inside_{ij} \times Distance_{ij} + \varepsilon_{isjt} \\ & s.t. \ -h \leq Distance_{ij} \leq h \end{aligned} \tag{7}$$

where the unit of observation is 1 km $\times$ 1 km grid cells. i denotes grid cells, s denotes boundary segments, j denotes airports, and t denotes years. $Inside_{ij}$ is an indicator equal to 1 if the centroid of grid cell i is inside the no burning zone around airport j and 0 otherwise. The sample used here only includes active airports, so the treatment variable $Inside_{ij}$ is time-invariant. I also utilize the panel structure of the data by including the years before an airport is active and estimating a difference-in-discontinuities specification in the robustness checks. $Distance_{ij}$ is the running variable, measuring the distance from the centroid of grid cell i to the no burning zone boundary around airport j (negative for observations outside the boundary). I include the interaction between $Inside_{ij}$ and $Distance_{ij}$ to allow for the coefficients on $Distance_{ij}$ to be different for each side of the boundary. The coefficient of interest is β_1, which captures the jump of outcomes at the boundary. To ensure the areas just outside the boundary are valid counterfactuals for areas just inside the boundary, I split the boundary into 10 segments of equal length and assign each grid cell to the nearest segment. I then include airport $\times$ segment fixed effects δ_{sj}, which ensures that I am comparing grid cells within a narrowly defined geographic area. Due to spatial adjacency, grid cells along the same segment should be ex-ante similar.

For most of the analysis, I report the results using three fixed bandwidths: 2 km, 4 km, 6 km, and also the optimal bandwidth calculated following the procedure proposed by Imbens and Kalyanaraman (2012).[24] The optimal bandwidths for different outcomes and specifications

[24]For specifications with fixed effects and predetermined covariates, the optimal bandwidth is selected based on

generally range from 2 km to 6 km with a mean value of about 4 km. I also check the robustness of the baseline results to a broader set of bandwidths ranging from 2 km to 10 km. I weight observations within the bandwidth using a uniform kernel for convenience and check robustness to a triangular kernel, as in practice the choice of kernels has little impact (Lee and Lemieux, 2010). I cluster the standard errors at the airport level to account for spatial and temporal correlation (Cameron and Miller, 2015). I also show robustness to clustering at the $2° \times 2°$ level and accounting for arbitrary correlation across spatially adjacent observations (Conley, 1999; Hsiang, 2010).

5.2 Identification Concerns

The key identification assumption of this RD design is that all relevant preexisting factors change very smoothly around a no burning zone boundary. This assumption is needed to ensure areas just outside the boundary are an appropriate counterfactual for those located just inside the boundary. Conceptually, this assumption is plausibly satisfied since the boundary is purely artificially drawn and does not overlap with other administrative boundaries. Nevertheless, I check for possible jumps at the boundary for a set of preexisting covariates, including geography, location, land usage, and economic structure. [25]

Another threat to the identification assumption is that there could be selective sorting of farmers around the boundary.[26] As mentioned before, long-term or permanent change of residence is hard for farmers due to the *hukou* system, largely alleviating the selective sorting concern. Nevertheless, selective sorting may still occur in other forms. First, cropland inside the boundary may be converted to commercial land, either by farmers to avoid regulation or by local governments to avoid accountability for agricultural fires. Conceptually, this is unlikely to happen since cropland is strictly protected by the central government.[27] For farmers, it is illegal to change the usage of cropland. For local governments, they need to get the approval

the residuals from regressing the dependent variable on these fixed effects and predetermined covariates (Lee and Lemieux, 2010).

[25] I do not compare climate variables such as temperature and precipitation because high resolution data are not available. Moreover, across the grid cells around an airport, such variables should not vary much.

[26] Selective sorting could happen as an outcome of the discontinuous policy, but then the interpretation would be different.

[27] See http://www.npc.gov.cn/npc/c2168/200011/048bf4075bee4b4da938a7681d63b58f.shtml.

from the central government for changes. Empirically, I check this possibility by looking at the change of land usage ex post. Second, local governments may implicitly encourage industrial firms to locate inside the boundary to transform the economic structure. I test this possibility by looking at the number and size of industrial firms using the ASIF data. Third, farmers inside the boundary may farm outside the boundary to avoid regulation. Conceptually, this is unlikely to happen since cropland is owned collectively and each farmer only owns the use rights of a s-mall piece of land near their house. To get cropland outside, one needs to negotiate with farmers there and get the approval of relevant authorities. Moreover, it would also be physically costly to farm in areas far away. Empirically, I adopt a "donut hole" specification by excluding areas close to the boundary (Barreca et al., 2011; Barreca, Lindo and Waddell, 2016), since farmers living in these areas are likely to have strong incentive to cross the boundary. Fourth, farmers inside the boundary may move out temporarily to seek for seasonal work in big cities. I test this possibility by looking at areas with greater ease of migration, which are areas with access to major transportation networks. Finally, selective sorting may also occur at the airport level. For the existing airport, it is nearly impossible to move since it is too costly. However, local governments may choose to locate new airports in some areas with very few agricultural fires to avoid accountability. I check this possibility by looking at airports established before the launch of the regulation in 1999.

5.3 Results

Baseline RD results Table 8 reports the RD results that identify the effects of no burning zones on agricultural fires using three fixed bandwidths: 2 km, 4 km, and 6 km, and the optimal bandwidth. Columns (1)-(3) examine the number of fires. The coefficients are small compared to the sample mean, and statistically insignificant. Moreover, they remain quite stable as more granular fixed effects are included. The coefficients change slightly when using larger bandwidths, but are still negligible compared to the sample mean. Columns (4)-(6) examine the probability that any fires occur. The patterns are quite similar to those in columns (1)-(3). Across all specifications and bandwidths, the coefficients range from -0.0004 to -0.0001, which are e-conomically small compared to the sample mean of 0.009 and are statistically insignificant.

Overall, these results suggest that the regulation fails to reduce agricultural fires. In Appendix Figure A6, I report the RD results using different bandwidths ranging from 2 km to 10 km. The results are not sensitive to alternative bandwidths and are still small and insignificant.

Figure 7 visualizes the RD results. The circles denote the average number of fires or probability that any fires occur within a distance bin, after partialling out the airport $\times$ segment fixed effects. The lines are fitted separately for grid cells on each side of the boundary, and the dashes are the associated 95% confidence intervals. In line with the RD estimates in Table 8, the jumps at the boundary are small. Moreover, the confidence intervals at the boundary are largely overlapped, implying that the difference in fires across the boundary is statistically indistinguishable from zero.

As discussed in Section 3, classical measurement errors in agricultural fires could decrease the precision of the estimates. However, as the economic magnitude is also close to zero, it is then hard to conclude that the zones really work. Even if the measurement errors are non-random due to farmers timing the burning, then this would only bias my findings against zero; as I do not find any reduction, this further strengthens my conclusion that the zones fail to reduce burning.

Preexisting balance I conduct a balance check for the preexisting covariates around the boundary, using the same specification as equation (7). Except for the distance to major transportation networks that is measured in 2000 due to data limitation,[28] all other time-varying variables are measured in 1998, one year before the launch of the regulation in 1999. The results are reported in Appendix Table A22, using three fixed bandwidths: 2 km, 4 km, and 6 km, and the optimal bandwidth. The coefficients are generally small, compared to the sample means, and statistically insignificant for most of the coefficients. A few covariates, such as the distance to city border, the share of settlement land, and the nighttime light intensity in 1998, show statistically significant jump at the boundary, but the magnitude of the jump is negligible compared to the sample mean. In Appendix Figure A4 and Appendix Figure A5, I show the RD plots for all these covariates using the optimal bandwidth. In line with the results in Appendix

[28] The transportation networks in 2000 can also be viewed as predetermined as they were likely to be planned several years before 2000.

Table A22, there are no discernible jumps at the boundary.

Selective sorting I check for possible selective sorting in Table 9, still adopting the same specification as equation (7) and using three fixed bandwidths: 2 km, 4 km, and 6 km, and the optimal bandwidth. In columns (1)-(2), I find no evidence of differences in land usage, either for farming or for settlement, across the boundary. The coefficients are small and statistically insignificant, alleviating the concern that farmers or local governments may change land usage as a response to the regulation. In columns (3)-(6), I test the difference in economic structure, measured by the number and size of industrial firms, across the boundary. Most of the coefficients are small and statistically insignificant. There is a statistically significant jump in the number of firms and employees when using the 2 km bandwidth, but the economic magnitude is relatively small. Moreover, the positive coefficients would bias the results towards finding a significant effect of the regulation. In columns (7)-(8), I exclude areas within 1 km of the boundary,[29] since farmers living in such areas have strong incentive to cross the boundary for farming to avoid regulation. This type of sorting would also bias the results towards finding a significant effect of the regulation. I do not find substantial changes after excluding such areas. In columns (9)-(12), I split the sample into two groups by the ease of migration. I define migration as easy for areas within 1 km of major roads and railways and hard for other areas. I still cannot find a significant effect in these two types of areas. Finally, I test whether the results are driven by airports sorting into areas with few agricultural fires. To this end, I focus on airports established before 1999 in columns (13)-(14)and still cannot find a significant effect. In sum, these findings suggest that the results are unlikely to be driven by various forms of selective sorting.

5.4 Robustness

Table 10 and Table 11 show the robustness of the baseline results. I report the results using the optimal bandwidth, except for a few specifications where fixed bandwidths are used due to the lack of a a widely accepted optimal bandwidth algorithm. Overall, the difference in agri-

[29] Appendix Table A23 shows that the results are similar if I exclude areas within 2 or 3 km of the boundary.

cultural fires across the boundary is small and statistically insignificant across different samples and specifications.

Difference-in-discontinuities To fully utilize the panel structure of the data and control for confounding factors by exploiting within grid cell variation, I use a difference-in-discontinuities specification following Grembi, Nannicini and Troiano (2016):

$$\begin{aligned} Y_{isjt} = & \delta_{sj} + \mu_{jt} + \beta_1 Inside_{ij} + \beta_2 Distance_{ij} + \beta_3 Inside_{ij} \times Distance_{ij} + \beta_4 Inside_{ij} \times Post_{jt} \\ & + \beta_5 Distance_{ij} \times Post_{jt} + \beta_6 Inside_{ij} \times Distance_{ij} \times Post_{jt} + \varepsilon_{isjt} \\ & s.t. \quad -h \leq Distance_{ij} \leq h \end{aligned} \tag{8}$$

where μ_{jt} is airport $\times$ year fixed effects, $Post_{jt}$ is a indicator equal to 1 if airport j is active at year t, and all other variables are as previously defined. I only include airports that opened between 2002 and 2015, in order to have at least one year of observations before the opening and at least one year of observations after the opening. The coefficient of interest is β_4, which measures the change in the discontinuities after the opening. I choose 4 km as the baseline bandwidth.[30] The results based on equation (8) are reported in columns (1) and (3) of Table 10. The coefficients on $Inside_{ij} \times Post_{jt}$ are still small and statistically insignificant. Columns (2) and (4) of Table 10 exploit within grid cell variation by further including grid cell fixed effects, and the results show no substantial change. Similar to difference-in-differences, this method requires that fires just inside and outside the boundary are on parallel trends before airport openings. I verify this assumption by estimating an event study variant of equation (8) and plot the results in Figure 8. I do not find that fires around the boundary exhibit differential pre-trends.

Alternative FEs and samples Columns (5)-(8) of Table 10 use alternative number of boundary segments. In the baseline specification, I split each boundary into 10 segments to ensure comparison of observations in close proximity. Here I instead split each boundary into 5 or 20 segments to evaluate the robustness of the results. The results are not affected.

[30]This is due to the lack of a widely accepted optimal bandwidth algorithm for the multi-dimensional RD design in the literature. In Appendix Table A24, I show that the results are robust to alternative bandwidths.

Columns (9)-(10) of Table 10 use only airports around which the share of cropland accounts for more than 10%. The insignificant difference in agricultural fires across the boundary might be driven by the lack of cropland around the airports. The results are still small and statistically insignificant. Columns (11)-(12) of Table 10 focus on airports located in major grain-producing provinces to check whether the insignificant results are driven by lower grain output, since agricultural fires come from burning crop residues.[31] The results barely change.

Two-dimensional RD Columns (1)-(2) of Table 11 use a two-dimensional linear RD polynomial in longitude and latitude (Dell, 2010). Although the treatment is purely determined by a one-dimensional variable that is the distance to the airport, such two-dimensional RD polynomial could flexibly control for geographic factors and provide checks on the baseline results. I choose 4 km as the baseline bandwidth and find that the coefficients are small and statistically insignificant. [32]

Weighting and quadratic polynomial Columns (3)-(4) of Table 11 use a triangular kernel, which gives higher weights to observations closer to the boundary.[33] The results are not affected. Columns (5)-(6) of Table 11 check robustness to quadratic polynomial in distance to the boundary.[34] The results are stable.

Including covariates Columns (7)-(10) of Table 11 add geographic covariates and other predetermined covariates used in the balance checks. In theory, adding such covariates should not affect the results but could improve the precision of the RD estimates (Imbens and Lemieux, 2008). The results stay small and statistically insignificant.

---

[31]The provinces are located in north, central, east, and northeast China, including Hubei, Hunan, Jiangxi, Jiangsu, Zhejiang, Hebei, Henan, Shandong, Shanxi, Heilongjiang, Jilin, and Liaoning.

[32]This is due to the lack of a widely accepted optimal bandwidth algorithm for the multi-dimensional RD design in the literature (Dell, Lane and Querubin, 2018). In Appendix Table A25, I show that the results are robust to alternative bandwidths.

[33]The weight is $1 - abs(Distance_{ij})/h$, where $Distance_{ij}$ denotes the distance from each observation to the boundary and h is the bandwidth

[34]I do not try orders higher than two since they are not recommended in RD designs and could generate undesirable results (Gelman and Imbens, 2019).

Spatial correlation Columns (11)-(12) of Table 11 cluster the standard errors at the $2^\circ \times 2^\circ$ level. Cluster (13)-(14) of Table 11 report the results using the Conley standard errors, which allow for spatial correlation for observations within a 200 km radius and serial correlation across 5 years. The results are not affected.

5.5 Heterogeneity

To check whether the insignificant effects of no burning zones on agricultural fires represent general policy failures across regions and years, I conduct RD estimation separately for each airport and for each year using the optimal bandwidth. The distribution of the coefficients is shown in Figure 9. The airport-specific coefficients are small and centered at 0. Moreover, the coefficients are statistically insignificant for the vast majority of airports. For example, 190 out of 205 airports ($\approx 93\%$) show insignificant coefficients in Panel A.[35] For the remaining 15 airports with significant coefficients, 4 airports show positive signs and 11 airports show negative signs. The year-specific coefficients also fluctuate around 0 and are generally insignificant. Overall, these results show that the weak effects of no burning zones on agricultural fires reflect a general policy failure.

I further test whether no burning zones work only at certain type of areas such as those subject to higher safety requirement or enjoying higher political status. Standard economic theory predicts that regulators would put more efforts into such areas, as the marginal benefits are higher. The results are reported in Table 12 using three fixed bandwidths and the optimal bandwidth. I first focus on upwind areas, which are defined as grid cells where the percentage of days upwind of the airport in a year is greater than the median percentage for that airport and year.[36] Such areas are viewed as subject to higher safety requirement, as agricultural fires in such areas have a larger threat to the take-off and landing of the planes. If regulators selectively target such areas, then one should see a significant effect in such areas. The results in columns (1)-(2) and (7)-(8), however, show that the zones work neither in upwind areas nor in other

[35]Results for Panel B are similar.

[36]The procedure to define upwind can be found in Section 3.2 and is also illustrated in Appendix Figure A1.

areas. I next test whether the zones work in provincial cities.[37] These cities are typically more economically developed than other cities, and local leaders in these cities enjoy higher political status than other cities. Therefore, local leaders should have stronger incentive to fight agricultural fires for both economic and political reasons. However, the results in columns (3)-(4) and (9)-(10) show that the zones work neither in provincial cities nor in other cities. Finally, I test whether the zones work in cities where local leaders have higher promotion incentives. As before, I define higher promotion incentives as cities where the party secretaries are younger than 57.[38] The results in columns (5)-(6) and (11)-(12) show that the effects are still insignificant for both high- and low-incentive cities, except in column (11) where I use a much smaller bandwidth (2 km) and measure agricultural fires using a binary variable. For this coefficient that is statistically significant, the economic magnitude is relatively small compared to the sample mean. Further, the statistical significance disappears when using the optimal bandwidth. Hence, it is hard to conclude that the effectiveness of the zones changes with promotion incentives.

I also check separately for fires occurring in the daytime and nighttime. There is a concern that farmers may postpone burning that would otherwise occur in the daytime until nighttime to escape regulation. If this is true, one may not see an overall reduction in fires inside the no burning zones but could expect to see a reduction in the daytime; given that there are much more flights in the daytime, a possible reduction in fire in the daytime could still be meaningful. The results in Appendix Table A27, however, show that the zones work neither in the daytime nor nighttime.

6 Discussion

So far I have documented that the biomass power plants are effective in reducing burning while the no burning zones have little impact on burning. In this section I discuss a few remaining issues. First, one may be concerned about potential side effects of the plants. Appendix F.4 shows through a case study that a typical plant imposes little negative environmental external-

[37]These include 27 provincial capitals, 4 cities directly controlled by the central government, and 5 cities under separate state planning. The results stay unchanged if I only include provincial capitals.

[38]In Appendix Table A26, I measure promotion incentives using mayors' ages, as mayors may be responsible for economic activities. The results are similar.

ities, such as air pollutants, water pollutants, and noise, on surrounding areas. Appendix H.1 shows that the plants can reduce ambient $PM_{2.5}$ concentrations by at least about 2%. Appendix H.2 shows that the plants do not deteriorate nearby water quality, which is consistent with the case study in Appendix F.4.

One may be interested in knowing whether areas falling into the overlapping catchment areas of more than one plant are different, which is, however, a priori unclear. The revenue from selling crop residues may not react to more than one potential buyer given the low and homogeneous price of electricity, which bounds the price of crop residues.[39] The explicit costs of collecting crop residues (e.g., labor and machine costs) may decrease with the number of potential buyers if more buyers lead to cheaper technologies and labor, although this may only happen when the number of buyers is sufficiently large. The explicit costs, which involve the gains from engaging in other activities, may increase if farmers work in the plants and earn higher wages in the presence of more than one plant, or if more plants lead to more prosperous local economic activities, although I empirically show that there is no change in employment and local economic activities. In sum, it is a priori unclear whether the economic benefits from selling crop residues would be different in the overlapping catchment areas. Consistent with such discussion, Appendix Table A19 demonstrates that these areas are no different from other areas.[40]

One may be concerned about differential crop types around the plants and airports, which may lead to the difference in crop residues and burning patterns. Although possible difference in crop types around the two types of areas does not affect the internal validity of my results, it may invalidate the contrast between these results. In Appendix I, I compare crop types around these two types of areas and do not find a substantial difference. Furthermore, the findings are robust to using a matching procedure that eliminates the difference in crop types across the plants and airports.

Throughout the analysis, grid cells are treated as if they are affected only by the plants or

[39] A plant may simply choose to not operate at its full capacity if the price of crop residues would be pushed up by more than one plant.

[40] Specifically, I include all plants and create an indicator *overlap*, which is equal to 1 if a grid cell around a certain plant also falls in the catchment areas of other plants that are active. I then fully interact this indicator with other terms in the baseline DID specification. The coefficients on the triple interaction terms are generally small and statistically insignificant, suggesting little difference between the overlapping areas and other areas.

airports. However, some grid cells around the plants (airports) may also be close to the airports (plants), which may raise the concern about compound treatment. To alleviate such concern, I redo the baseline analysis focusing on a sample in which the catchment areas of the plants do not overlap with those of the airports. Specifically, I drop all plants (airports) that are placed within 30 km of airports (plants). This trimmed sample includes 148 plants (78%) and 166 airports (81%). As shown in Appendix Table A21 and A29, the results are very similar to the baseline results.

Finally, one may also worry about the external validity of the spatial RD design around airports. Specifically, the no burning zones may be effective in curtailing fires at areas far away from the boundaries, which cannot be captured by the RD estimates. I therefore complement the RD design with a DID strategy that takes into consideration areas far from the boundaries in Appendix J. The results are small and statistically insignificant. Such results, together with the RD results, could be caused by a number of factors such as difficulty in patrolling the rural areas, lack of rule-abiding norms in rural areas, laziness, or incompetence of local officials. The current data do not allow me to disentangle these potential mechanisms. Instead, the goal of this paper is to provide credible evidence on the effects of different instruments on reducing agricultural fires.

7 Conclusion

In this paper, I test the effects of economic incentives and command and control on reducing agricultural fires using satellite-based high-resolution fire measures and two quasi-experiments: the opening of biomass power plants purchasing agricultural residues from nearby areas and the drawing of no burning zones banning agricultural fires. The results suggest that the plants could reduce fires by at least 30%. Such reduction is consistent with the plants increasing farmers' economic incentives to collect crop residues, as demonstrated not only by narrative evidence depicting farmers' reaction to the plants, but also by quantitative evidence showing smaller effects in areas that are more costly to collect crop residues. Using household survey data, I corroborate such mechanism by showing an increase in farmers' income after a plant opening.

Turning to the airports, I find no discernible difference in fires between areas just inside the no burning zones and areas just outside. Such finding is further complemented by a DID design that considers all areas inside the no burning zones, which extends the external validity of the baseline findings. By examining each airport and each year separately, I confirm that the results are not an artifact of potential heterogeneity across airports or years. Further heterogeneity analysis shows that the zones are dysfunctional even in areas with stringent regulatory enforcement, suggesting that the results represent a general policy failure.

A key contribution of this paper is that it examines the effectiveness of two common policy instruments in curbing agricultural fires rigorously using microdata, which has practical implications given that agricultural fires are still ubiquitous in the world. Particularly, in many developing countries plagued by agricultural fires, resorting to regulations or laws seems to be a straightforward and economical way to fight agricultural fires. The results in the China context show that, however, even in a country with strong state capacity and in areas with high safety requirements (i.e., areas around airports), a command and control policy could still be ineffective in curtailing agricultural fires. In contrast, the results show the success of instruments based on economic incentives. Note that the economic incentives created by the plants are slightly different from those created by traditional cash transfers, as the latter may incur considerable enforcement costs, which may undermine the policy effectiveness. In this vein, it would be interesting for future research to contrast the effectiveness of various types of economic incentive-based policy instruments in curbing agricultural fires and explore how the effectiveness interacts with enforcement costs.

References

Anderson, Michael L. 2020. "As the wind blows: The effects of long-term exposure to air pollution on mortality." *Journal of the European Economic Association*, 18(4): 1886–1927.

Bai, Ying, and Ruixue Jia. 2016. "Elite recruitment and political stability: the impact of the abolition of China's civil service exam." *Econometrica*, 84(2): 677–733.

Barreca, Alan I, Jason M Lindo, and Glen R Waddell. 2016. "Heaping-induced bias in regression-discontinuity designs." *Economic Inquiry*, 54(1): 268–293.

Barreca, Alan I, Melanie Guldi, Jason M Lindo, and Glen R Waddell. 2011. "Saving babies? Revisiting the effect of very low birth weight classification." *The Quarterly Journal of Economics*, 126(4): 2117–2123.

Barwick, Panle Jia, Dave Donaldson, Shanjun Li, Yatang Lin, and Deyu Rao. 2021. "Improved Transportation Infrastructure Facilitates Adaptation to Pollution and Temperature Extremes." *Available at SSRN 3814060*.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics*, 119(1): 249–275.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing." *Journal of Development Economics*, 97(2): 339–351.

Callaway, Brantly, and Pedro HC Sant'Anna. 2020. "Difference-in-differences with multiple time periods." *Journal of Econometrics*.

Cameron, A Colin, and Douglas L Miller. 2015. "A practitioner's guide to cluster-robust inference." *Journal of Human Resources*, 50(2): 317–372.

Cheng, Zhen, Shuxiao Wang, Xiao Fu, John G Watson, Jingkun Jiang, Qingyan Fu, Changhong Chen, Bingye Xu, Jianqiao Yu, Judith C Chow, et al. 2014. "Impact of biomass burning on haze pollution in the Yangtze River delta, China: a case study in summer 2011." *Atmospheric Chemistry and Physics*, 14(9): 4573–4585.

Chen, Jianmin, Chunlin Li, Zoran Ristovski, Andelija Milic, Yuantong Gu, Mohammad S Islam, Shuxiao Wang, Jiming Hao, Hefeng Zhang, Congrong He, et al. 2017. "A review of biomass burning: Emissions and impacts on air quality, health and climate in China." *Science of the Total Environment*, 579: 1000–1034.

Chen, Shuai, Paulina Oliva, and Peng Zhang. 2017. "The effect of air pollution on migration: evidence from China." *National Bureau of Economic Research*.

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of Econometrics*, 92(1): 1–45.

Currie, Janet, and Reed Walker. 2011. "Traffic congestion and infant health: Evidence from E-ZPass." *American Economic Journal: Applied Economics*, 3(1): 65–90.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker. 2015. "Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings." *American Economic Review*, 105(2): 678–709.

Cutler, David M, and Edward L Glaeser. 1997. "Are ghettos good or bad?" *The Quarterly Journal of Economics*, 112(3): 827–872.

Davis, Lucas W. 2008. "The effect of driving restrictions on air quality in Mexico City." *Journal of Political Economy*, 116(1): 38–81.

de Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review*, 110(9): 2964–96.

Dell, Melissa. 2010. "The persistent effects of Peru's mining mita." *Econometrica*, 78(6): 1863–1903.

Dell, Melissa, and Benjamin A Olken. 2020. "The development effects of the extractive colonial economy: The dutch cultivation system in java." *The Review of Economic Studies*, 87(1): 164–203.

Dell, Melissa, Nathan Lane, and Pablo Querubin. 2018. “The historical state, local collective action, and economic development in Vietnam.” *Econometrica*, 86(6): 2083–2121.

Dominici, Francesca, Roger D Peng, Michelle L Bell, Luu Pham, Aidan McDermott, Scott L Zeger, and Jonathan M Samet. 2006. “Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases.” *The Journal of the American Medical Association*, 295(10): 1127–1134.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. “Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India.” *The Quarterly Journal of Economics*, 128(4): 1499–1545.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2018. “The value of regulatory discretion: Estimates from environmental inspections in India.” *Econometrica*, 86(6): 2123–2160.

Fang, Meng, Peng Zhang, and Zhe Xu. 2006. “The Application of the 3S Technique to the Management of Crop Residue Burning.” *Remote Sensing for Land & Resources*, 69(3).

Gelman, Andrew, and Guido Imbens. 2019. “Why high-order polynomials should not be used in regression discontinuity designs.” *Journal of Business & Economic Statistics*, 37(3): 447–456.

Giglio, Louis, Jacques Descloitres, Christopher O Justice, and Yoram J Kaufman. 2003. “An enhanced contextual fire detection algorithm for MODIS.” *Remote Sensing of Environment*, 87(2-3): 273–282.

Goodman-Bacon, Andrew. 2021. “Difference-in-differences with variation in treatment timing.” *Journal of Econometrics*.

Graff Zivin, Joshua, Tong Liu, Yingquan Song, Qu Tang, and Peng Zhang. 2020. “The unintended impacts of agricultural fires: Human capital in China.” *Journal of Development Economics*, 147: 102560.

Greenstone, Michael, and B Kelsey Jack. 2015. “Envirodevonomics: A research agenda for an emerging field.” *Journal of Economic Literature*, 53(1): 5–42.

Greenstone, Michael, and Rema Hanna. 2014. “Environmental regulations, air and water pollution, and infant mortality in India.” *American Economic Review*, 104(10): 3038–72.

Greenstone, Michael, Guojun He, Shanjun Li, and Eric Yongchen Zou. 2021. “Chinas War on Pollution: Evidence from the First 5 Years.” *Review of Environmental Economics and Policy*, 15(2): 281–299.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano. 2016. “Do fiscal rules matter?” *American Economic Journal: Applied Economics*, 1–30.

Guo, Shiqi. 2020. “How Does Straw Burning Affect Urban Air Quality in China?” *American Journal of Agricultural Economics*.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. “Identification and estimation of treatment effects with a regression-discontinuity design.” *Econometrica*, 69(1): 201–209.

Harrison, Ann, Benjamin Hyman, Leslie Martin, and Shanthi Nataraj. 2015. “When do Firms Go Green? Comparing Command and Control Regulations with Price Incentives in India.” *National Bureau of Economic Research*.

He, Guojun, Shaoda Wang, and Bing Zhang. 2020. “Watering down environmental regulation in China.” *The Quarterly Journal of Economics*, 135(4): 2135–2185.

He, Guojun, Tong Liu, and Maigeng Zhou. 2020. “Straw burning, PM2. 5, and death: Evidence from China.” *Journal of Development Economics*, 102468.

Hsiang, Solomon M. 2010. “Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America.” *Proceedings of the National Academy of sciences*, 107(35): 15367–15372.

Huang, Zhangkai, Lixing Li, Guangrong Ma, and Lixin Colin Xu. 2017. "Hayek, local information, and commanding heights: Decentralizing state-owned enterprises in China." *American Economic Review*, 107(8): 2455–78.

Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal bandwidth choice for the regression discontinuity estimator." *The Review of Economic Studies*, 79(3): 933–959.

Imbens, Guido W, and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics*, 142(2): 615–635.

Keele, Luke J, and Rocio Titiunik. 2015. "Geographic boundaries as regression discontinuities." *Political Analysis*, 23(1): 127–155.

Lee, David S, and Thomas Lemieux. 2010. "Regression discontinuity designs in economics." *Journal of Economic Literature*, 48(2): 281–355.

Lipscomb, Molly, and Ahmed Mushfiq Mobarak. 2016. "Decentralization and pollution spillovers: evidence from the re-drawing of county borders in Brazil." *The Review of Economic Studies*, 84(1): 464–502.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics*, 142(2): 698–714.

Nickell, Stephen. 1981. "Biases in dynamic models with fixed effects." *Econometrica*, 1417–1426.

Nunn, Nathan, and Diego Puga. 2012. "Ruggedness: The blessing of bad geography in Africa." *Review of Economics and Statistics*, 94(1): 20–36.

Oates, Wallace E, Paul R Portney, and Albert M McGartland. 1989. "The net benefits of incentive-based regulation: a case study of environmental standard setting." *The American Economic Review*, 79(5): 1233–1242.

Oliva, Paulina. 2015. "Environmental regulations and corruption: Automobile emissions in Mexico City." *Journal of Political Economy*, 123(3): 686–724.

Oster, Emily. 2019. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics*, 37(2): 187–204.

Rangel, Marcos A, and Tom S Vogl. 2019. "Agricultural fires and health at birth." *Review of Economics and Statistics*, 101(4): 616–630.

Sager, Lutz. 2019. "Estimating the effect of air pollution on road safety using atmospheric temperature inversions." *Journal of Environmental Economics and Management*, 98: 102250.

Souza-Rodrigues, Eduardo. 2019. "Deforestation in the Amazon: A unified framework for estimation and policy analysis." *The Review of Economic Studies*, 86(6): 2713–2744.

Tanaka, Shinsuke. 2015. "Environmental regulations on air pollution in China and their impact on infant mortality." *Journal of Health Economics*, 42: 90–103.

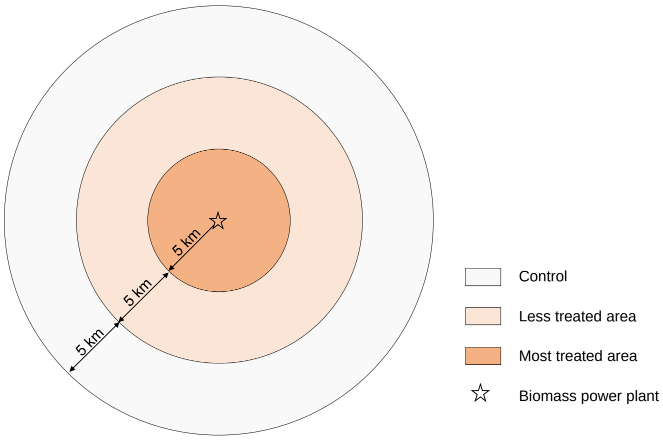
Viard, V Brian, and Shihe Fu. 2015. "The effect of Beijing's driving restrictions on pollution and economic activity." *Journal of Public Economics*, 125: 98–115.

Villar, MC, V Petrikova, M Dıaz-Ravina, and T Carballas. 2004. "Changes in soil microbial biomass and aggregate stability following burning and soil rehabilitation." *Geoderma*, 122(1): 73–82.

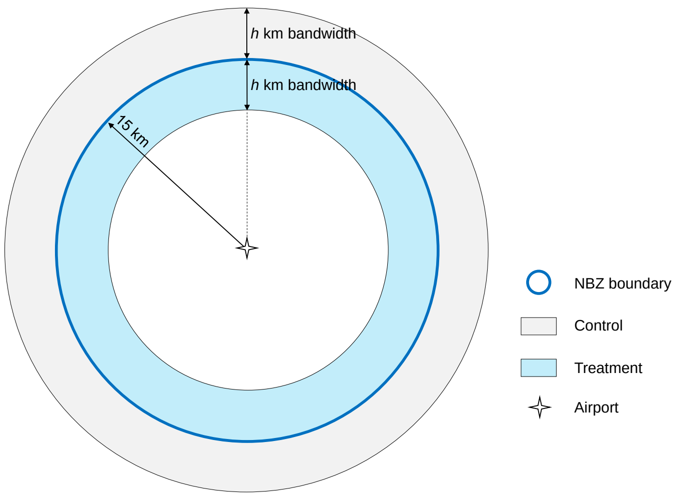
Figures and Tables

Figure 1: Illustration of Identification Strategy

A: Effects of biomass power plants: spatial difference-in-differences



B: Effects of no burning zones: spatial regression discontinuity design



Notes: The unit of observation is 1 km × 1 km grid cells. This figure illustrates the construction of treatment and control groups used in testing the effects of biomass power plants and no burning zones on agricultural fires, respectively.

Figure 2: Distribution of Biomass Power Plants and Airports

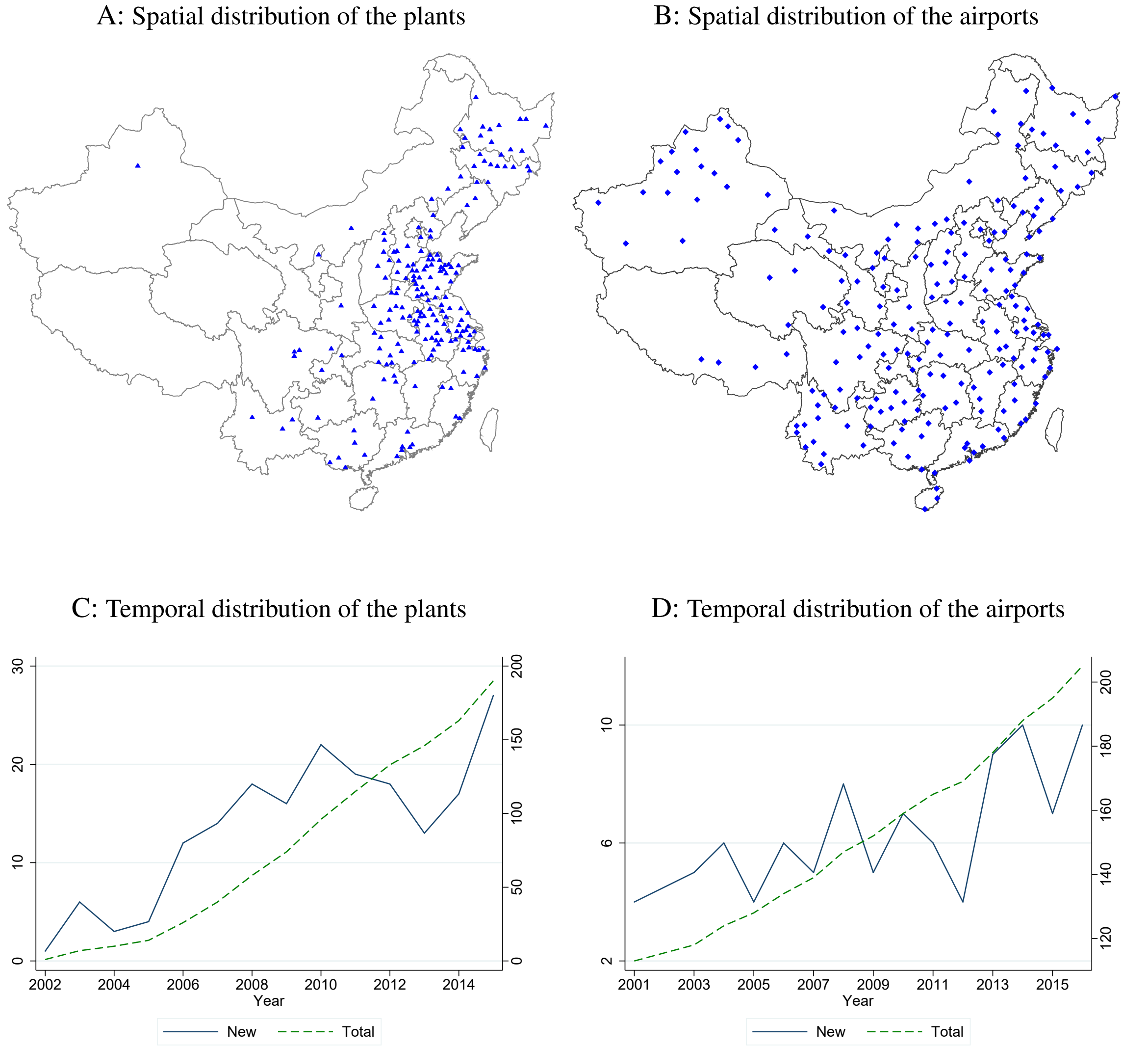
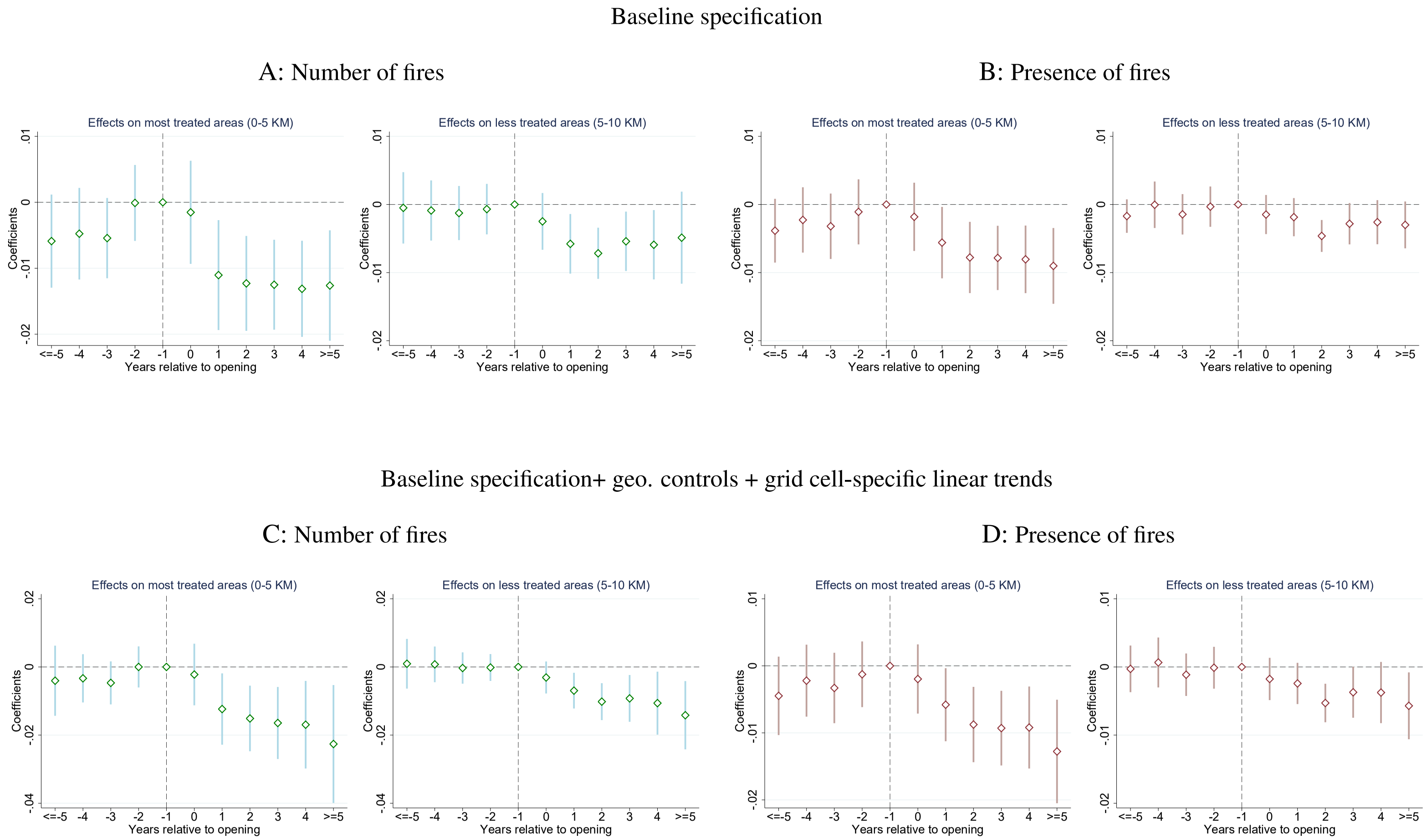


Figure 3: Agricultural Fires around Biomass Power Plants and No Burning Zone Boundaries

Notes: Panel A shows the number of agricultural fires within 15 km of biomass power plants. Panel B shows the number of agricultural fires within 4 km of the no burning zone boundaries. Panel C and D break down them by treatment status.

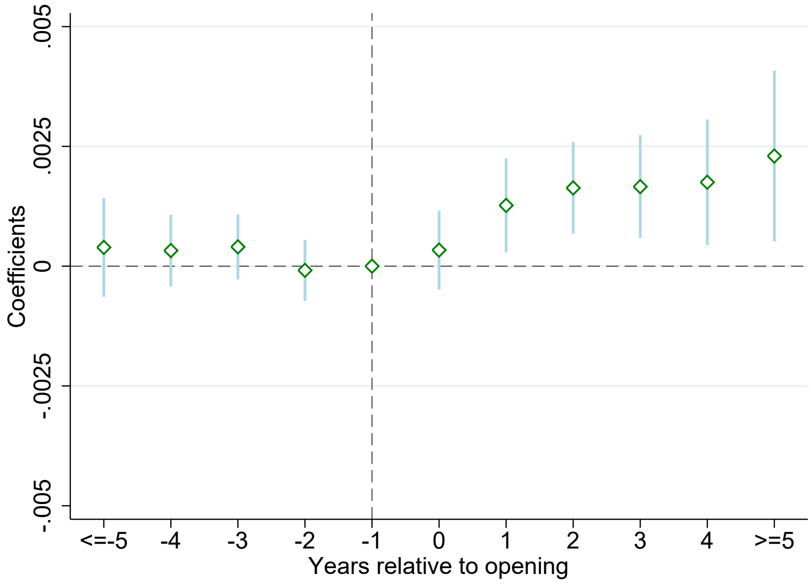
Figure 4: Biomass Power Plants: Dynamic Effect on Agricultural Fires



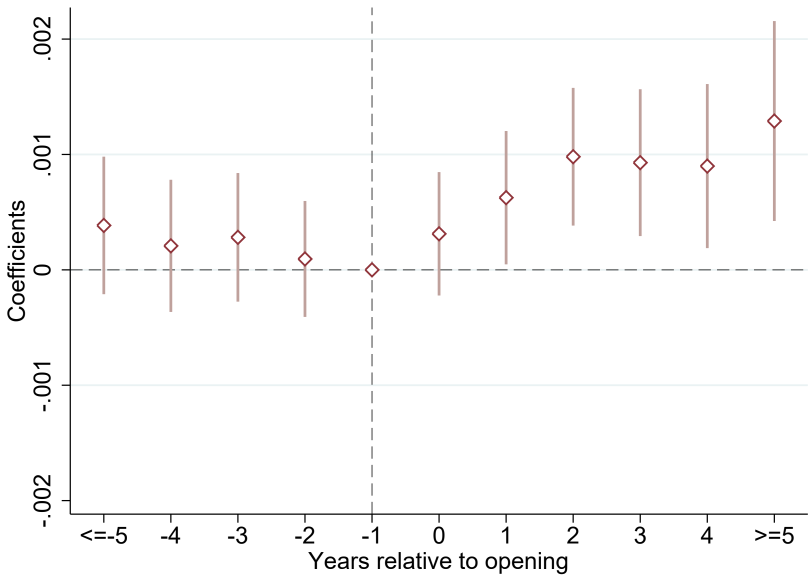
Notes: The unit of observation is 1 km × 1 km grid cells. This figure checks the dynamic effect of biomass power plants on agricultural fires by conducting an event study plot. *Presence of fires* is an indicator for observing at least one fire. The year before a plant opening is omitted as the reference group. Panel A and B only control for grid cell fixed effects and plant × year fixed effects. Panel C and D further include grid cell-specific linear trends and a set of geographic controls interacted with year fixed effects. The diamonds denote coefficients and bars denote the associated 95% confidence intervals, where standard errors used to construct the confidence intervals are clustered at the plant level.

Figure 5: Biomass Power Plants: Dynamic Effect Using Continuous Treatment Intensity

A: Number of fires



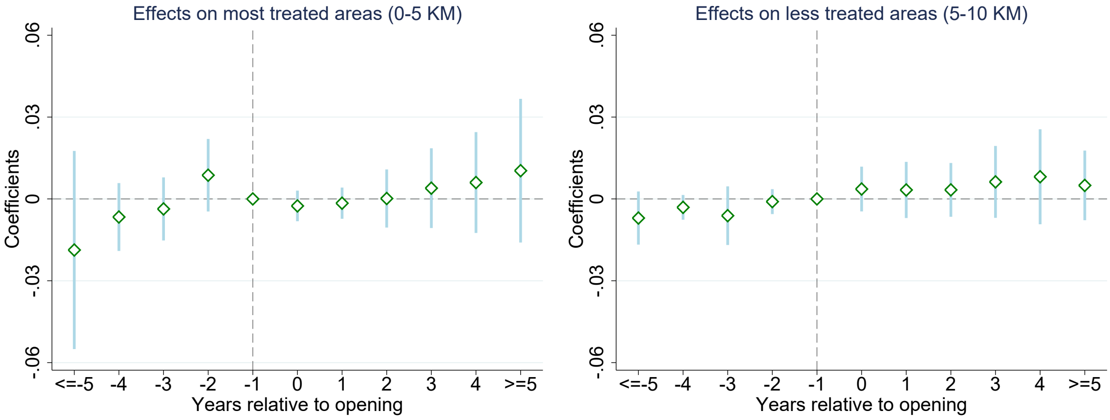
B: Presence of fires



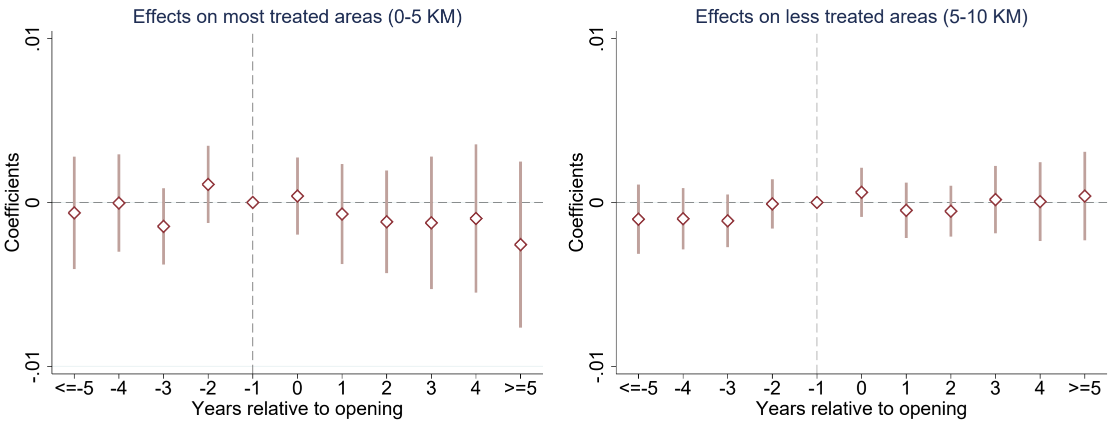
Notes: The unit of observation is 1 km × 1 km grid cells. This figure checks the dynamic effect of biomass power plants on agricultural fires by conducting an event study plot with continuous treatment intensity, which is the distance to the plant. *Presence of fires* is an indicator for observing at least one fire. The year before a plant opening is omitted as the reference group. Both Panel A and B control for grid cell fixed effects, plant × year fixed effects, grid cell-specific linear trends and a set of geographic controls interacted with year fixed effects. The diamonds denote coefficients and bars denote the associated 95% confidence intervals, where standard errors used to construct the confidence intervals are clustered at the plant level.

Figure 6: Biomass Power Plants: Dynamic Effect on Non-agricultural Fires

A: Number of fires

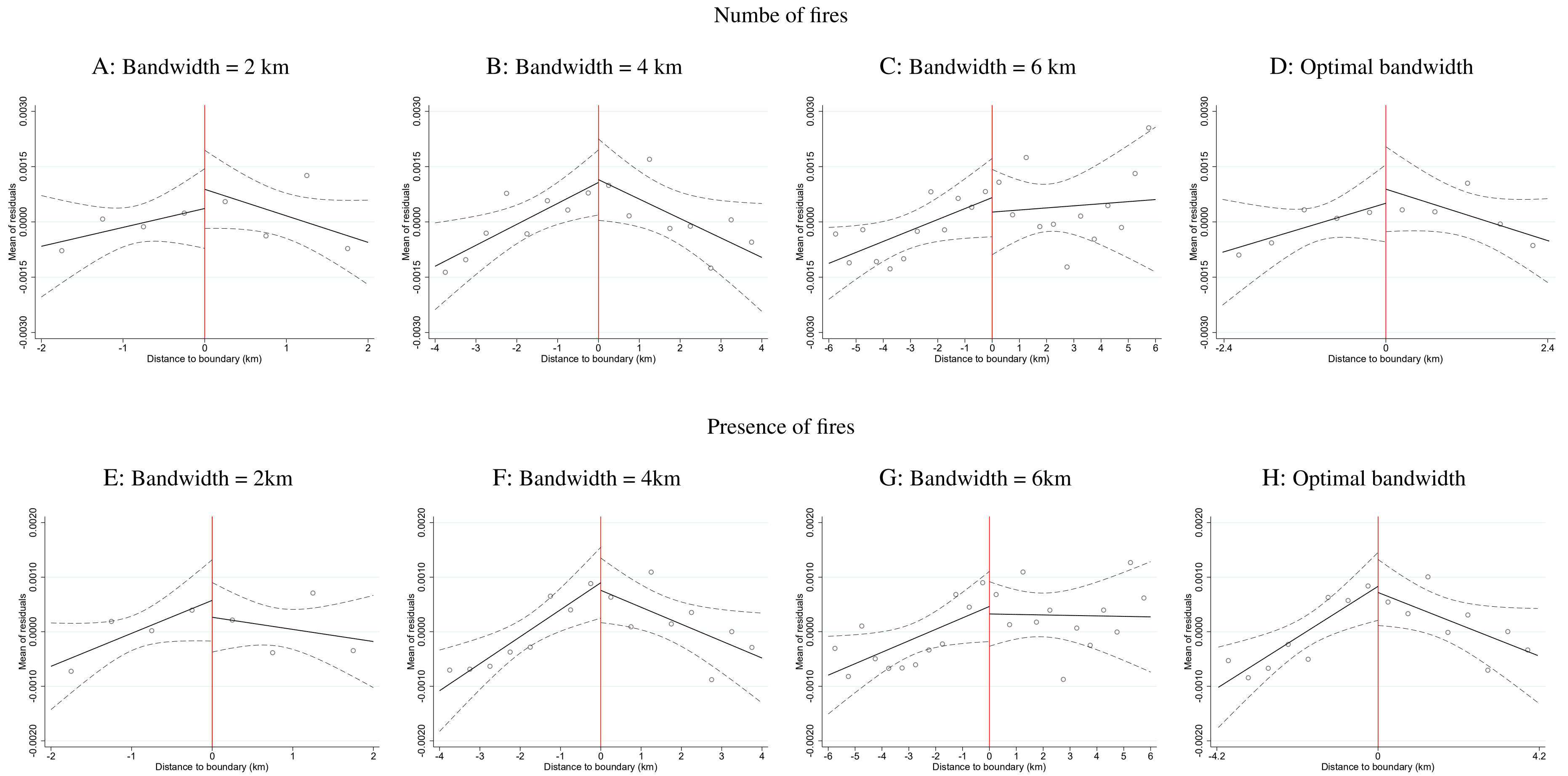


B: Presence of fires



Notes: The unit of observation is 1 km × 1 km grid cells. This figure checks the dynamic effect of biomass power plants on non-agricultural fires by conducting an event study plot. *Presence of fires* is an indicator for observing at least one fire. The year before a plant opening is omitted as the reference group. Both Panel A and B control for grid cell fixed effects, plant × year fixed effects, grid cell-specific linear trends, and a set of geographic controls interacted with year fixed effects. The diamonds denote coefficients and bars denote the associated 95% confidence intervals, where standard errors used to construct the confidence intervals are clustered at the plant level.

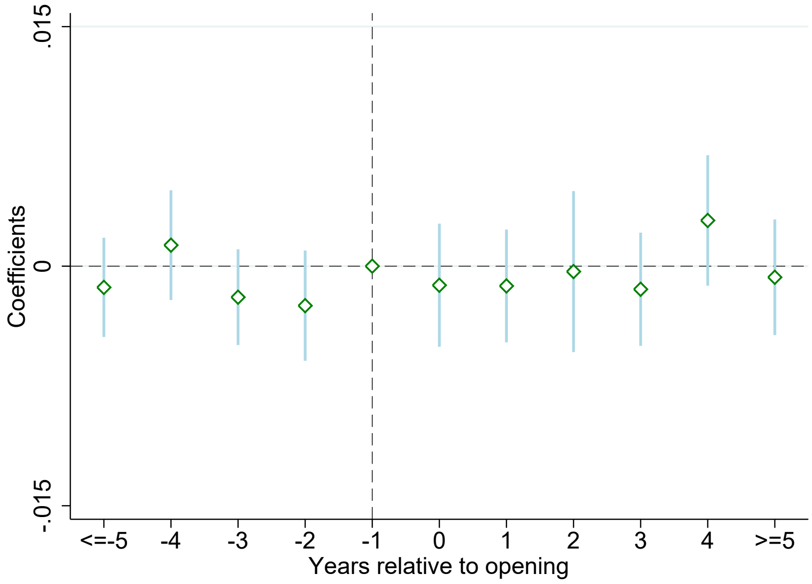
Figure 7: No Burning Zones: Insignificant Effect on Agricultural Fires - RD Plot



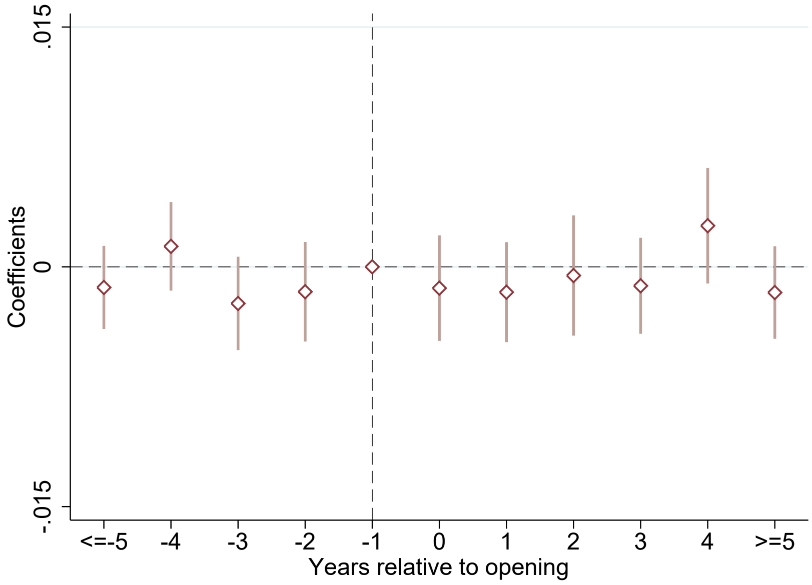
Notes: The unit of observation is 1 km $\times$ 1 km grid cells. The no burning zones are areas within 15 km of an airport. This figure shows that the no burning zones have little effect on agricultural fires. *Presence of fires* is an indicator for observing at least one fire. The circles denote the average number of fires or probability that any fires occur within a distance bin, after partialling out the airport $\times$ boundary segment fixed effects. The lines are fitted separately for grid cells on each side of the boundary, and the dashes are the associated 95% confidence intervals, where standard errors used to construct the confidence intervals are clustered at the airport level. The optimal bandwidth in Pane D and H are selected using the procedure proposed by Imbens and Kalyanaraman (2012).

Figure 8: No Burning Zones: Event Study Plot for Difference-in-discontinuities Estimation

A: Number of fires

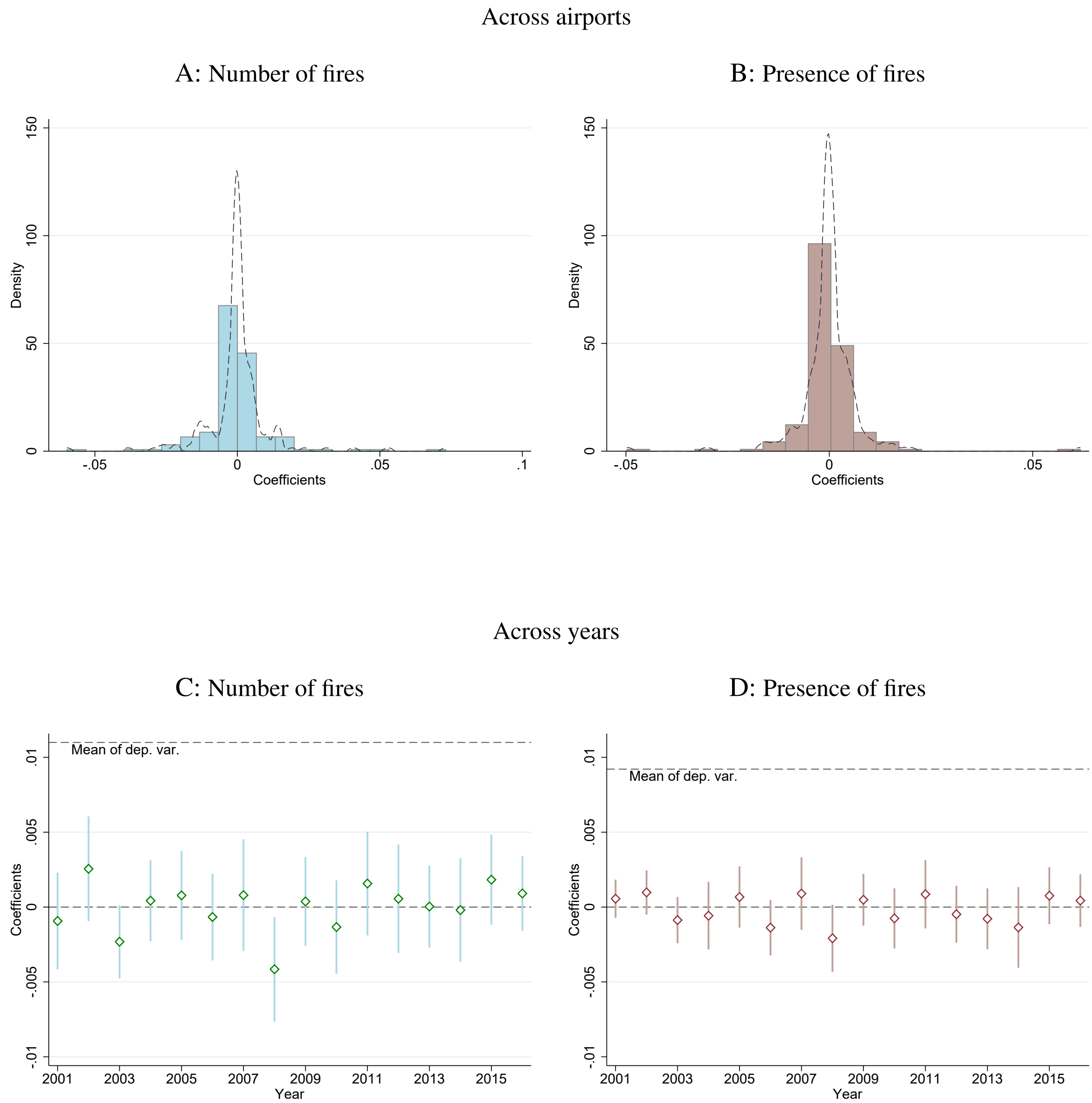


B: Presence of fires



Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This figure conducts an event study plot for the difference-in-discontinuities estimation based on a 4 km bandwidth, utilizing the staggered opening of airports. *Presence of fires* is an indicator for observing at least one fire. The year before an airport opening is omitted as the reference group. The diamonds denote coefficients and bars denote the associated 95% confidence intervals, where standard errors used to construct the confidence intervals are clustered at the airport level.

Figure 9: No Buring Zones: Heterogeneity of RD Estimates



Notes: The unit of observation is 1 km $\times$ 1 km grid cells. The no burning zones are areas within 15 km of an airport. This figure conducts RD estimation separately for each airport and each year, using the optimal bandwidth calculated following the procedure proposed by Imbens and Kalyanaraman (2012), and shows little heterogeneity across airports or over years. *Presence of fires* is an indicator for observing at least one fire. Panel A and B show the distribution of airport-specific RD coefficients. Panel C and D show the year-specific RD coefficients (the diamonds) and the associated 95% confidence intervals (the bars), where standard errors used to construct the confidence intervals are clustered at the airport level.

Table 1: The Effect of Biomass Power Plants on Agricultural Fires

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var. =	Number of fires				Presence of fires			
$<$ 5 km of the plant $\times$ Post	-0.0054**	-0.0085***	-0.0088***	-0.0127***	-0.0033**	-0.0052***	-0.0055***	-0.0065***
	(0.0025)	(0.0020)	(0.0020)	(0.0028)	(0.0014)	(0.0012)	(0.0012)	(0.0015)
5-10 km of the plant $\times$ Post	-0.0027*	-0.0049***	-0.0049***	-0.0068***	-0.0018**	-0.0018**	-0.0019**	-0.0027***
	(0.0015)	(0.0015)	(0.0015)	(0.0018)	(0.0008)	(0.0008)	(0.0008)	(0.0009)
$<$ 5 km of the plant	0.0014				0.0001			
	(0.0022)				(0.0011)			
5-10 km of the plant	0.0003				0.0012**			
	(0.0011)				(0.0006)			
Grid cell FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Plant $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls $\times$ Year FE	No	No	Yes	Yes	No	No	Yes	Yes
Grid cell trends	No	No	No	Yes	No	No	No	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant
Observations	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576
R-squared	0.0187	0.2729	0.2731	0.5196	0.0428	0.1615	0.1617	0.2378
Mean of dep. var.	0.0220	0.0220	0.0220	0.0220	0.0174	0.0174	0.0174	0.0174

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This table shows that biomass power plants can reduce nearby agricultural fires, using a spatial difference-in-differences strategy. *Presence of fires* is an indicator for observing at least one fire. *$<$ 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 2: The Effect of Biomass Power Plants on Agricultural Fires: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Treatment intensity		Controlling for spillovers		Controlling for city shocks		Controlling for 2013 air quality monitoring	
Dep. var. =	# fires	# fires > 0	# fires	# fires > 0	# fires	# fires > 0	# fires	# fires > 0
Distance to the plant × Post	0.0013***	0.0007***						
	(0.0003)	(0.0002)						
< 5 km of the plant × Post			-0.0132***	-0.0065***	-0.0122***	-0.0064***	-0.0127***	-0.0065***
			(0.0033)	(0.0017)	(0.0030)	(0.0015)	(0.0028)	(0.0015)
5-10 km of the plant × Post			-0.0072***	-0.0028***	-0.0067***	-0.0027***	-0.0068***	-0.0027***
			(0.0021)	(0.0011)	(0.0018)	(0.0010)	(0.0018)	(0.0009)
Distance to neighbor plant × Year FE	No	No	Yes	Yes	No	No	No	No
City × Year FE	No	No	No	No	Yes	Yes	No	No
Dist. to city center × Post 2013	No	No	No	No	No	No	Yes	Yes
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant
Observations	1,986,576	1,986,576	1,611,456	1,611,456	1,986,560	1,986,560	1,966,912	1,966,912
R-squared	0.5196	0.2378	0.5341	0.2430	0.5203	0.2392	0.5196	0.2377
Mean of dep. var.	0.0220	0.0174	0.0242	0.0190	0.0220	0.0174	0.0222	0.0176

Notes: The unit of observation is 1 km × 1 km grid cells. This table shows the robustness of the baseline results. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Columns (1)-(2) use a continuous treatment intensity where areas closer to a plant are assumed to have greater treatment. Columns (3)-(4) control for the spillover effects from neighbor plants. Columns (5)-(6) control for shocks at the city level by including city × year fixed effects. Columns (7)-(8) consider the large-scale air quality monitoring program launched in 2013. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 3: The Effect of Biomass Power Plants on Agricultural Fires: Robustness (Cont.)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Controlling for persistence of fires		Clustering at: $2° \times 2°$		Conley standard errors		Placebo: fires in non-cropland	
Dep. var. =	# fires	# fires > 0	# fires	# fires > 0	# fires	# fires > 0	# fires	# fires > 0
< 5 km of the plant × Post	-0.0124***	-0.0065***	-0.0127***	-0.0065***	-0.0127***	-0.0065***	-0.0048	-0.0012
	(0.0025)	(0.0016)	(0.0032)	(0.0014)	(0.0025)	(0.0015)	(0.0037)	(0.0011)
5-10 km of the plant × Post	-0.0059***	-0.0026***	-0.0068***	-0.0027***	-0.0068***	-0.0027***	0.0028	-0.0005
	(0.0015)	(0.0010)	(0.0018)	(0.0010)	(0.0016)	(0.0009)	(0.0046)	(0.0006)
Lagged dep. var. included	Yes	Yes	No	No	No	No	No	No
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	$2° \times 2°$	$2° \times 2°$	-	-	Plant	Plant
Observations	1,862,415	1,862,415	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576
R-squared	0.5551	0.2560	0.5196	0.2378	0.5196	0.2378	0.8725	0.3193
Mean of dep. var.	0.0232	0.0183	0.0220	0.0174	0.0220	0.0174	0.0164	0.0071

Notes: The unit of observation is 1 km × 1 km grid cells. This table shows the robustness of the baseline results. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Columns (1)-(2) control for the persistence of fires by including a one-period lag of fires. Columns (3)-(4) cluster the standard errors at the $2° \times 2°$ level. Columns (5)-(6) adopt the Conley standard errors, assuming spatial correlation for observations within a 200 km radius and temporal correlation across 5 years. Columns (7)-(8) conduct a placebo test by replacing agricultural fires with non-agricultural fires. Standard errors in parentheses are clustered at the plant level in columns (1)-(2) and (7)-(8). * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 4: Biomass Power Plants: Testing the Economic Incentive Mechanism

Dep. var. =	(1) # Fires	(2) # Fires > 0	(3) # Fires	(4) # Fires > 0	(5) # Fires	(6) # Fires > 0
< 5 km of the plant × Post	-0.0138***	-0.0073***	-0.0139***	-0.0073***	-0.0225***	-0.0147***
	(0.0032)	(0.0017)	(0.0032)	(0.0017)	(0.0071)	(0.0036)
5-10 km of the plant × Post	-0.0074***	-0.0029***	-0.0074***	-0.0029***	-0.0183***	-0.0132***
	(0.0020)	(0.0011)	(0.0021)	(0.0011)	(0.0045)	(0.0031)
Slope × < 5 km of the plant × Post	0.0009**	0.0007**				
	(0.0005)	(0.0003)				
Slope × 5-10 km of the plant × Post	0.0004	0.0002				
	(0.0003)	(0.0002)				
Ruggedness × < 5 km of the plant × Post			0.0000**	0.0000**		
			(0.0000)	(0.0000)		
Ruggedness × 5-10 km of the plant × Post			0.0000	0.0000		
			(0.0000)	(0.0000)		
# Fragments created by rivers × Post					-0.0008	-0.0011
					(0.0032)	(0.0018)
# Fragments created by rivers × < 5 km of the plant × Post					0.0088*	0.0074***
					(0.0048)	(0.0026)
# Fragments created by rivers × 5-10 km of the plant × Post					0.0104***	0.0096***
					(0.0034)	(0.0024)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant
Observations	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576
R-squared	0.5196	0.2378	0.5196	0.2378	0.5196	0.2378
Mean of dep. var.	0.0220	0.0174	0.0220	0.0174	0.0220	0.0174

Notes: The unit of observation is 1 km × 1 km grid cells. This table shows that the reduction of agricultural fires after a plant opening is consistent with an economic incentive explanation. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Columns (1)-(2) and (3)-(4) show smaller effects in areas with greater difficulty in collecting crop residues, where difficulty is measured by slope and ruggedness, respectively. Columns (5)-(6) also show smaller effects in areas with greater difficulty in collecting crop residues, where difficulty is measured by the number of disconnected fragments created by rivers in a grid cell. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 5: Biomass Power Plants: Testing the Structural Transformation Mechanism

	(1)	(2)	(3)	(4)	(5)	(6)
	Land usage		Industrial sector dynamics			
Dep. var. =	% Crop	% Settlement	# Firms	# Employees	log(Capital+1)	log(Output+1)
$<$ 5 km of the plant $\times$ Post	0.0006	-0.0009	0.0555	1.9697	0.0585	0.0274
	(0.0010)	(0.0013)	(0.0820)	(14.7630)	(0.0453)	(0.0734)
5-10 km of the plant $\times$ Post	-0.0001	-0.0005	-0.0143	-5.5121	0.0522	0.0655
	(0.0006)	(0.0007)	(0.0734)	(10.6314)	(0.0376)	(0.0444)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes
Plant $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant
Observations	1,986,576	1,986,576	869,127	869,127	869,127	744,966
R-squared	0.9943	0.9810	0.8010	0.8452	0.8676	0.8859
Mean of dep. var.	0.734	0.0761	0.203	42.32	0.474	0.548

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This table shows that the reduction of agricultural fires after a plant opening is not caused by changes of economic structure. *$<$ 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 6: Biomass Power Plants: Testing the Migration Mechanism

	(1)	(2)	(3)	(4)	(5)	(6)
	Downwind is measured as:				Road access:	
	% days downwind of the plant in a year		1 (% days downwind > plant-year median)		1(within 1 km of railways or major roads)	
Dep. var. =	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0
< 5 km of the plant × Post	-0.0146***	-0.0067***	-0.0139***	-0.0066***	-0.0107***	-0.0066***
	(0.0048)	(0.0024)	(0.0039)	(0.0017)	(0.0025)	(0.0016)
5-10 km of the plant × Post	-0.0116***	-0.0036**	-0.0084***	-0.0028**	-0.0052***	-0.0021**
	(0.0034)	(0.0018)	(0.0024)	(0.0012)	(0.0019)	(0.0010)
Downwind	-0.0006	-0.0017	0.0009	0.0003		
	(0.0045)	(0.0030)	(0.0013)	(0.0008)		
Downwind × Post	-0.0083	0.0021	-0.0016	0.0005		
	(0.0104)	(0.0044)	(0.0025)	(0.0012)		
Downwind × < 5 km of the plant	-0.0069	-0.0023	-0.0021	-0.0007		
	(0.0083)	(0.0051)	(0.0025)	(0.0015)		
Downwind × 5-10 km of the plant	-0.0028	-0.0011	-0.0024	-0.0012		
	(0.0060)	(0.0040)	(0.0016)	(0.0011)		
Downwind × < 5 km of the plant × Post	0.0071	0.0006	0.0023	0.0001		
	(0.0123)	(0.0068)	(0.0037)	(0.0019)		
Downwind × 5-10 km of the plant × Post	0.0196*	0.0034	0.0032	0.0002		
	(0.0100)	(0.0058)	(0.0025)	(0.0016)		
Road access × Post					0.0108	0.0025
					(0.0067)	(0.0018)
Road access × < 5 km of the plant × Post					-0.0114	-0.0002
					(0.0090)	(0.0033)
Road access × 5-10 km of the plant × Post					-0.0095	-0.0031
					(0.0060)	(0.0027)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant
Observations	1,963,011	1,963,011	1,963,011	1,963,011	1,986,576	1,986,576
R-squared	0.5215	0.2387	0.5215	0.2387	0.5196	0.2378
Mean of dep. var.	0.0219	0.0173	0.0219	0.0173	0.0220	0.0174

Notes: The unit of observation is 1 km × 1 km grid cells. This table shows that the reduction of agricultural fires after a plant opening is not caused by potential pollution-induced migration. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Columns (1)-(4) show that the effects are not larger in downwind areas of a plant that are perceived to suffer more from potential pollution. Columns (5)-(6) show no differential effects in areas with greater ease of migration, proxied by access to transportation networks. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 7: Biomass Power Plants: Testing the Enhanced Inspection Mechanism

	(1)	(2)	(3)	(4)	(5)	(6)
			Other city: 1(grid cell and the plant are located in different cities)		High incentive: 1(city party secretary is younger than 57)	
Dep. var. =	# Fires	# Fires > 0	# Fires	# Fires > 0	# Fires	# Fires > 0
< 5 km of the plant × Post	-0.0145***	-0.0072***	-0.0059***	-0.0120***	-0.0102***	-0.0064***
	(0.0049)	(0.0022)	(0.0014)	(0.0030)	(0.0037)	(0.0024)
5-10 km of the plant × Post	-0.0099***	-0.0044***	-0.0024**	-0.0068***	-0.0055**	-0.0021*
	(0.0035)	(0.0015)	(0.0010)	(0.0019)	(0.0026)	(0.0013)
Distance to city center × Post	-0.0001	-0.0000				
	(0.0001)	(0.0000)				
Distance to city center × < 5 km of the plant × Post	0.0000	0.0000				
	(0.0001)	(0.0000)				
Distance to city center × 5-10 km of the plant × Post	0.0001	0.0000				
	(0.0000)	(0.0000)				
Other city × Post			0.0029	0.0043		
			(0.0031)	(0.0066)		
Other city × < 5 km of the plant × Post			-0.0085	-0.0104		
			(0.0100)	(0.0131)		
Other city × 5-10 km of the plant × Post			-0.0022	0.0027		
			(0.0060)	(0.0074)		
High incentive					0.0041	0.0034
					(0.0031)	(0.0021)
High incentive × Post					-0.0016	-0.0031
					(0.0046)	(0.0029)
High incentive × < 5 km of the plant					0.0005	-0.0002
					(0.0022)	(0.0014)
High incentive × 5-10 km of the plant					0.0027	0.0013
					(0.0025)	(0.0010)
High incentive × < 5 km of the plant × Post					-0.0034	0.0004
					(0.0034)	(0.0022)
High incentive × 5-10 km of the plant × Post					-0.0018	-0.0005
					(0.0029)	(0.0014)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant
Observations	1,966,912	1,966,912	1,986,576	1,986,576	1,858,853	1,858,853
R-squared	0.5196	0.2377	0.2378	0.5196	0.5248	0.2414
Mean of dep. var.	0.0220	0.0174	0.0174	0.0220	0.0220	0.0174

Notes: The unit of observation is 1km × 1km grid cells. This table shows that the reduction of agricultural fires after a plant opening is not caused by enhanced inspection by local officials following the arrival of the plant. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Columns (1)-(2) show no differential effects in areas with greater penetration of local officials (areas closer to city center). Columns (3)-(4) show no differential effects in areas outside the city where the plant is located. Columns (5)-(6) show no differential effects in areas where the city party secretary has greater promotion incentives (younger than 57). Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 8: The Effect of No Burning Zones on Agricultural Fires

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var. =	Number of fires			Presence of fires		
	Panel A: Bandwidth = 2 km					
Inside	0.0004	0.0006	0.0005	-0.0004	-0.0002	-0.0003
	(0.0008)	(0.0007)	(0.0007)	(0.0005)	(0.0005)	(0.0005)
Observations	914,902	914,902	914,902	914,902	914,902	914,902
Mean of dep. var.	0.0115	0.0115	0.0115	0.0097	0.0097	0.0097
	Panel B: Bandwidth = 4 km					
Inside	0.0000	0.0001	0.0001	-0.0002	-0.0001	-0.0001
	(0.0007)	(0.0007)	(0.0007)	(0.0004)	(0.0004)	(0.0004)
Observations	1,829,849	1,829,849	1,829,849	1,829,849	1,829,849	1,829,849
Mean of dep. var.	0.0110	0.0110	0.0110	0.0092	0.0092	0.0092
	Panel C: Bandwidth = 6 km					
Inside	-0.0004	-0.0004	-0.0004	-0.0002	-0.0001	-0.0001
	(0.0008)	(0.0008)	(0.0008)	(0.0004)	(0.0004)	(0.0004)
Observations	2,743,315	2,743,315	2,743,315	2,743,315	2,743,315	2,743,315
Mean of dep. var.	0.0109	0.0109	0.0109	0.0092	0.0092	0.0092
	Panel D: Optimal bandwidth (km)					
Inside	0.0007	0.0008	0.0004	-0.0001	-0.0001	-0.0001
	(0.0007)	(0.0007)	(0.0007)	(0.0004)	(0.0004)	(0.0004)
Observations	1,030,204	1,016,363	1,107,607	1,790,391	1,799,071	1,902,847
Mean of dep. var.	0.0114	0.0115	0.0114	0.0092	0.0092	0.0092
Bandwidth	2.2541	2.2222	2.4227	3.9146	3.9343	4.1610
RD polynomial	Linear	Linear	Linear	Linear	Linear	Linear
Airport FE	No	Yes	Yes	No	Yes	Yes
Airport × Segment FE	No	No	Yes	No	No	Yes
Cluster level	Airport	Airport	Airport	Airport	Airport	Airport

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table shows that no burning zones do not reduce agricultural fires, by conducting a spatial RD estimation around the no burning zone boundaries. *Presence of fires* is an indicator for observing at least one fire. *Inside* is an indicator for areas within the no burning zones. The optimal bandwidth in Panel D is selected using the procedure proposed by Imbens and Kalyanaraman (2012). Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 9: The Effect of No Burning Zones on Agricultural Fires: Testing for Sorting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Checking the land usage and economic structure						Donut hole specification			Ease of migration:			Only airports built before 1999	
									Hard	Easy	Hard	Easy		
Dep. var. =	% Cropland	% Settlement land	# Firms	# Employees	log(1+capital)	log(1+output)	# Fires	# Fires > 0	# Fires	# Fires	# Fires > 0	# Fires > 0	# Fires	# Fires > 0
Panel A: Bandwidth = 2km														
Inside	0.0012	-0.0009	0.0485**	9.1964*	-0.0121	-0.0136	0.0014	0.0006	0.0000	0.0033	-0.0006	0.0012	0.0005	-0.0003
	(0.0034)	(0.0016)	(0.0241)	(5.5490)	(0.0266)	(0.0316)	(0.0013)	(0.0007)	(0.0007)	(0.0027)	(0.0005)	(0.0014)	(0.0010)	(0.0006)
Observations	914,902	914,902	327,552	327,552	327,552	280,801	915,503	915,503	757,879	157,020	757,879	157,020	615,136	615,136
Mean of dep. var.	0.4670	0.0686	0.2330	48.6800	0.5220	0.5970	0.0111	0.0092	0.0106	0.0158	0.0090	0.0128	0.0126	0.0106
Panel B: Bandwidth = 4km														
Inside	0.0002	0.0002	0.0189	2.5191	-0.0085	-0.0116	0.0001	0.0005	-0.0006	0.0035	-0.0004	0.0015	0.0004	0.0002
	(0.0039)	(0.0018)	(0.0225)	(5.1880)	(0.0249)	(0.0285)	(0.0011)	(0.0005)	(0.0008)	(0.0024)	(0.0005)	(0.0013)	(0.0009)	(0.0005)
Observations	1,829,849	1,829,849	655,155	655,155	655,155	561,664	1,828,918	1,828,918	1,514,760	315,088	1,514,760	315,088	1,230,096	1,230,096
Mean of dep. var.	0.4680	0.0686	0.2310	47.3100	0.5090	0.5830	0.0107	0.0090	0.0101	0.0151	0.0086	0.0121	0.0119	0.0100
Panel C: Bandwidth = 6km														
Inside	-0.0001	0.0014	0.0047	1.8606	-0.0148	-0.0189	-0.0013	-0.0001	-0.0004	0.0013	-0.0002	0.0010	0.0003	0.0002
	(0.0041)	(0.0022)	(0.0223)	(4.3612)	(0.0238)	(0.0272)	(0.0013)	(0.0005)	(0.0008)	(0.0025)	(0.0005)	(0.0011)	(0.0009)	(0.0005)
Observations	2,743,315	2,743,315	982,154	982,154	982,154	842,002	2,742,818	2,742,818	2,270,713	472,600	2,270,713	472,600	1,844,016	1,844,016
Mean of dep. var.	0.4680	0.0679	0.2260	46.4000	0.5060	0.5790	0.0110	0.0091	0.0100	0.0155	0.0085	0.0123	0.0117	0.0100
Panel D: Optimal bandwidth (km)														
Inside	0.0008	0.0002	0.0290	-0.4042	-0.0207	-0.0211	0.0002	0.0001	-0.0005	0.0030	-0.0004	0.0012	0.0008	-0.0000
	(0.0036)	(0.0014)	(0.0232)	(4.9267)	(0.0260)	(0.0288)	(0.0011)	(0.0005)	(0.0008)	(0.0023)	(0.0005)	(0.0011)	(0.0010)	(0.0006)
Observations	1,175,095	1,320,551	500,061	549,861	844,202	752,374	1,454,386	2,399,172	1,311,801	363,877	1,508,651	459,522	945,760	1,100,704
Mean of dep. var.	0.4668	0.0689	0.2292	46.9729	0.5106	0.5816	0.0108	0.0090	0.0102	0.0150	0.0086	0.0123	0.0122	0.0101
Bandwidth	2.5692	2.8894	3.0525	3.3595	5.1565	5.3616	3.1809	5.2457	3.4649	4.6211	3.9845	5.8391	3.0770	3.5788
RD polynomial	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear
Airport × Segment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table checks potential sorting to validate the RD estimation around the no burning zone boundaries. *Inside* is an indicator for areas within the no burning zones. Columns (1)-(2) rule out the sorting of land usage. Columns (3)-(6) rule out the sorting of firms. Columns (7)-(8) rule out the sorting of farmers by excluding areas within 1 km of the no burning zone boundaries. Columns (9)-(12) rule out the sorting of farmers by looking at areas with differential ease of migration. Columns (13)-(14) rule out the sorting of airports by looking at preexisting airports. The optimal bandwidth in Panel D is selected using the procedure proposed by Imbens and Kalyanaraman (2012). Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 10: The Effect of No Burning Zones on Agricultural Fires: Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Difference-in-Discontinuities analysis				Alternative number of boundary segments				% Cropland > 10%		Major grain provinces	
Dep. var. =	#Fires	#Fires	#Fires>0	#Fires>0	#Fires	#Fires	#Fires>0	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0
Inside	0.0006		0.0003		0.0007	0.0004	-0.0001	-0.0001	0.0003	-0.0001	0.0014	-0.0002
	(0.0005)		(0.0004)		(0.0008)	(0.0007)	(0.0004)	(0.0004)	(0.0008)	(0.0005)	(0.0018)	(0.0009)
Inside × Post	0.0002	0.0006	-0.0004	0.0001								
	(0.0009)	(0.0008)	(0.0006)	(0.0005)								
RD polynomial	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear
RD polynomial × Post	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No
Airport × Segment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell FE	No	Yes	No	Yes	No	No	No	No	No	No	No	No
Airport × Year FE	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No
Number of segments	10	10	10	10	5	5	20	20	10	10	10	10
Cluster level	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport
Observations	1,110,368	1,110,368	1,110,368	1,110,368	1,021,851	1,089,873	1,778,652	1,876,313	970,182	1,648,696	492,724	674,229
Mean of dep. var.	0.0068	0.0068	0.0060	0.0060	0.0115	0.0114	0.0093	0.0092	0.0129	0.0105	0.0192	0.0153
Bandwidth (km)	4	4	4	4	2.2353	2.3849	3.8903	4.1023	2.4165	4.1061	2.5869	3.5339

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table checks the robustness of the RD estimation around the no burning zone boundaries. *Inside* is an indicator for areas within the no burning zones. Columns (1)-(4) use a difference-in-discontinuities specification utilizing the staggered opening of airports. Columns (5)-(8) use alternative number of boundary segments. Columns (9)-(10) use only airports around which the share of cropland accounts for more than 10%. Columns (11)-(12) focus on airports located in major grain-producing provinces. Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 11: The Effect of No Burning Zones on Agricultural Fires: Robustness (Cont.)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	2-dimension RD		Triangular kernel		Quadratic polynomial		Geo. covariates		All covariates		Cluster: 2°×2°		Conley SE	
Dep. var. =	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0
Inside	0.0003	0.0003	0.0001	-0.0004	0.0006	0.0002	0.0003	-0.0001	-0.0006	-0.0001	0.0004	-0.0001	0.0004	-0.0001
	(0.0006)	(0.0004)	(0.0008)	(0.0004)	(0.0008)	(0.0005)	(0.0007)	(0.0004)	(0.0008)	(0.0004)	(0.0006)	(0.0004)	(0.0006)	(0.0003)
RD polynomial	Linear	Linear	Linear	Linear	Quadratic	Quadratic	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear
RD dimension	2	2	1	1	1	1	1	1	1	1	1	1	1	1
RD kernel	Uniform	Uniform	Triangle	Triangle	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform	Uniform
Geo. controls	No	No	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No
All controls	No	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No
Airport × segment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	2°×2°	2°×2°	-	-
Observations	1,829,849	1,829,849	1,409,705	2,420,320	2,223,090	3,002,163	1,100,422	1,876,714	672,077	762,225	1,107,607	1,902,847	1,107,607	1,902,847
Mean of dep. var.	0.0110	0.0092	0.0112	0.0092	0.0109	0.0092	0.0114	0.0092	0.0080	0.0066	0.0114	0.0092	0.0114	0.0092
Bandwidth (km)	4	4	3.0822	5.2938	4.8627	6.5673	2.4063	4.1029	4.1038	4.6593	2.4227	4.1610	2.4227	4.1610

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table checks the robustness of the RD estimation around the no burning zone boundaries. *Inside* is an indicator for areas within the no burning zones. Columns (1)-(2) use a two-dimensional RD polynomial in latitude and longitude. Columns (3)-(4) use a triangular kernel, giving more weights to observations closer to the boundaries. Columns (5)-(6) use quadratic RD polynomial in distance to the boundaries. Columns (7)-(10) include the covariates used in balance checks. Columns (11)-(12) cluster the standard errors at the 2° × 2° level. Columns (13)-(14) adopt the Conley standard errors, assuming spatial correlation for observations within a 200 km radius and temporal correlation across 5 years. Standard errors in parentheses in columns (1)-(10) are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table 12: The Effect of No Burning Zones on Agricultural Fires: Heterogeneity

	(1) Upwind area	(2) Non-upwind area	(3) Provincial city	(4) Non-prov. city	(5) High incentive	(6) Low incentive	(7) Upwind area	(8) Non-upwind area	(9) Provincial city	(10) Non-prov. city	(11) High incentive	(12) Low incentive
Dep. var. =	Number of fires						Presence of fires					
Panel A: Bandwidth = 2 km												
Inside	-0.0002	0.0011	0.0018	0.0003	-0.0000	0.0030	-0.0008	-0.0000	-0.0015	-0.0001	-0.0011**	0.0021
	(0.0011)	(0.0010)	(0.0017)	(0.0008)	(0.0008)	(0.0020)	(0.0006)	(0.0007)	(0.0011)	(0.0005)	(0.0005)	(0.0014)
Observations	427,125	455,768	131,710	783,192	629,566	141,339	427,125	455,768	131,710	783,192	629,566	141,339
Mean of dep. var.	0.0117	0.0111	0.0171	0.0106	0.0113	0.0193	0.0099	0.0092	0.0135	0.0090	0.0096	0.0159
Panel B: Bandwidth = 4 km												
Inside	-0.0010	0.0012	0.0010	-0.0001	0.0002	0.0014	-0.0009	0.0005	-0.0004	-0.0001	-0.0004	0.0018
	(0.0010)	(0.0010)	(0.0022)	(0.0008)	(0.0008)	(0.0019)	(0.0006)	(0.0006)	(0.0012)	(0.0004)	(0.0005)	(0.0011)
Observations	854,726	910,774	261,819	1,568,030	1,258,846	281,844	854,726	910,774	261,819	1,568,030	1,258,846	281,844
Mean of dep. var.	0.0113	0.0105	0.0159	0.0102	0.0109	0.0180	0.0094	0.0088	0.0125	0.0087	0.0092	0.0148
Panel C: Bandwidth = 6 km												
Inside	-0.0013	0.0005	-0.0001	-0.0004	-0.0001	-0.0004	-0.0008	0.0005	-0.0007	-0.0000	-0.0002	0.0012
	(0.0010)	(0.0011)	(0.0032)	(0.0007)	(0.0008)	(0.0020)	(0.0006)	(0.0006)	(0.0015)	(0.0004)	(0.0005)	(0.0011)
Observations	1,282,559	1,364,611	390,642	2,352,673	1,888,507	421,521	1,282,559	1,364,611	390,642	2,352,673	1,888,507	421,521
Mean of dep. var.	0.0110	0.0105	0.0155	0.0102	0.0109	0.0178	0.0093	0.0088	0.0124	0.0087	0.0092	0.0147
Panel D: Bandwidth = Optimal bandwidth (km)												
Inside	-0.0009	0.0018	0.0023	0.0003	0.0007	0.0039*	-0.0007	0.0006	-0.0008	0.0001	-0.0001	0.0017
	(0.0010)	(0.0011)	(0.0017)	(0.0008)	(0.0009)	(0.0021)	(0.0006)	(0.0006)	(0.0012)	(0.0005)	(0.0005)	(0.0011)
Observations	881,969	557,823	215,318	902,363	993,730	197,732	953,146	972,217	266,888	2,014,149	1,571,582	262,851
Mean of dep. var.	0.0112	0.0110	0.0165	0.0104	0.0110	0.0192	0.0093	0.0088	0.0125	0.0086	0.0092	0.0149
Bandwidth	4.1242	2.4510	3.2760	2.3074	3.1583	2.8007	4.4602	4.2723	4.0744	5.1380	4.9945	3.7288
RD polynomial	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear
Airport × Segment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table checks the heterogeneity of the RD estimation around the no burning zone boundaries. *Presence of fires* is an indicator for observing at least one fire. *Inside* is an indicator for areas within the no burning zones. *Upwind area* denotes grid cells where the percentage of days upwind of the airport in a year is greater than the median percentage for that airport and year. *Provincial city* denotes provincial capitals and other cities which share roughly the same political status as provincial capitals. *High incentive* denotes that the city party secretary is younger than 57, as local leaders older than 57 have very low probability of being promoted. The optimal bandwidth in Panel D is selected using the procedure proposed by Imbens and Kalyanaraman (2012). Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Online Appendix

Incentives, Penalties, and Rural Air Pollution: Evidence from Satellite Data

Yongwei Nian

Contents

A Determining treatment and control group **203**

B Staggered Treatment Correction **204**

C Topography-Based Distance Measure **205**

D Additional Selection Issues **207**

E Additional Data Issues **209**

E.1 Data Quality 209

E.2 Map Censoring 210

E.3 Sample Selection 213

E.4 Treatment Timing 214

F Additional Mechanism Tests **215**

F.1 Qualitative Evidence on Economic Incentives 215

F.2 Quantitative Evidence on Economic Incentives 218

F.3 Changes in Small Firms 220

F.4 Cropland Difference 221

F.5 Farmer Adaption to Plants 223

F.6 Potential Fire Bans under Transmission Towers 224

G Other City-level Regulations **225**

H Air and Water Pollution **227**

H.1 Air Pollution 227

H.2 Water Pollution 230

I Comparing Plants to Airports **231**

J Difference-in-differences Estimation around Airports **232**

K Appendix Figures and Tables **239**

A Determining treatment and control group

To bolster the choice of treatment and control group in the baseline specification, I estimate a flexible variant of equation (1). Specifically, I focus on areas within 55 km of a plant and divide them into 11 distance bins, which have the same width of 5 km. Namely, bin 1= [0 km, 5 km), bin 2=[5 km, 10 km), $\cdots$, and bin 11=[50 km, 55 km). Bin 11 will be omitted as the reference group, which is motivated by the fact that 50 km is a sufficiently large catchment radius for a typical large plant.[1] I then estimate a DID specification as follows:

$$Y_{ijt} = \delta_i + \lambda_{jt} + \beta^k \sum_{k=1}^{10} DistanceBin_{ij}^k \times Post_{jt} + \varepsilon_{ijt} \tag{A1}$$

where the unit of observation is 1 km $\times$ 1 km grid cells. Y_{ijt} denotes the number of agricultural fires in grid cell i around plant j at year t. δ_i denotes grid cell fixed effects, controlling for time-invariant regional characteristics at the grid cell level such as topography. λ_{jt} denotes plant $\times$ year fixed effects, controlling for time-varying factors affecting all the grid cells around a plant such as weather conditions, plant performance, and local economic prosperity. $DistanceBin_{ij}^k$ is an indicator equal to 1 if the distance from the centroid of grid cell i to plant j is within the kth bin, and 0 otherwise. $Post_{jt}$ is an indicator equal to 1 for years after the opening of plant j and 0 otherwise. In line with the baseline specification, I cluster the standard errors at the plant level (Bertrand, Duflo and Mullainathan, 2004; Cameron and Miller, 2015).

The coefficient β^k measures the change in agricultural fires in areas within the kth bin after a plant opening, relative to areas within the 11th bin. I plot the results in Appendix Figure A7. The effect is the largest for the first bin that is the closest to the plant and fades away quickly with distance. For the third bin and all farther bins, the effect is small and statistically insignificant.[2] Therefore, for the remaining analysis, I focus only on areas within the first three bins (0-15 km). Specifically, I define areas within the first bin (0-5 km) as the most treated areas and areas within the second bin (5-10 km) as the less treated areas. Areas within the third bin

[1] According to a report on these plants by World Bank, a typical large plant only requires 2-5 percent of the agricultural land within a 50 km radius to provide the biomass (see https://openknowledge.worldbank.org/handle/10986/2897). The report also recommends a maximum catchment radius of 50 km for such plants.

[2] This does not necessarily mean that a plant can only utilize crop residues up to 10 km. Instead, this suggests that only farmers within a 10 km radius change burning behavior in response to the plant opening.

(10-15 km) serve as the control group.

B Staggered Treatment Correction

Recent econometric literature shows that the estimates from conventional staggered DID designs, such as those using equations with a treatment indicator and two-way fixed effects, are not easily interpretable (e.g., Callaway and Sant'Anna, 2020; de Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021). The key concern is that these estimates equal the weighted average of all possible 2×2 DID estimates (two periods, one treated and one control group). When treatment effects vary over time, one may get opposite signs compared to the true average treatment effects. The intuition is that some early-treated units could serve as controls for late-treated units. The changes in the outcomes of these early-treated units, which are possibly much larger, are then subtracted from those late-treated units and generate estimates with misleading signs. My DID design, however, can largely alleviate such issue as I include plant $\times$ year fixed effects to ensure that I am doing comparisons around the same plant, instead of comparing plants with differential treatment timing.

Nevertheless, it may still be difficult to interpret the baseline estimates as the weights on each 2×2 DID estimate may not be proportional to the size of that subsample. I therefore follow Callaway and Sant'Anna (2020) to create more interpretable estimates.[3] Specifically, I first classify plants into cohorts, where plants opening in the same year belong to the same cohort. I then estimate equation (1) separately for each cohort. The estimates from each cohort-specific regression represent the average effect of receiving the treatment among units in each cohort. Because there is no variation in treatment timing within each cohort, the cohort-specific estimates do not suffer from any interpretation issue. Then I average the cohort-specific estimates by the size of each cohort, which is the number of plants in each cohort, to summarize the overall average effect of plant openings. In this respect, the final estimates I get have the same interpretation as the ATT in the canonical 2×2 DID setup. The results are presented in Appendix Figure A8. The estimates are similar to the baseline estimates, suggesting that

[3]The estimators proposed by de Chaisemartin and d'Haultfoeuille (2020) cannot be readily used in my setting as I have multiple treatments in one equation.

staggered treatment timing is not a major concern in my setting.

C Topography-Based Distance Measure

Although straight line distance could measure the most fundamental cost of moving across a space, topographic terrain features of the space may also be important. For example, movement on flat terrain is much easier and faster than on uphill terrain. To take into consideration the additional cost induced by terrain impedance, I construct a topography-based distance, which measures the least cost distance from a grid cell to the nearest plant, along three steps.[4] The first step is to generate a cost surface; that is, one needs to assign a value denoting the cost of crossing each grid cell planimetrically. I assign a per unit distance cost of 1 to grid cells containing no rivers. As rivers are rarely utilized for transportation purpose in rural China, they are viewed as an impediment to movement.[5] Hence, I assign a per unit distance cost of 2 to grid cells containing rivers and check robustness to a larger value of 5. These values are taken from Gillings, Hacıgüzeller and Lock (2020) who collect terrain cost factors from various sources.

The next step is to impose penalties for ascents and descents when the movement involves elevation changes. I adopt the slope-dependent cost function from Tobler (1993), which is derived from empirical data on walking across mountainous areas. This function is widely used in routing and other geographic analyses. The dashed line in Appendix Figure A9 shows this function, where negative slopes indicate descents.[6] The cost for zero slope is normalized to 1. In general, it shows a larger cost for moving upward (or downward) than moving on a level surface.[7] The cost is asymmetric about zero slope as walking on a slightly downward sloping surface is generally easier than on a level surface. I also adopt another cost function from Pingel

[4]The construction can be done readily in ArcGIS using the *Path Distance* tool, although one still needs to prescribe some key parameters.

[5]According to various household surveys conducted in recent years that recorded travel modes in the rural areas, water transportation was rare (e.g., see the China Household Finance Survey at https://chfs.swufe.edu.cn/). Rural rivers are generally clogged by sands and agricultural wastes, and some even have dried up. Hence, they are not suitable for ships to pass through. There are many reasons for this, such as lack of maintenance, illegal sand mining, and rapid expansion of urbanization and road system (see an official report about the current issues with the rural water system at http://slt.ah.gov.cn/xwzx/ztzl/zhz/gzdt/120249971.html).

[6]Slopes steeper than 20 degrees are clipped in the figure, as the slope between two neighboring grid cells never exceeds 20 degrees in my data.

[7]The generally larger cost for moving downward than moving on a level surface is because one needs to be more cautious and exert more body control when moving downward.

(2010), which is derived from driving data using modern vehicles. This may be more relevant for measuring the cost of transporting crop residues, although hand carts are still not rare in the rural areas. The cost function is denoted by the solid line in Appendix Figure A9. Compared to the walking cost function, it is more symmetric about zero degree. Also, the cost function is much steeper as vehicles are more sensitive to slopes.

Finally, the cost distance of moving from one grid cell a to one of its neighbors b can be defined as follows:

$$CostDistance_{ab} = \frac{1}{2}(Cost_a^{unit} + Cost_b^{unit}) \times SurfaceDistance_{ab} \times CostFunction_{ab} \tag{A2}$$

$Cost_a^{unit}$ and $Cost_b^{unit}$ denote the unit cost of moving planimetrically through grid cell a to b, respectively; $SurfaceDistance_{ab}$ measures the distance traveled from the center of grid cell a to the center of b, accounting for diagonal movement and elevation changes; $CostFunction_{ab}$ denotes the penalties for ascents or descents from grid cell a to b, using the two aforementioned cost functions. After repeating this calculation for each grid cell, one can generate the least cost distance from each grid cell to the grid cell where the nearest plant is located using the well-known Dijkstra's algorithm (Dijkstra et al., 1959). Appendix Figure A10 shows the least cost distance for grid cells around a typical plant located in the North China Plain, one of China's major grain-producing areas, using different cost functions and river costs. The least cost distance gradient is insensitive to cost function and river cost choice; it generally increases as one moves away from the plant. This is expected as the terrain is unlikely to change dramatically within a small area. Such pattern also lends some support to the use of straight line distance in my main empirical specifications, as the least cost distance gradient can be largely captured by straight line distance. Appendix Table A6 estimates a DID specification similar to equation (6). The coefficients are all positive and significant, suggesting that the plant generates a smaller effect on burning in areas with higher least cost distance. Furthermore, the estimates are quantitatively stable across different measures of terrain and river costs.

D Additional Selection Issues

Selection on crop residues There is a concern that the plants may select into locations with a high amount of crop residues. Such location characteristic may further correlate with fires and confound my findings. To address such concern, I use a global gridded dataset on agricultural production from Yu et al. (2020) to estimate and control for the amount of crop residues in each area. The data record the harvested area and yields per unit area for almost all crops at the 5 arcmin $\times$ 5 arcmin grid cell level for 2000, 2005 and 2010.[8] I focus on the initial year 2000 and match such data to the data on the straw to grain ratio for each crop in different regions in China, collected from the Ministry of Agriculture of China.[9] The amount of crop residues in a grid cell then can be expressed as follows:

$$CropResidues_i^{2000} = \sum_j HarvestedArea_{ij} \times Yield_{ij} \times StrawGrainRatio_{ij} \tag{A3}$$

where i and j denote grid cell and crop type, respectively. The distribution of crop residues in 2000 calculated from this formula is shown in Appendix Figure A12. I then reestimate the baseline regressions by adding the interaction between the amount of crop residues in 2000 and year fixed effects. In this way, I can flexibly control for the selection on the amount of crop residues. As shown in Appendix Table A7, the results are virtually identical to the baseline results.

Selection on land ownership Additionally, the plants may also affect land availability and select into areas with concentrated land ownership for better coordination. Concentrated land ownership may imply fewer fires as in this case it would be less costly to collect residues due to economies of scale. To gauge to what extent the plants may affect land availability, I collect a technical report assessing a typical plant's environmental impacts.[10] The report should be

[8] 5 arcmin $\times$ 5 arcmin is about 7.2 km $\times$ 9 km at the center of China.

[9] See http://www.moa.gov.cn/nybgb/2019/201902/201905/t20190518_6309472.htm. One caveat is that such data cover fewer types of crops than the gridded dataset on agricultural production, but the major crops in China are included.

[10] See http://www.nantong.gov.cn/jcms/jcms_files/jcms1/web14/site/attach/0/160429135950571.pdf. The plant is located in East China, one of China's major grain-producing areas, and has a typical size of 30 MW.

representative as the plants are quite homogeneous.[11] I find no evidence that the plant could affect land availability in areas several hundred meters away.[12] Furthermore, as China strictly controls the conversion of cropland into other usage, it is also impossible for a plant to affect land use in a large area. Hence, if a location choice regarding land ownership exists, it should be limited to areas very close to a plant (i.e., within several hundred meters). I then drop areas within 1 or 2 km of a plant, where the fire reduction may be explained by location choice, and redo the baseline analysis. The results are stable as shown in Appendix Table A8.

Selection on unobservables Finally, there may also exist other unobserved factors that explain both the arrival of a plant and the subsequent reduction in fires. I use the method proposed by Oster (2019) to assess how sensitive the results are to selection on unobservables. The method assumes that selection on unobservables is proportional to selection on observables and derives a bias-adjusted coefficient formula using the coefficient and R^2 obtained with and without observed controls:

$$\beta_{Adjusted} \approx \beta_{Control} - \delta \left[\beta_{NoControl} - \beta_{Control}\right] \frac{R^2_{Max} - R^2_{Control}}{R^2_{NoControl} - R^2_{Control}} \tag{A4}$$

where δ measures the degree of selection on unobservables relative to observables and R^2_{Max} denotes the hypothetical R^2 with all observables and unobservables included. As suggested by Oster (2019), I set $R^2_{Max} = 1.3R^2_{Control}$.[13] I treat the specification (1) with only grid cell and plant $\times$ year fixed effects as with no controls, and the most demanding specification (2) as with controls. I focus on negative values of δ as only in this case my results would be biased towards zero.[14] As shown in Appendix Figure A14, the coefficients only change slightly towards zero, if selection on unobservables is equally important as selection on observables (i.e., $|\delta|$=1). To explain away my findings, $|\delta|$ needs to be larger than 3 (actually it needs to be larger than 5 for 3 out of the 4 coefficients). Given that the rural areas are relatively homogeneous and that I have included very granular geographic factors affecting agricultural production, it is unlikely

[11] They typically use the same technology from Denmark and have roughly the same size between 20-30 MW.

[12] The plant covers an area of about 0.04 km^2, including the warehouse used to store crop residues.

[13] Oster (2019) shows that it is unrealistic to set $R^2_{Max} = 1$ and derives the 1.3 multiplication factor from RCT papers such that 90% of the results would survive.

[14] This means that selection on unobservables is negatively related to selection on observables.

that selection on unobservables would be several times larger than selection on observables.

E Additional Data Issues

E.1 Data Quality

There is a concern about the representativeness of the data collected from third-party sources. To check this, I collect an official report published by the National Energy Administration of China.[15] According to the report, 254 biomass power plants with agricultural waste as inputs had been built by the end of 2016. In my dataset, I have 253 such plants by the end of 2016 (the final sample for regression includes 190 plants as I only keep those opening between 2002 and 2015 and also drop about 24% plants with overlapping 15 km buffers). As stated in the main text, these 253 plants were active plants by the end of 2016. As the official number of all plants is 254, this means that there may be 1 plant that opened and closed before 2016. Such small number of closed plants is reasonable since these plants belong to the strategic industries supported by the government. Compared to the 190 plants in my final sample, omitting one closed plant is unlikely to cause any substantial impact on my estimates. To see this, note that a closed plant was likely to perform poorly than the average plant, implying that it was likely to fall behind the average plant in collecting crop residues. Hence, the treatment effect for such a plant should be weaker. Assume zero treatment effect for conservation. Then the average treatment effect including such plant should be roughly equal to $(1 \times 0 + 190 \times \hat{\beta})/(1 + 190) = 0.99\hat{\beta}$. Hence, the bias associated with not including this closed plant is obviously negligible.

The remaining data quality concerns may include measurement error in the physical address and year of opening of a plant. Measurement error in the physical address of a plant can bias the translated longitude and latitude in any direction and by any amount. Then such error can be viewed as random and bias my estimates towards zero; this would imply that the plants are even more effective in reducing burning than I show. Furthermore, as I am estimating the average treatment effects across the entire 15 km buffer, a potential non-random measurement error in the longitude and latitude of a plant location with reasonable size—for example, about 1 or 2

[15] See http://www.jlnyxx.com/shengwu/5473.html.

km—would only change the treatment status of a small proportion of grid cells. As treatment effects are unlikely to differ dramatically across spatially adjacent grid cells, correcting for the treatment status of such affected grid cells would only change the results by a small amount. Measurement error in the year of opening is also possible. However, as my data period covers 16 years, a reasonable measurement error, say 1 or 2 years, is unlikely to change my results substantially. As discussed in Appendix E.4, the results are robust to excluding the year of opening or a three-year window around the year of opening.

E.2 Map Censoring

When translating Chinese physical addresses into geographic coordinates using the *geocoding* APIs provided by map providers in China (e.g., Baidu, AutoNavi, and Google),[16] the output coordinates use the GCJ02 coordinate system, which was developed by the Chinese State Bureau of Surveying and Mapping in 2002.[17] This coordinate system is typically called the censored coordinate system as it adds some offsets, up to a few hundred meters, to the widely used WGS84 coordinates (Xu, Lewis and Guan, 2019).[18] To match such coordinates to other data that use the WGS84 coordinates, I use a simple iteration procedure, which only relies on a few tools provided by Baidu Maps, to transform the GCJ02 coordinates to WGS84 coordinates.[19] Before outlining the procedure, I will first introduce some technical details regarding map censoring in China, which forms the basics of the procedure.

The key feature of the censoring in the GCJ02 coordinate system is that the offsets it adds to WGS84 coordinates are *not* random: to preserve the relative positions of different locations at the very local level, the offsets (both the directions and the lengths) should be virtually identical across different locations within a small area (e.g., within 1 km). Actually, this is the key

[16]For Google Maps, it had a customized site for China: ditu.google.cn, which was closed in 2019. Currently one can only use maps.google.com. Both sites use the GCJ02 coordinate system for China, as Google's street map for China came from a Chinese map provider.

[17]Note that the outputs from Baidu Maps' *geocoding* API use two types of coordinate systems: BD09 (default output) and GCJ02. The BD09 is created by Baidu by adding additional offsets to the GCJ02 coordinates and is mainly used for display purpose. To avoid such trouble, one needs to modify the parameters to get the GCJ02 coordinates.

[18]WGS84 is the reference coordinate system used by the Global Positioning System and is also widely used by almost all geographic data providers in the world.

[19]See https://lbsyun.baidu.com/index.php?title=webapi/guide/webservice-geocoding; and https://lbsyun.baidu.com/index.php?title=webapi/guide/changeposition.

property of any map censoring algorithm; otherwise, at the very local level, the relative positions of locations in the censored map would differ significantly from those in the real world, making the censored map useless for any practical purpose such as navigation. In contrast, across locations that are distant from each other, such as two cities, the offsets could vary significantly: the directions may be totally different and the lengths may differ by several hundred meters. This is reasonable as such difference in the offsets is several orders of magnitudes smaller than the distance between two cities, creating little impact on between-city navigation or other uses that require coarser resolutions. More formally, denote the WGS84 coordinates of two nearby locations (e.g., within 1 km) by $Z_1 = (x_1, y_1)$ and $Z_2 = (x_2, y_2)$,[20] and their offsets by $O_1 = (u_1, v_1)$ and $O_2 = (u_2, v_2)$. Then the following two properties hold: (1) $O_1 \approx O_2$. Specifically, the difference is about a few meters;[21] (2) as the difference between Z_1 and Z_2 decreases, the difference between O_1 and O_2 also roughly decreases.[22]

To better illustrate the offsets, I create Appendix Figure A11 by overlaying the street map (shown in translucent color) onto the satellite map in Google Maps.[23] Google Maps creates the satellite map of China using the WGS84 coordinate system, but it uses the street map with the GCJ02 coordinate system. These two types of maps work well individually. However, a discrepancy emerges when simultaneously loading the two maps. Such discrepancy provides the chance to visually gauge the offset by comparing the locations of some unique buildings in the two maps. Panel A shows a place at the center of Beijing and Panel B shows a place 8 km east of the place in Panel A. In each place, there are two red arrows with the starts denoting the locations of two nearby buildings (within 1 km) in the WGS84 satellite map while the ends denoting their mappings into the GCJ02 street map. The lengths of the arrows measured in meters are also labeled in red on the arrows. Two findings are noteworthy: (1) at two nearby locations within the same place (i.e., within Panel A or Panel B), the offsets are virtually identical: the directions are parallel and lengths differ by less than 1 meter; (2) across two places that are

[20] Note that here the coordinates can also be denoted by projected coordinates (in meters), as at the very local level there is an affine transformation from geographic coordinates (in degrees) to projected coordinates (in meters).

[21] One can gauge this visually using Google Maps as I will describe below. The exact amount of difference does not affect the iteration procedure I use.

[22] The "roughly" is used here because the iteration procedure does not require a strict decrease. A rough decrease may only increase the number of iterations.

[23] See https://www.google.com/maps/place/Tiananmen+Square/@39.9053084,116.381517,14z.

several km away from each other (i.e., Panel A versus Panel B), the difference in the offsets becomes larger by about 10 meters (the directions are still quite parallel). These findings are consistent with the two properties described above.[24]

Additionally, the Chinese State Bureau of Surveying and Mapping provides an official tool that can be used to transform WGS84 coordinates into GCJ02 coordinates (but not reverse). This tool is typically called *coordination conversion* and one can access it through major map providers such as Baidu Maps.[25] Although the tool is a complete black box and one can only know the outputs (GCJ02 coordinates) from inputs (WGS84 coordinates), it helps to gauge the degree of censoring. With this tool and the two properties described above, one can use the following procedure to iteratively gauge the censoring and approximate the true WGS84 coordinates:

1. Specify a tolerance level $\ell = 1 \times 10^{-6}$ and input a GCJ02 coordinate $G_0 = (x_0, y_0)$, which is measured in degrees.

2. Set the k-th guess of G_0's WGS84 counterpart $W_k = (x_k, y_k)$. Set the initial guess $W_1 = G_0 = (x_0, y_0)$.

3. Use the official tool through Baidu Maps to get W_k's GCJ02 counterpart $G_k = (\tilde{x}_k, \tilde{y}_k)$ and calculate the offset $O_k = G_k - W_k$.

4. If $|\tilde{x}_k - x_0| < \ell$ and $|\tilde{y}_k - y_0| < \ell$, stop the iteration and return $W_{stop} = W_k$. Else, go back to step 2 and update the guess to $W_{k+1} = G_0 - O_k$

In this procedure, one inputs a GCJ02 coordinate G_0 and starts from an initial guess of its WGS84 counterpart W_1. To ensure that the guess is sufficiently close, the initial guess W_1 is set to be numerically equal to G_0; thus, W_1 points to a location a few hundred meters away from the location that G_0 points to.[26] One can then use the official tool to get W_1's GCJ02 counterpart G_1 and the offset O_1. O_1 would only differ from the offset of the location that G_0 points to by a few meters due to first property described above.[27] By subtracting O_1 from G_0 one gets

[24] One can also check for other places in other cities. The results are similar.

[25] See https://lbsyun.baidu.com/index.php?title=webapi/guide/changeposition.

[26] Remember that the degree of censoring is about a few hundred meters at anywhere in China.

[27] The exact amount of difference does not matter (e.g., it could also be dozens of meters), as such difference would be reduced quickly through iterations.

a better new guess of G_0's WGS84 counterpart W_2, which also differs from G_0's true WGS84 counterpart by a few meters. Then one proceeds to the next iteration and gets an offset O_2 that differs much less from the offset of the location that G_0 points to (due to the second property described above). This new offset can be further used to get a much better new guess of G_0's WGS84 counterpart W_3. The stopping rule is that one reaches a WGS84 coordinate W_{stop} whose GCJ02 counterpart is highly close to G_0 (differs only by ℓ).[28] Because the censoring needs to preserve the relative positions of different locations at the very local level, then W_{stop} can be used as a highly accurate estimate of G_0's true WGS84 counterpart, with an error at the same order of ℓ. To have a sense of the accuracy, note that 1 degree of longitude or latitude is at most 111 km at anywhere on the earth. This means that the error is at an order of 0.1 meter (i.e., $1 \times 10^{-6} \times 111 \times 1000$). Given that the unit of my analysis is 1 km by 1 km, such error is obviously negligible.

E.3 Sample Selection

The baseline specification excludes about 24% plants that have overlapping 15 km buffers in order to clearly define the treatment and control areas. To assess the potential bias created by such sample selection process, I include these plants in two ways. First, following Currie et al. (2015), I include all plants and grid cells in their 15 km buffers. This means that a grid cell may be linked to more than one plant in a certain year: it can be treated by a plant but can also serve as control for another plant, if these two plants have overlapping 15 km buffers. Second, following Muehlenbachs, Spiller and Timmins (2015), I include all plants and non-overlapping grid cells in their 15 km buffers. This means that, compared to the first approach, here I exclude grid cells that are within 15 km of more than one plant. Appendix Table A9 shows the results for both exercises. Both the economic magnitude and statistical significance are similar to those in the baseline specification, suggesting that sample selection is not a threat to my findings.

[28] In practice, this procedure converges quickly and one can reach the tolerance level ℓ in only several iterations.

E.4 Treatment Timing

As I do not observe when during the year a plant opens, the indicator $Post_{jt}$ in the baseline specification takes the value of 1 for years after opening and 0 otherwise. However, one may be concerned about possible misspecification about the treatment value in the year of opening: a plant opening in January may enjoy a treatment value of 1 while a plant opening in December may enjoy a treatment value of 0, since crop harvest and burning typically occur in the summer and autumn. Intuitively, this would only be a threat if the treatment effects are concentrated around the year of opening. The event study graphs, however, show that the treatment effects are relatively small in the year of opening and are larger and quite persistent after the year of opening. Hence, it is unlikely that the DID estimates would be affected by such issues.

To formally assess how sensitive the results are to treatment timing, I conduct two tests following Currie et al. (2015). First, I exclude the year of opening, or a three-year window around the year of opening (i.e., the year before opening, the year of opening, and the year after opening). The remaining years then can be unambiguously defined as treated or untreated. Second, I assign an "expected" treatment value to the year of opening, which takes the following form:

$$E(T_j) = \sum_{m=1}^{12} \left(p_j^m \times \sum_{n=m}^{12} ResidueShare^n \right) \tag{A5}$$

where p_j^m denotes the probability that plant j opens in month m and $ResidueShare^n$ denotes the share of crop residues generated in month n. The intuition is that if more crop residues are left after month m, then the plant would generate a larger impact and enjoy a larger treatment value.[29] By construction, $E(T_j)$ takes the value between 0 and 1. To calculate $E(T_j)$, I first assume that plant openings are evenly distributed by month throughout a year. That is, p_j^m =1/12. This is a plausible assumption given that opening a large plant in China, especially one in strategic industries such as utilities, requires a lengthy bureaucratic procedure. Hence, the month of opening can be viewed as random. I then assume that farmers burn a fixed percentage of crop residues so that I can infer the distribution of crop residues across months from the entire agricultural fire data from 2001 to 2016. I estimate such distribution separately for each province

[29] As crop residues would affect subsequent cultivation, crop residues generated in month m are assumed to disappear in the next month.

as crop types and harvest calendar may differ across provinces.[30] As shown in Appendix Figure A13, the expected treatment value ranges from 0.36 to 0.8 and varies significantly across provinces. I then reestimate the baseline equation (2) by replacing the value of $Post_{jt}$ in the year of opening with $E(T_j)$. As shown in Appendix Table A10, the results from such exercises are similar to the baseline results.

F Additional Mechanism Tests

F.1 Qualitative Evidence on Economic Incentives

To better understand how the plants collect crop residues and how this can motivate farmers to reduce burning, I collect several cases from technical reports and newspapers and summarize the relevant information below.

The collection of crop residues is generally done by the plants in a decentralized way.[31] Specifically, the plants first set up several groups of brokers formed by workers and some farmers in their primes. Then each group of brokers would be equipped with trucks provided by the plants and based in a few villages or towns to collect crop residues. As young people migrate to work in urban areas, most people engaged in agricultural production are women and the elderly, who are generally unable to transport the cumbersome crop residues. This is the main reason why these brokers exist. The local farmers then only need to collect and package the crop residues from the cropland and wait for the brokers to come, which still requires some efforts but is relatively doable for these farmers. Of course, they can also directly send the crop residues to where the brokers are based or where the plants are located, although this is less common. But no matter who performs the transport, farmers' revenue from the crop residues would generally decrease with the distance between the cropland and the plant. At areas sufficiently far away from the plant, the revenue would become low enough so that farmers may not even want to collect and package the crop residues.

The price of crop residues typically ranges from 200 to 400 *yuan* (about 31-62 USD) per

[30]Results are similar if I allow the distribution to vary by $2° \times 2°$ grid cells.

[31]See http://www.nantong.gov.cn/jcms/jcms_files/jcms1/web14/site/attach/0/160429135950571.pdf for how a typical plant works. Also see https://openknowledge.worldbank.org/handle/10986/2897 for a general report on these plants.

ton, depending on the type of crop residues and whether the residues are sold directly to the plants or through the brokers (see the following quotes for the prices).[32] To benchmark such price, consider the price of 250 *yuan* per ton for maize recorded in a report in 2009 (paid to farmers through the brokers).[33] One *mu* (about 0.067 hectare) of maize cropland can produce 0.4 ton of residues, amounting to 100 *yuan*. According to a survey conducted by the National Development and Reform Commission of China in 2009, one *mu* of cropland can produce about 0.575 ton of maize with a cost of 300 *yuan*.[34] Combined with the price of maize of 1,670 *yuan* per ton at that time,[35] the revenue from maize is $1,670 \times 0.575 - 300 = 660$ *yuan* per *mu*. Then the revenue from selling crop residues accounts for 100/660 = 15% of the revenue from crop production. Alternatively, as migrant workers earned a daily wage of 47 *yuan* (1,417/30) in 2009,[36] then the revenue from selling crop residues amounts to about two days' wages. In sum, the revenue from selling crop residues is certainly nontrivial to farmers. Furthermore, as in this case the crop residues are transported by the brokers, farmers only need to collect and package them from the cropland. This is a relatively easy task and will generally cost no longer than two days. Hence, selling crop residues should be attractive to at least some farmers.

To provide further evidence regarding the price of crop residues and farmers' willingness to sell crop residues, I present some quotes below extracted from various national and local newspapers in China:

> *On June 12, farmers in Nigou Town, Taierzhuang District, Zaozhuang City, Shandong Province lined up to sell wheat straw, which would be used by the power plant for power generation. Selling straw can increase farmers' income by more than 160 yuan per ton, and at the same time avoid the pollution caused by straw burning. (Xinhua News Agency, 2008)*[37]

[32] For the price, see http://www.moa.gov.cn/ztzl/jsshzyxnc/zjnnsr/200911/t20091118_1385166.htm; http://fgw.sh.gov.cn/resource/19/1933ede6596e407a81d515bfcf84384b/5b04d6aadde5605ef1deb160d908062c.doc; and the following quotes from newspapers.

[33] See http://www.moa.gov.cn/ztzl/jsshzyxnc/zjnnsr/200911/t20091118_1385166.htm.

[34] See https://www.ndrc.gov.cn/fzggw/jgsj/njs/sjdt/200910/t20091012_1194509.html?code=&state=123. The cost is for raw materials and does not include labor cost. The survey gives a yield range of 0.55-0.6 ton per *mu*. I take the average of 0.575 ton per *mu*.

[35] See http://lswz.hebei.gov.cn/lysc/hqsp/200911/t20091110_26611.html.

[36] See http://www.stats.gov.cn/ztjc/ztfx/fxbg/201003/t20100319_16135.html.

[37] See http://www.gov.cn/jrzg/2007-06/13/content_646599.htm.

In Gaotang County, Wang Shouhua, a farmer from Liusi Town, who was waiting in line for selling the crop residues, told the reporter happily: "This year I had a good harvest of cotton. Due to frequent rains and cotton's high straw-to-grain ratio, I planted 10 mu more cotton this year. As a result, my revenue would exceed last year by about 1,200 yuan." (Farmers' Daily, 2009)[38]

"Wheat straw can be sold at 240 yuan per ton, which not only helps us to dispose of the cumbersome straw, but can also increase our income." Li Xingang, a villager in Linquanzhuang Village, Wangcheng Street, told the reporter. (Economic Daily, 2018) [39]

On June 10, in Beishi Village, Dong'e Town, Pingyin County, 74-year-old Mrs. Liu went out early in the morning with a tricycle full of straw. "This 1 ton of straw might sell for 300 yuan." Said Mrs. Liu. (Jinan Times, 2016) [40]

Zheng Shengguo, a farmer in Beixiangyang Village, Tangyuan Town, Tangyuan County, planted 3 hectares of maize this year. A large amount of straw was left behind after harvest. While he was wondering how to dispose of the straw, he heard that the newly-built biomass power plant in the county bought straw with relatively high price, so he found a local broker and delivered all the straw to the plant. Zheng Shengguo earned more than 2,000 yuan from a total of more than 20 tons of straw. (Ministry of Agriculture and Rural Affairs, 2008)[41]

"In the past, after the wheat was harvested, the straw was either burned or wasted. Nowadays, a biomass power plant has been built in my county. Selling the straw to the brokers is not only profitable but also beneficial to the environment. It is one stone for two birds!" Said happily Wang Ronghua, a farmer in Wulong Town, Shangcai County, who took over 300 yuan at his just harvested cropland from the straw broker not long ago. (Zhumadian Daily, 2013) [42]

[38]See http://www.moa.gov.cn/ztzl/jsshzyxnc/zjnnsr/200911/t20091118_1385166.htm.
[39]See http://www.xinhuanet.com/energy/2018-01/22/c_1122292671.htm.
[40]See https://news.e23.cn/jnnews/2016-06-16/2016061600605.html.
[41]See http://www.moa.gov.cn/ztzl/jsshzyxnc/zjnnsr/200811/t20081119_1175488.htm.
[42]See http://www.nzdb.com.cn/hy/50026.jhtml.

"Sir, have you sold your maize straw? We will purchase at a high price." "You are late as I have already sold to the biomass power plant in the county." On October 14, the straw broker Wang Haiming in Guantao County was busy purchasing straw, which was abandoned by the farmers in the past but is now in great demand. Since the opening of the biomass power plant this year, many straw brokers have noticed the high demand for straw by the plant and gone to the villages to collect straw, which can be sold at 180 yuan per ton. (Hebei Daily, 2010)[43]

"During the slack season, I will collect some straw, peanut shells, and bark to sell here to earn some money." Wang Yuqian, a farmer from Ninglaozhuang Town, Yingquan District, said while preparing for weighing, "I can handle 1.5 tons per time and earn more than 300 yuan minus expenses such as fuel costs." (Fuyang Daily, 2014)[44]

Shen Ruihua, deputy general manager of the biomass power plant, told reporters that in recent years, with the improvement of people's environmental and economic awareness, more and more farmers had begun to send straw and other wastes generated in agricultural production to the plant. "We price the straw by quality. Straw sent by farmers with moisture within 20% and gray scale within 8% is 320 yuan per ton, rice husk and peanut husk are 400 yuan per ton, and the bark is 310 yuan per ton. " Said Shen. (Fuyang Daily, 2014)[45]

In sum, these quotes show that farmers are well aware of the economic and environmental benefits of selling crop residues. They also view the price paid by the plants as attractive.

F.2 Quantitative Evidence on Economic Incentives

To further corroborate the economic incentives mechanism, in this section I directly test the effect of plant openings on farmers' income using household income data from the China Family Panel Studies (CFPS). Because survey data like the CFPS typically does not report detailed

[43] See http://hebei.hebnews.cn/2010-10/15/content_1102948.htm.
[44] See https://news.ifeng.com/a/20140515/40310152_0.shtml.
[45] See https://news.ifeng.com/a/20140515/40310152_0.shtml.

household addresses, I am not able to exploit hyperlocal variation around the same plant as in the baseline specification. Instead, I focus on plant openings at the city level. As it may not be realistic to assume that cities with and without plant openings are ex ante comparable, I estimate a triple difference-in-differences specification, which further exploits within-city variation:

$$Y_{hcw} = \lambda_c + \mu_w + \beta_1 Opening_{cw} + \beta_2 Opening_{cw} \times Farmer_h + \beta_3 Farmer_h + \varepsilon_{hcw} \tag{A6}$$

where Y_{hcw} denotes the constant price income of household h in city c surveyed in wave w. I use four waves of the survey data: 2010, 2012, 2014, and 2016. λ_c and μ_w denote city fixed effects and survey wave fixed effects, respectively. $Opening_{cw}$ is an indicator equal to 1 if there are any plant openings in city c by wave w. $Farmer_h$ is an indicator equal to 1 if household h lives in the rural areas.[46] The coefficient of interest is β_2, which measures the differential effect of plant openings on farmers' income relative to non-farmers' income. The underlying identification assumption is that farmers' income would evolve the same way as non-farmers' income in the absence of plant openings, which will be supported by an event study. I will also partially relax such assumption by allowing for farmer-specific survey wave trends. [47] In contrast, β_1 measures the effect of plant openings on non-farmers' income, which may simply reflect diverging trends between cities with and without plant openings. In some specifications, I will add city by survey wave fixed effects to fully control for the selective entry of plants into cities. I cluster standard errors at the city level for all specifications.

Appendix Table A11 presents the results. Column (1) shows that relative to non-farmers, plant openings significantly increase farmers' income by 3,693 *yuan*, which amounts to 12% of farmers' mean income. This is comparable to the 15% increase in income from selling crop residues as shown in the back-of-the-envelope calculation in Section F.1. In columns (2)-(3), I provide an intuitive solution to the recent econometric concerns about staggered DID designs (de Chaisemartin and d'Haultfoeuille, 2020).[48] Specifically, column (2) compares cities with

[46]Results are similar if I define farmers as those who were engaged in agricultural production in a survey year, although such measure may be endogenous to plant openings.

[47]An additional concern is that the plants may differentially affect non-farmers' income through labor market effects. I view this as a minor concern due to two reasons: first, an average plant typically hires around one hundred employees, which are minimal relative to the entire population in a city; second, even if the labor market effects are substantial, the estimates from equation (A6) can be viewed as a lower bound of the effect on farmers' income.

[48]Also see Callaway and Sant'Anna (2020) and Goodman-Bacon (2021).

plant openings in wave 2 to other cities that never have any plant openings; column (3) compares cities with plant openings in wave 3 to other cities that never have any plant openings. The results are slightly larger than that in column (1). In column (4), I add city by survey wave fixed effects to account for selective entry of plants into cities. The results also become slightly larger. In column (5), I further add farmer-specific survey wave trends to relax the assumption that farmers and non-farmers would be on the same trajectory in the absence of plant openings. This decreases the effect by half, but the magnitude is still economically meaningful (about 6% of farmers' mean income). Finally, column (6) estimates an event study and shows that the leads are small in magnitude and statistically insignificant; in contrast, the lags are relatively larger and some are statistically significant. This bolsters the identification assumption of no pre-trends between farmers and non-farmers. In sum, these results suggest that plant openings have a differential impact on farmers' income with a magnitude of at least 6%, which is a non-trivial fraction and lends additional support to the economic incentives mechanism.

F.3 Changes in Small Firms

To address the concern that farmers may work in very small firms that cannot be captured by the ASIF data, I collect additional data from the 2004 and 2008 economic censuses conducted by the National Bureau of Statistics of China. Such data cover the entire population of industrial firms. Because I only have two periods of data, I focus on plants that opened between 2004 and 2008 and estimate a long-difference model similar to Faber (2014):[49]

$$Y_{ij}^{2008} - Y_{ij}^{2004} = \lambda_j + \beta_1 Distance_{ij}^{<5km} + \beta_2 Distance_{ij}^{5-10km} + \mathbf{X}_i + \varepsilon_{ij} \tag{A7}$$

where Y_{ij} denotes the number of industrial firms or employees in grid cell i around plant j,[50] λ_j denotes plant fixed effects, $\mathbf{X}_i$ denotes a vector of geographic controls identical to those used in the baseline regression, and $Distance_{ij}^{<5km}$ and $Distance_{ij}^{5-10km}$ are indicators for the most treated areas and less treated areas, respectively. Standard errors are clustered at the plant

[49] This is equivalent to the difference-in-differences specification in equation (2) , except that here it is impossible to estimate grid cell-specific trends with 2 periods of data.

[50] The economic census in 2008 does not contain relevant variables for calculating the output and capital as reported in Table 5.

level. The coefficients β_1 and β_2 are essentially difference-in-differences estimators capturing the changes over the two periods in the treated areas relative to the control areas. As shown in Appendix Table A12, there is little evidence that structural transformation occurred, as the coefficients are neither economically nor statistically significant.

F.4 Cropland Difference

There is a concern that the drop in fires may be driven by cropland and agricultural production near the plants changing differentially relative to those elsewhere. The most plausible reason for such change is that the plants may impose negative externalities on agricultural production through various types of pollution. To assess such impact, I collect a technical report on the environmental impacts of a typical biomass power plant, written by a professional environmental assessment company and reviewed by the provincial government.[51] The report should be representative as such plants typically use the same technology from Denmark and have roughly the same size. According to the report, the plant mainly generates air pollution and noise.

For air pollution, the report measures the air pollutants generated in an hour using monitors and then uses an atmospheric diffusion model to derive the pollution concentration at various distances downwind of the plant.[52] The results are visualized in Appendix Figure A15 Panel A. The pollutant concentrations peak at areas 900 meters away from the plant and then decrease quite quickly.[53] To benchmark the magnitudes, I compare the maximum concentrations to the ambient air quality standards recommended by the Ministry of Environmental Protection.[54] As shown in the figure, the maximum concentrations account for less than 1% of the national standards, except for NO_x, which is about 2.6%. Such minimal magnitudes are consistent with the fact the plant can reduce pollutants released from burning using dedusting and desulfurization equipment.

[51] See http://www.nantong.gov.cn/jcms/jcms_files/jcms1/web14/site/attach/0/160429135950571.pdf. The plant is located in East China, one of China's major grain-producing areas, and has a typical size of 30 MW.

[52] The report does not directly report $PM_{2.5}$ and PM_{10}. I assume that their proportions in PM are 10% and 40%, respectively. These numbers are taken from the maximum proportions for different power plants. See https://www.mee.gov.cn/gkml/hbb/bgth/201401/W020140124409250921321.pdf.

[53] The hump shape is mainly driven by the plant using a tall chimney to emit pollutants.

[54] See https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/dqhjbh/dqhjzlbz/201203/t20120302_224165.shtml.

To translate such an increase in air pollution—although negligible—into change in fires, I use estimates from the literature on the effect of air pollution on crop yields. Lobell and Burney (2021) show that one standard deviation increase in key air pollutants (ozone, particulate matter, sulfur dioxide, and nitrogen dioxide) can reduce crop yields (maize and soybean) by about 2%. Assume that the number of fires is proportional to the amount of crop yields. As in their data one standard deviation for various pollutants is generally much larger than the maximum concentrations for all pollutants shown in Panel A, then such an increase in air pollution around the plant should at most reduce fires by 2%. As I find a much larger drop in fires (more than 30%), then my findings cannot be explained by air pollution. Furthermore, as the pollution is mainly concentrated in areas within 1 km of the plant, in Appendix Table A8 I show that the results remain virtually unaffected when I drop such areas.

The noise generated by the machines in the plant generally ranges from 80 to 90 dB. Occasionally, the noise generated by the boiler can reach 130 dB, which lasts for less than one minute and can be reduced to below 90 dB by the deadener installed in the plant. Hence, the maximum noise level can be viewed as about 90 dB. The report uses monitors to measure noise level near the plant and then uses a noise decay function to create predicted values at farther distance. As shown in Appendix Figure A15 Panel B, the noise level decays quickly with distance. At areas about 400 meters away from the plant, the noise level is 38 dB, which is smaller than the 45 dB recommended by the WHO for getting good sleep in dwellings.[55] Hence, noise is also unlikely to affect agricultural production and fires.

One may also be concerned about water pollution. However, it is unlikely to drive my results due to several reasons. First, water is mainly used in the cooling system and also recycled in the plant. Hence, waste water mainly comes from human activities in the plant and can be viewed as minimal as such a plant only has about one hundred employees. Second, the impact of water pollution should be restricted to areas close to rivers, which can be well controlled by the interaction between proximity to rivers and year fixed effects. Finally, I do not find any impact of plant openings on water pollution as shown in Appendix H.2.

There may also exist other channels through which plant openings affect cropland and agri-

[55] See http://whqlibdoc.who.int/hq/1999/a68672.pdf.

cultural production. Such impact should be reflected in changes in crop share and the value of crop production. As shown in Appendix F.5, there is little change in the share of major crops and the value of crop production after a plant opens. Hence, these results combined suggest that my findings are unlikely to be driven by differential changes in cropland near the plants.

F.5 Farmer Adaption to Plants

Farmers may adapt to the plants on many dimensions of agricultural production, such as crop choice, harvest techniques, labor usage, timing of harvest/planting or even productivity; this may not only explain the reduction in fires but could also be interesting in itself. Although rigorously testing this requires highly disaggregated data—which is nearly impossible—on agricultural production and is also a bit out of the scope of this paper, I can still shed some light using the 2005 and 2010 global gridded dataset on agricultural production described in Appendix D. I keep grid cells whose centroids are within 15 km of a plant. To check crop choice, I focus on the share of the top five major crops harvested, as described in the previous section. For harvest techniques, labor usage, and timing of harvest/planting, I cannot test them directly due to lack of data. I instead focus on the total value of production, which is reported in the dataset. The total value of production can be viewed as a function of different inputs described above; if there is any change in these inputs, then it should be largely captured by the change in the total value of production. Finally, I use the total value of production per hectare harvested to measure productivity.

I focus on plants that opened between 2005 and 2010 and estimate a standard long-difference model similar to equation (A7):

$$Y_{ij}^{2010} - Y_{ij}^{2005} = \lambda_j + \beta_1 Distance_{ij}^{<5km} + \beta_2 Distance_{ij}^{5-10km} + \mathbf{X}_i + \varepsilon_{ij} \tag{A8}$$

where Y_{ij} denotes crop share for the top five crops or total value of production (per hectare) in grid cell i around plant j, λ_j denotes plant fixed effects, and $Distance_{ij}^{<5km}$ and $Distance_{ij}^{5-10km}$ are indicators for the most treated areas and less treated areas, respectively. Similar to those used in the baseline regression, $\mathbf{X}_i$ denotes a vector of geographic controls, although they are measured at a larger geographic unit due to the coarse resolution of the crop data. Standard errors are clustered at the plant level. The coefficients β_1 and β_2 are essentially difference-in-differences estimators capturing the changes over the two periods in the treated areas relative to the control areas. Columns (1)-(5) of Appendix Table A13 shows little change in major crop shares. If anything, there is a small and marginally significant increase in rice share in the less treated areas; as rice is one of the major sources of crop residues, this goes against my findings that there is a reduction in agricultural fires. Column (6) shows little change in the total value of production. Column (7) shows similar results for productivity as measured by the total value of production per hectare harvested. These results combined suggest that it is unlikely that farmers adapt to the plants by changing agricultural production along various dimensions.

F.6 Potential Fire Bans under Transmission Towers

There is a concern that the baseline results may be driven by potential fire bans under transmission towers. Conceptually, this is unlikely to explain my results for the following reasons. First, given that the electrification rate in China is close to 100% during the sample period,[56] it is plausible that the plants would connect to the distribution grid of electricity through existing transmission towers. Hence, even if fires are banned under transmission towers, this will not confound my results as long as these transmission towers are not time-varying. Second, even if there are newly-built transmission towers, the areas under transmission towers would be smaller than the treatment areas by several orders of magnitude. Hence, it is unlikely that the sizable reduction in fires would be driven by a few transmission towers.

To empirically rule out such possibility, ideally, one would like to have the exact locations of the transmission towers, but unfortunately this is unavailable. I therefore test this indirectly. The transmission of the power and the newly built transmission towers, if any, are likely located in a ray whose initial point is at the plant. Given that this ray would lie in either half of the catchment area, then if the results are driven by the fact that fires are banned under the newly built transmission towers, one could split the catchment area in half and expect to see differential reduction in fires for each half of the catchment area. I repeat this process for each plant and

[56]See https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=CN.

find differential results for only 12 (about 6.3%) plants, and my baseline results barely change after excluding these 12 plants (Appendix Table A14, columns 1-2).

Alternatively, to check if this explanation holds, one can also test if the reduction in fires is smaller in areas with higher frequency of lightning strike, as transmission towers are less likely to be placed in areas susceptible to lightning strike due to the damage of lightning on electrical systems and power transmission lines (Said, Cohen and Inan, 2013).[57] This is in the spirit of Manacorda and Tesei (2020) who use lightning strike as instrument for the placement of communication towers. To this end, I use high resolution lightning data from Albrecht et al. (2016). As seen from columns (3)-(4) of Appendix Table A14, the reduction in fires is not smaller in areas with higher frequency of lightning strike. Taken together, it is unlikely that the baseline results are driven by fires being banned under transmission towers.

G Other City-level Regulations

There is a concern that the results on the plants could be confounded by other environmental regulations at the city level.[58] Conceptually, as my DID design utilizes spatial variations around the same plant in the same year and is also robust to including city $\times$ year fixed effects, then the results are unlikely to be confounded by city-level regulations. To further show the robustness, I examine a number of major regulations during my sample period in this section, in addition to the 2013 air quality monitoring campaign examined in Table 2. I then model the impacts of these regulations to see whether the results are stable in Appendix Table A15.

Two Control Zones The Two Control Zones (TCZ) policy was initiated in 1998 to control for acid rain and SO_2 pollution.[59] A wide literature has tested its environmental and economic impacts (Tanaka, 2015; Cai et al., 2016; Chen, Li and Lu, 2018). The policy designated 175 cities as TCZ cities (see Appendix Figure A16 Panel A), which were required to take various actions, such as closing coal mines with high sulfur, to achieve the national standards on acid

[57] Also see https://www.inmr.com/hazards-lightning-transmission-lines/.

[58] As other city-level regulations are unlikely to vary discontinuously at the no burning zone boundaries, I will only consider the threat of these regulations to the results on the plants.

[59] See http://www.gov.cn/zhengce/content/2010-11/22/content_5181.htm.

rain and SO_2 by 2010. The policy may also affect agricultural fires as they can also release SO_2 and NO_x—though with minimal amount— that contribute to acid rain. In column (1) of Appendix Table A15, I add the interaction between the distance from a grid cell to the center of the nearest TCZ city and year fixed effects to control for this policy. The results barely change.

Key Cities for Air Pollution Control The Key Cities for Air Pollution Control (KCAPC) policy was another major policy in the early 2000s,[60] although less examined in the literature (Liu, Tan and Zhang, 2021). It was launched in two batches in 1998 and 2001, with the first batch covering 47 cities and the second batch covering 66 cities (see Appendix Figure A16 Panel B). In addition to acid rain and SO_2, the policy also targeted particulate matter. In column (2) of Appendix Table A15, I add the interaction between the distance from a grid cell to the center of the nearest KCAPC city and year fixed effects to control for this policy. The results barely change.

Two Control Zones after 2005 The effect of the TCZ policy was shown to be small and temporary. To ensure more stringent enforcement, in late 2005 the central government linked the enforcement of the TCZ policy to the performance evaluation system of local leaders,[61] which drove local leaders to trade off economic growth to comply with the policy (Chen, Li and Lu, 2018). In column (3) of Appendix Table A15, I add the interaction between the distance from a grid cell to the center of the nearest TCZ city and an indicator for years after 2005 to control for this policy.[62] The results barely change.

Regulation during the Beijing 2008 Olympic Games To prepare for the 2008 Beijing Olympic Games, China launched a radical air pollution regulation in all host cities and some neighbor cities starting from 2007 (see Appendix Figure A16 Panel C), such as closing or temporarily shutting down some power plants and chemical plants.[63] A few studies have doc-

[60] See http://fgcx.bjcourt.gov.cn:4601/law?fn=chl266s052.txt.

[61] See http://www.gov.cn/zwgk/2007-11/26/content_815498.htm.

[62] Results are similar if I interact the distance to policy city center with year fixed effects. The use of a post-2005 indicator is simply to more directly model the policy effect starting from 2005.

[63] See http://www.gov.cn/xwfb/2008-02/27/content_903668.htm.

umented the air quality improvements brought by such regulation (Chen et al., 2013; He, Fan and Zhou, 2016). To rule out its confounding effect, in column (4) of Appendix Table A15 I add the interaction between the distance from a grid cell to the center of the nearest host city (or neighbor city) and an indicator for years 2007-2008. The results barely change.

Key Regions for Air Pollution Control In 2012, China designated 117 cities as Key Regions for Air Pollution Control (KRAPC) (see Appendix Figure A16 Panel D).[64] This policy not only targeted primary pollution as specified in previous policies but also emphasized the importance of controlling secondary pollution formed by chemical or photochemical reactions. Furthermore, it explicitly put controlling $PM_{2.5}$ at the first place. To check whether this policy confounds my results, in column (5) of Appendix Table A15 I add the interaction between the distance from a grid cell to the center of the nearest KRAPC city and an indicator for years after 2012. The results barely change.

In sum, these policies have little effect on my results. This is consistent with the fact that my identification strategy has already largely taken into consideration many city-level regulations or confounding factors through plant $\times$ year fixed effects.

H Air and Water Pollution

H.1 Air Pollution

To assess the impact of plant openings on air pollution, I focus on fine particulate matter ($PM_{2.5}$), as it is both the primary pollutant released from agricultural fires and also the primary pollutant for many Chinese cities (Simões Amaral et al., 2016; Lv et al., 2017). I use a state-of-the-art global surface $PM_{2.5}$ data product from Van Donkelaar et al. (2019) and Hammer et al. (2020).[65] Such data combine aerosol optical depth (AOD) retrievals from satellites and a

[64] See http://www.gov.cn/jrzg/2012-12/05/content_2283138.htm.

[65] See Fowlie, Rubin and Walker (2019) for a discussion of such data. I do not directly use ground-based monitoring station readings for several reasons. First, before 2013, only an air pollution index (API) for some major cities were reported, which cannot differentiate air pollution within a small area; additionally, the API only partially reflected the true air quality as it did not incorporate the major pollutant $PM_{2.5}$ and was also proved to

chemical transport model, with *in situ* monitoring station readings as training data, to predict $PM_{2.5}$ concentrations at the $0.01^\circ \times 0.01^\circ$ level (about 1 km $\times$ 1 km).[66] I extract the $PM_{2.5}$ data for each grid cell from 2001 to 2016 and average them at the 10 km buffer level for a plant. In this way, I can compare $PM_{2.5}$ concentrations *across* plants utilizing their differential timing of opening. I do not estimate the baseline equation (2), which compares areas around the same plant, for two reasons. First, as $PM_{2.5}$ particles are light and can travel at a speed of 10 mile per hour (Barwick et al., 2021), they could quickly spread from a certain fire near a plant to the other areas within several km and thereby blur the difference in $PM_{2.5}$ concentrations across different areas around the same plant. Put differently, such spatial spillover of particles violates the stable unit treatment value assumption (SUTVA) and could bias the estimates towards zero even in otherwise valid quasi-experimental settings (Deschênes and Meng, 2018). In contrast, comparisons across plants could largely alleviate such spillover issue. Second, as noted by Van Donkelaar et al. (2019), the data are gridded at the $0.01^\circ \times 0.01^\circ$ level simply to better allow for aggregation; they cannot fully reflect the $PM_{2.5}$ gradient at such resolution due to influence by information sources at coarser resolution. Hence, it is better to aggregate the data at a coarser resolution. I estimate the following difference-in-differences specification:

$$Y_{it} = \delta_i + \lambda_t + \delta_i \times Year_t + \beta Post_{it} + \mathbf{X}_{it}\gamma + \varepsilon_{it} \tag{A9}$$

where Y_{it} measures the average $PM_{2.5}$ concentrations around plant i in year t and $Post_{it}$ is an indicator for years after a plant opening. The plant fixed effects δ_i control for all time-invariant confounding factors at the plant level—such as crop suitability that may affect both the arrival of a plant and burning patterns —and thereby ensure that the coefficients are identified from changes in plant openings over time. The year fixed effects λ_t control for time-varying shocks common to all plants such as national policies regarding burning. The plant-specific linear time trends $\delta_i \times Year_t$ could control for confounding factors at the very granular level (i.e., plant) that

be largely manipulated (Ghanem and Zhang, 2014). Second, after 2013, fine-scale air quality data were published based on a national monitoring system suffering little from manipulation. However, these stations are in the urban areas and therefore are also unable to capture the pollution gradient in the rural areas. Third, in terms of the timing, most cities were covered by the end of 2014. This means that I would only have one year of data before a plant opening for most cities, which is insufficient to draw any credible causal inference.

[66]The training data only include those after 2014, which are less likely to be manipulated (Greenstone et al., 2020, 2021)

involve approximately linearly. In particular, it could control for the fact that underdeveloped areas may catch up with developed areas over time in terms of economic growth and hence pollution (Brock and Taylor, 2005). $\mathbf{X}_{it}$ is a vector of weather controls include temperature, precipitation, wind direction, wind speed, and dew point temperature.[67] To allow for nonlinear effects of weather, for each plant $\times$ year observation I include the share of days falling into each of the 10 quantiles derived from the overall daily weather distribution. As the distribution of daily precipitation is highly right-skewed, I simply use its annual average. Standard errors are clustered at the plant level.

As shown by the recent econometric literature (de Chaisemartin and d'Haultfoeuille, 2020),[68] estimates from specifications with two-way fixed effects and staggered treatment timing would be difficult to interpret. Particularly, these specifications would inevitably use units treated early as controls for those treated late. In the presence of heterogeneous treatment effects over time, such comparison would bias the coefficient β and may even lead to an opposite sign. I therefore use the DID_M estimators proposed by de Chaisemartin and d'Haultfoeuille (2020), which compare units that switch treatment status from $t-1$ to t, to units that do not switch in both periods. The causal identification of β relies on the common trends assumption; that is, $PM_{2.5}$ concentrations around different plants should exhibit the same trends in the absence of plant openings. I use their DID_M^{pl} estimators to test such assumption, which compare switchers and non-switchers from the $l+1$th to the lth period before switchers' treatment status changes, in the sample of units whose treatment status does not change from the $l+1$th to the last period before the switch. The results using equation (A9) are reported in Appendix Table A16, which shows that plant openings reduce annual average $PM_{2.5}$ concentrations by about 1 $\mu g/\text{m}^3$, amounting to about 2% of the same mean. I view this as a lower bound of the true effect, as agricultural fires mainly happen in the summer and autumn whereas the $PM_{2.5}$ concentrations are averaged over the entire year. Furthermore, the sample mean could be inflated by the higher pollution concentrations in winter, possibly due to the winter heating system (Ebenstein et al., 2017;

[67] These weather variables come from the daily weather station readings from the NOAA. For wind speed and wind direction, I use the nearest station readings within 100 km of a plant. For other weather variables, I take the inverse-distance weighted average of station readings within 200 km of a plant (Deschênes and Greenstone, 2011).

[68] Also see Callaway and Sant'Anna (2020) and Goodman-Bacon (2021).

Fan, He and Zhou, 2020).[69] Appendix Figure A17 shows the result using the DID_M estimator, which is very similar to the previous result in terms of point estimate and statistical significance (coefficient=-0.99, p-value=0.087). The figure further shows no pre-trends as the estimates from the DID_M^{pl} estimators are small and statistically insignificant.

H.2 Water Pollution

In this section I check whether the plants deteriorate water quality. This is particularly relevant for the rural areas as many people there still rely on surface water (Ebenstein, 2012; He and Perloff, 2016). To this end, I collect data on surface water pollution at the monitoring station level from the China Environmental Yearbooks. The data include station readings on several individual pollutants, such as chemical oxygen demand (COD) and dissolved oxygen (DO), and also a composite water quality index based on these pollutants. The water quality index ranges from 1 to 6, with lower values denoting higher quality. I use the period from 2004 to 2010, which covers nearly 500 monitoring stations along ten major river systems.[70] Because these stations are sparsely distributed, I estimate the following difference-in-differences specification:

$$Y_{it} = \delta_i + \lambda_t + \delta_i \times Year_t + \beta Opening_{it}^{<20km} + \mathbf{X}_{it}\gamma + \varepsilon_{it} \tag{A10}$$

where Y_{it} denotes either the water quality index or various pollutant readings at station i in year t and $Opening_{it}^{<20km}$ is an indicator if there are any plant openings within 20 km of station i by year t.[71] Similar to equation (A9), I include station fixed effects δ_i and year fixed effects λ_t, which control for time-invariant confounding factors at the station level and time-varying shocks common to all stations, respectively. I also include station-specific linear time trends $\delta_i \times Year_t$ to remove trends at the station level. $\mathbf{X}_{it}$ is a vector of weather controls include temperature, precipitation, wind direction, wind speed, and dew point temperature, which are

[69] Also see a report on the temporal distribution of $PM_{2.5}$ concentrations by the National Bureau of Statistics of China at http://www.stats.gov.cn/tjzs/tjsj/tjcb/dysj/201803/t20180312_1587456.html, which shows that the $PM_{2.5}$ concentrations are much higher in winter than in summer.

[70] The data are available from 2002 to 2010. After 2010, the yearbooks stopped publishing such station-level water pollution data. The data before 2004 are not used due to lack of readings on different pollutants and a much smaller number of stations.

[71] In my sample, there are 26 plant openings within 10 km of a station and 74 plant openings within 20 km of a station. The results are not sensitive to these distance choices.

constructed in a similar way to those in equation (A9). Standard errors are clustered at the station level. As discussed in the previous section, estimates from specifications with two-way fixed effects and staggered treatment timing would be biased in the presence of heterogeneous treatment effects. I therefore use the DID_M and DID_M^{pl} estimators from de Chaisemartin and d'Haultfoeuille (2020). Appendix Table A17 shows that water quality does not deteriorate in response to nearby plant openings, as the estimates are small and statistically insignificant. Appendix Figure A18 verifies this by showing a similar estimate using the DID_M estimator. The figure further shows no pre-trends as the estimates from the DID_M^{pl} estimators are also small and statistically insignificant. Appendix Table A18 zooms in on various individual pollutants and the results are similar.

I Comparing Plants to Airports

There is a concern that crops harvested around the airports may differ systematically from those around the plants. This will not affect the internal validity of my findings, but may invalidate the contrast between the airports and plants. To alleviate such concern, I use the 2005 global gridded dataset on agricultural production from Yu et al. (2020) to check the baseline difference.[72] The data contain the harvested area for almost all crops at the 5 arcmin $\times$ 5 arcmin grid cell level.[73] I extract the grid cells whose centroids are within 15 km of the plants or within 10-20 km of the airports.[74] To increase statistical power, I focus on the top 5 major crops: wheat, rice, maize, soybean, and vegetables. These crops combined account for about 70% of the total harvested area. I then calculate each crop's share of the total harvested area around the plants and airports. Appendix Figure A19 Panel A shows the mean difference between the airports and the plants. The difference in crop share is relatively small and only statistically significant for wheat and maize. I further adopt a matching procedure to balance the crop share (Abadie and Imbens, 2006). Specifically, I match the plants to the airports without replacement on these five crop share variables and discard those lacking common support in the propensity

[72] As stated in Appendix D, the data are also available for 2000 and 2010. The 2000 data are not used in my analysis because one major type of crops—vegetables—is missed.

[73] 5 arcmin $\times$ 5 arcmin is about 7.2 km $\times$ 9 km at the center of China.

[74] The 10-20 km range ensures that the grid cells are roughly within the same bandwidth as used in the RD analysis.

score of being a plant.[75] The matched sample contains 144 plants and 144 airports. Panel B shows the mean difference in the matched sample, which now becomes not only economically small but also statistically insignificant. I then redo the baseline DID and RD analysis using the matched plants and airports. As shown in Appendix Table A20 and Appendix Table A28, the results are similar to those using the full sample.

J Difference-in-differences Estimation around Airports

The spatial RD estimates are only informative of the treatment effects at the boundaries. In reality, regulators may strategically focus on areas close to an airport instead of areas close to the zone boundaries. Following Burlig and Preonas (2022), I overcome such limitation by complementing the spatial RD design with a difference-in-differences model that takes into consideration areas far away from the boundaries. Specifically, I estimate:

$$Y_{ijt} = \delta_i + \mu_{jt} + \beta Inside_{ij} \times Post_{jt} + \varepsilon_{ijt} \tag{A11}$$

where i, j, and t denote grid cells, airports, and years, respectively. The baseline sample includes areas as far as 10 km away from the zone boundaries, although the results are robust to alternative bandwidth such as 15 km. δ_i and μ_{jt} denote grid cell fixed effects and airport $\times$ year fixed effects, respectively. $Inside_{ij}$ is an indicator equal to 1 if grid cell i is within the no burning zones of airport j and 0 otherwise, and $Post_{jt}$ is an indicator equal to 1 if airport j is active at year t and zero otherwise. I only include airports that opened between 2002 and 2015, in order to have at least one year of observations before the opening and at least one year of observations after the opening. Standard errors are clustered at the airport level. The coefficient of interest is β, which measures the change in fires after the opening in areas inside the zones, relative to areas outside. The results using this specification are reported in Panel A of Appendix Table A30. In Panel B I include all areas within 15 km of the zone boundaries. The results are consistently small and statistically insignificant across all specifications. Appendix Figure A20

[75]Ideally, one may want to match on preexisting variables such as those measured in 2000; this is not possible for me as the data for 2000 lack a major type of crops. I compare the crop share in 2005 and 2010 and find little change over time. Hence, the data for 2005 should largely capture the preexisting crop share.

further addresses the recent econometric concerns about staggered DID designs and verifies the common trends assumption (de Chaisemartin and d'Haultfoeuille, 2020). In sum, both the RD and DID results points to the failure of the regulation in reducing fires.

Appendix References

Abadie, Alberto, and Guido W Imbens. 2006. "Large sample properties of matching estimators for average treatment effects." *Econometrica*, 74(1): 235–267.

Albrecht, R, S Goodman, D Buechler, R Blakeslee, and H Christian. 2016. "LIS 0.1 Degree very high resolution gridded lightning climatology data collection." *Data sets available online [https://ghrc. nsstc. nasa. gov/pub/lis/climatology/LIS/] from the NASA Global Hydrology Resource Center DAAC, Huntsville, Alabama, USA doi: http://dx. doi. org/10.5067/LIS/LIS/DATA306*.

Barwick, Panle Jia, Shanjun Li, Deyu Rao, and Nahim Bin Zahur. 2021. "The Healthcare Cost of Air Pollution: Evidence from the Worlds Largest Payment Network." National Bureau of Economic Research.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How much should we trust differences-in-differences estimates?" *The Quarterly Journal of Economics*, 119(1): 249–275.

Brock, William A, and M Scott Taylor. 2005. "Economic growth and the environment: a review of theory and empirics." *Handbook of Economic Growth*, 1: 1749–1821.

Burlig, Fiona, and Louis Preonas. 2022. "Out of the darkness and into the light? Development effects of rural electrification." *Journal of Political Economy*.

Cai, Xiqian, Yi Lu, Mingqin Wu, and Linhui Yu. 2016. "Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China." *Journal of Development Economics*, 123: 73–85.

Callaway, Brantly, and Pedro HC Sant'Anna. 2020. "Difference-in-differences with multiple time periods." *Journal of Econometrics*.

Cameron, A Colin, and Douglas L Miller. 2015. "A practitioner's guide to cluster-robust inference." *Journal of Human Resources*, 50(2): 317–372.

Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi. 2013. "The promise of Beijing: Evaluating the impact of the 2008 Olympic Games on air quality." *Journal of Environmental Economics and Management*, 66(3): 424–443.

Chen, Yvonne Jie, Pei Li, and Yi Lu. 2018. "Career concerns and multitasking local bureaucrats: Evidence of a target-based performance evaluation system in China." *Journal of Development Economics*, 133: 84–101.

Ciacci, Riccardo, and Mara Micaela Sviatschi. 2021. "The Effect of Adult Entertainment Establishments on Sex Crime: Evidence from New York City." *The Economic Journal*.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker. 2015. "Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings." *American Economic Review*, 105(2): 678–709.

de Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review*, 110(9): 2964–96.

Deschênes, Olivier, and Kyle C Meng. 2018. "Quasi-experimental methods in environmental economics: Opportunities and challenges." *Handbook of Environmental Economics*, 4: 285.

Deschênes, Olivier, and Michael Greenstone. 2011. "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US." *American Economic Journal: Applied Economics*, 3(4): 152–85.

Dijkstra, Edsger W, et al. 1959. "A note on two problems in connexion with graphs." *Numerische mathematik*, 1(1): 269–271.

Ebenstein, Avraham. 2012. "The consequences of industrialization: evidence from water pollution and digestive cancers in China." *Review of Economics and Statistics*, 94(1): 186–201.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. "New evidence on the impact of sustained exposure to air pollution on life expectancy from Chinas Huai River Policy." *Proceedings of the National Academy of Sciences*, 114(39): 10384–10389.

Faber, Benjamin. 2014. "Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System." *Review of Economic Studies*, 81(3): 1046–1070.

Fan, Maoyong, Guojun He, and Maigeng Zhou. 2020. "The winter choke: coal-fired heating, air pollution, and mortality in China." *Journal of health economics*, 71: 102316.

Fowlie, Meredith, Edward Rubin, and Reed Walker. 2019. "Bringing satellite-based air quality estimates down to earth." *AEA Papers and Proceedings*, 109: 283–88.

Ghanem, Dalia, and Junjie Zhang. 2014. "Effortless Perfection:Do Chinese cities manipulate air pollution data?" *Journal of Environmental Economics and Management*, 68(2): 203–225.

Gillings, Mark, Piraye Hacıgüzeller, and Gary Lock. 2020. *Archaeological spatial analysis: a methodological guide.* Routledge.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics*.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. 2020. "Can Technology Solve the Principal-Agent Problem? Evidence from Chinas War on Air Pollution." National Bureau of Economic Research.

Greenstone, Michael, Guojun He, Shanjun Li, and Eric Yongchen Zou. 2021. "Chinas War on Pollution: Evidence from the First 5 Years." *Review of Environmental Economics and Policy*, 15(2): 281–299.

Hammer, Melanie S, Aaron van Donkelaar, Chi Li, Alexei Lyapustin, Andrew M Sayer, N Christina Hsu, Robert C Levy, Michael J Garay, Olga V Kalashnikova, Ralph A Kahn, et al. 2020. "Global estimates and long-term trends of fine particulate matter concentrations (1998–2018)." *Environmental Science & Technology*, 54(13): 7879–7890.

He, Guojun, and Jeffrey M Perloff. 2016. "Surface water quality and infant mortality in China." *Economic Development and Cultural Change*, 65(1): 119–139.

He, Guojun, Maoyong Fan, and Maigeng Zhou. 2016. "The effect of air pollution on mortality in China: Evidence from the 2008 Beijing Olympic Games." *Journal of Environmental Economics and Management*, 79: 18–39.

Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal bandwidth choice for the regression discontinuity estimator." *The Review of Economic Studies*, 79(3): 933–959.

Liu, Mengdi, Ruipeng Tan, and Bing Zhang. 2021. "The costs of blue sky: environmental regulation, technology upgrading, and labor demand in China." *Journal of Development Economics*, 150: 102610.

Lobell, David B, and Jennifer A Burney. 2021. "Cleaner air has contributed one-fifth of US maize and soybean yield gains since 1999." *Environmental Research Letters*, 16(7): 074049.

Lv, Baolei, Jun Cai, Bing Xu, and Yuqi Bai. 2017. "Understanding the rising phase of the PM 2.5 concentration evolution in large China cities." *Scientific Reports*, 7(1): 1–12.

Manacorda, Marco, and Andrea Tesei. 2020. "Liberation technology: Mobile phones and political mobilization in Africa." *Econometrica*, 88(2): 533–567.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. 2015. "The housing market impacts of shale gas development." *American Economic Review*, 105(12): 3633–59.

Oster, Emily. 2019. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics*, 37(2): 187–204.

Pingel, Thomas J. 2010. "Modeling slope as a contributor to route selection in mountainous areas." *Cartography and Geographic Information Science*, 37(2): 137–148.

Said, RK, MB Cohen, and US Inan. 2013. "Highly intense lightning over the oceans: Estimated peak currents from global GLD360 observations." *Journal of Geophysical Research: Atmospheres*, 118(13): 6905–6915.

Simões Amaral, Simone, João Andrade de Carvalho, Maria Angélica Martins Costa, and Cleverson Pinheiro. 2016. "Particulate matter emission factors for biomass combustion." *Atmosphere*, 7(11): 141.

Tanaka, Shinsuke. 2015. "Environmental regulations on air pollution in China and their impact on infant mortality." *Journal of Health Economics*, 42: 90–103.

Tobler, Waldo. 1993. "Three presentations on geographical analysis and modeling." National Center for Geographic Information and Analysis Technical Report 93-1.

Van Donkelaar, Aaron, Randall V Martin, Chi Li, and Richard T Burnett. 2019. "Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors." *Environmental Science & Technology*, 53(5): 2595–2611.

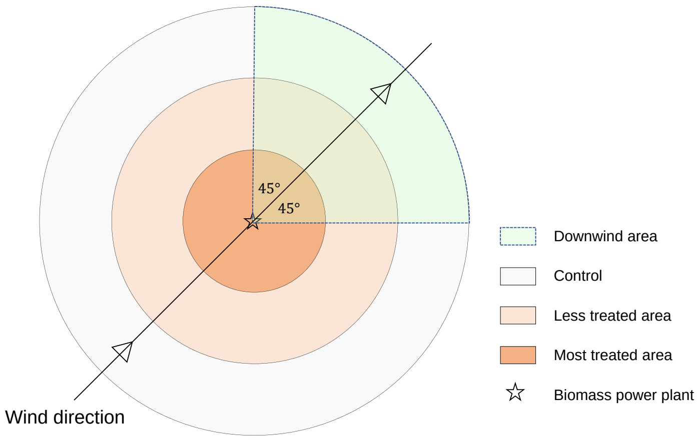
Xu, Yongming, Benjamin Lewis, and Weihe Wendy Guan. 2019. "Developing the Chinese Academic Map Publishing Platform." *ISPRS International Journal of Geo-Information*, 8(12): 567.

Yu, Qiangyi, Liangzhi You, Ulrike Wood-Sichra, Yating Ru, Alison KB Joglekar, Steffen Fritz, Wei Xiong, Miao Lu, Wenbin Wu, and Peng Yang. 2020. "A cultivated planet in 2010–Part 2: the global gridded agricultural-production maps." *Earth System Science Data*, 12(4): 3545–3572.

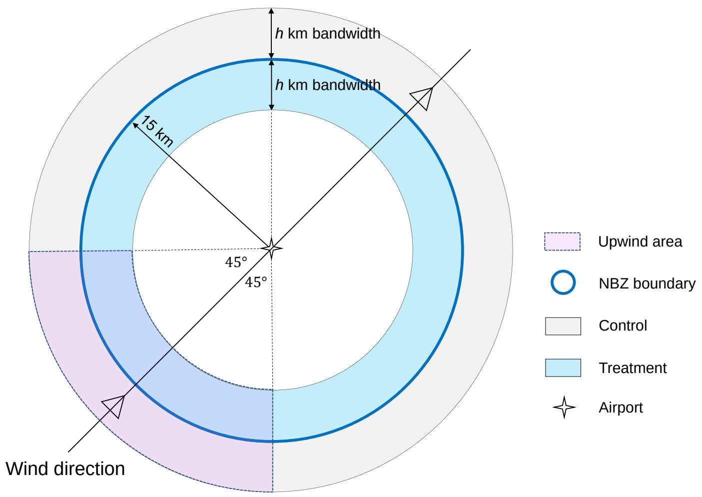
K Appendix Figures and Tables

Figure A1: Illustration of Downwind Area and Upwind Area

A: Downwind area of biomass power plants



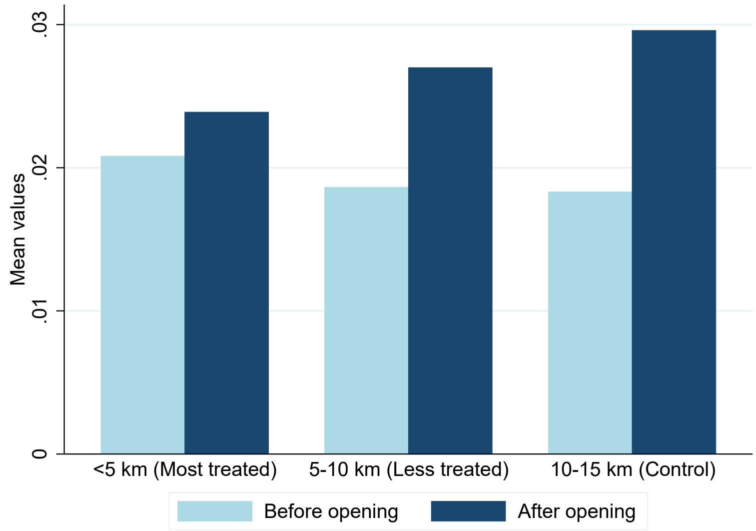
B: Upwind area of airports



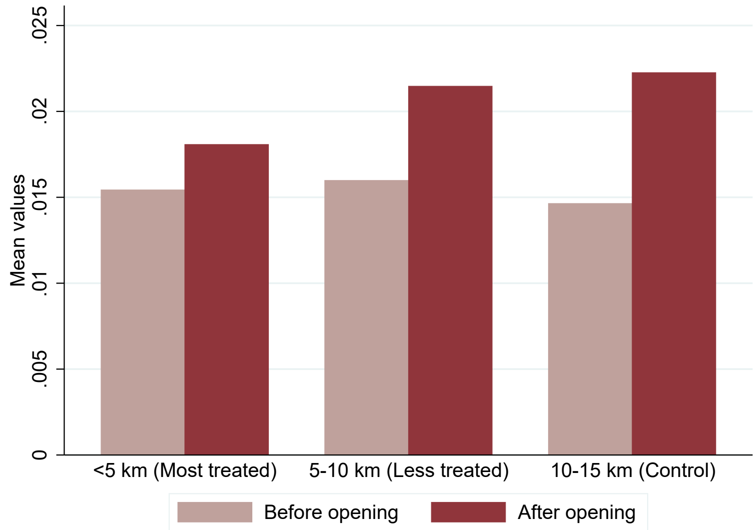
Notes: The unit of observation is 1 km × 1 km grid cells. Panel A illustrates the downwind areas of a biomass power plant, which are perceived to suffer more from potential pollution from the plant. Panel B illustrates the upwind areas of an airport, which are supposed to have more stringent enforcement of the no burning ban due to safety requirements. The wind direction denotes daily wind direction determined using the vector decomposition method and the wind direction data at three-hour frequencies from the NOAA.

Figure A2: Biomass Power Plants: Changes in Agricultural Fires by Treatment Status

A: Number of fires



B: Presence of fires

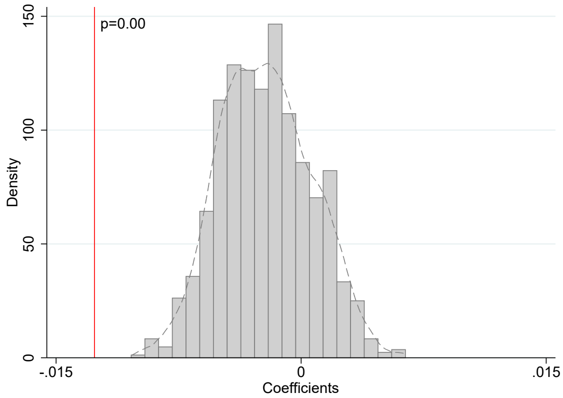


Notes: This figure shows the mean outcomes in the treatment and control groups, before and after the opening of a biomass power plant. *Presence of fires* is an indicator for observing at least one fire.

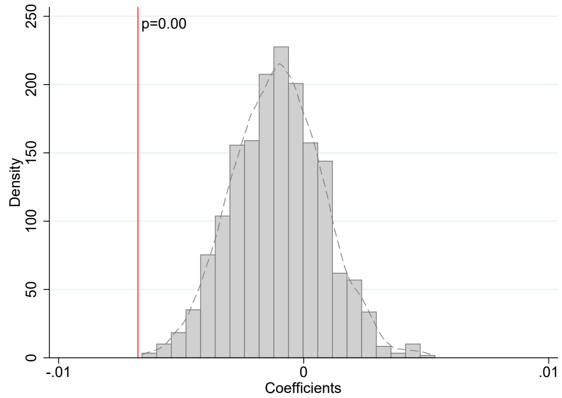
Figure A3: Biomass Power Plants: Permutation Test

Number of fires

A: Effect on most treated areas

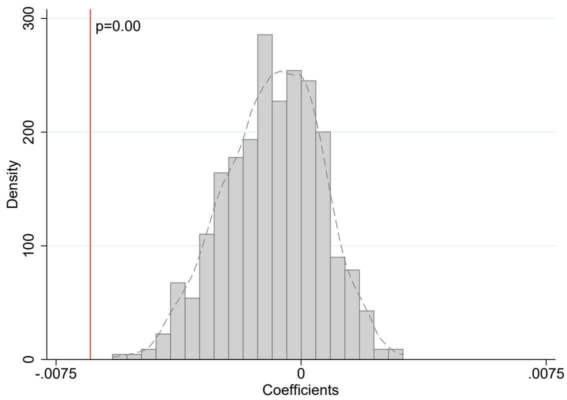


B: Effect on less treated areas

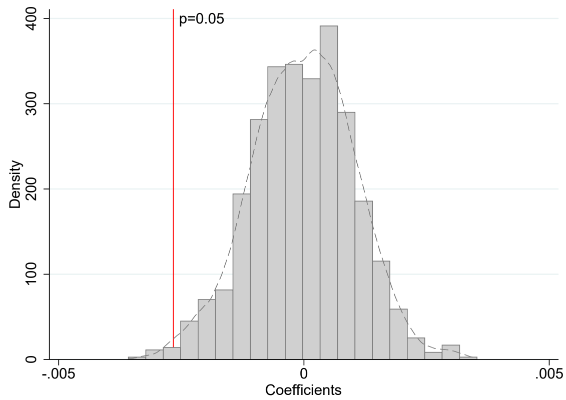


Presence of fires

C: Effect on most treated areas

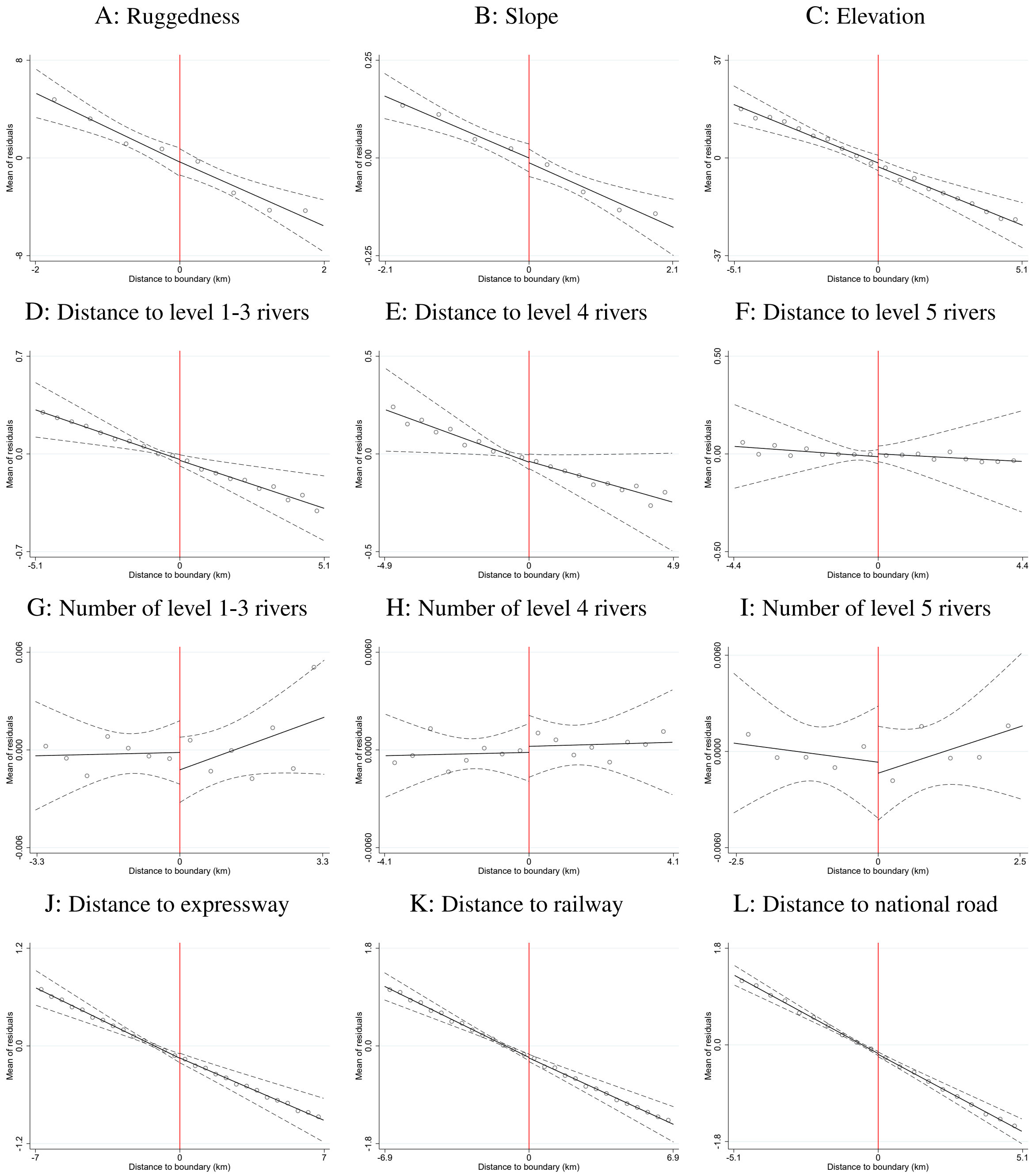


D: Effect on less treated areas



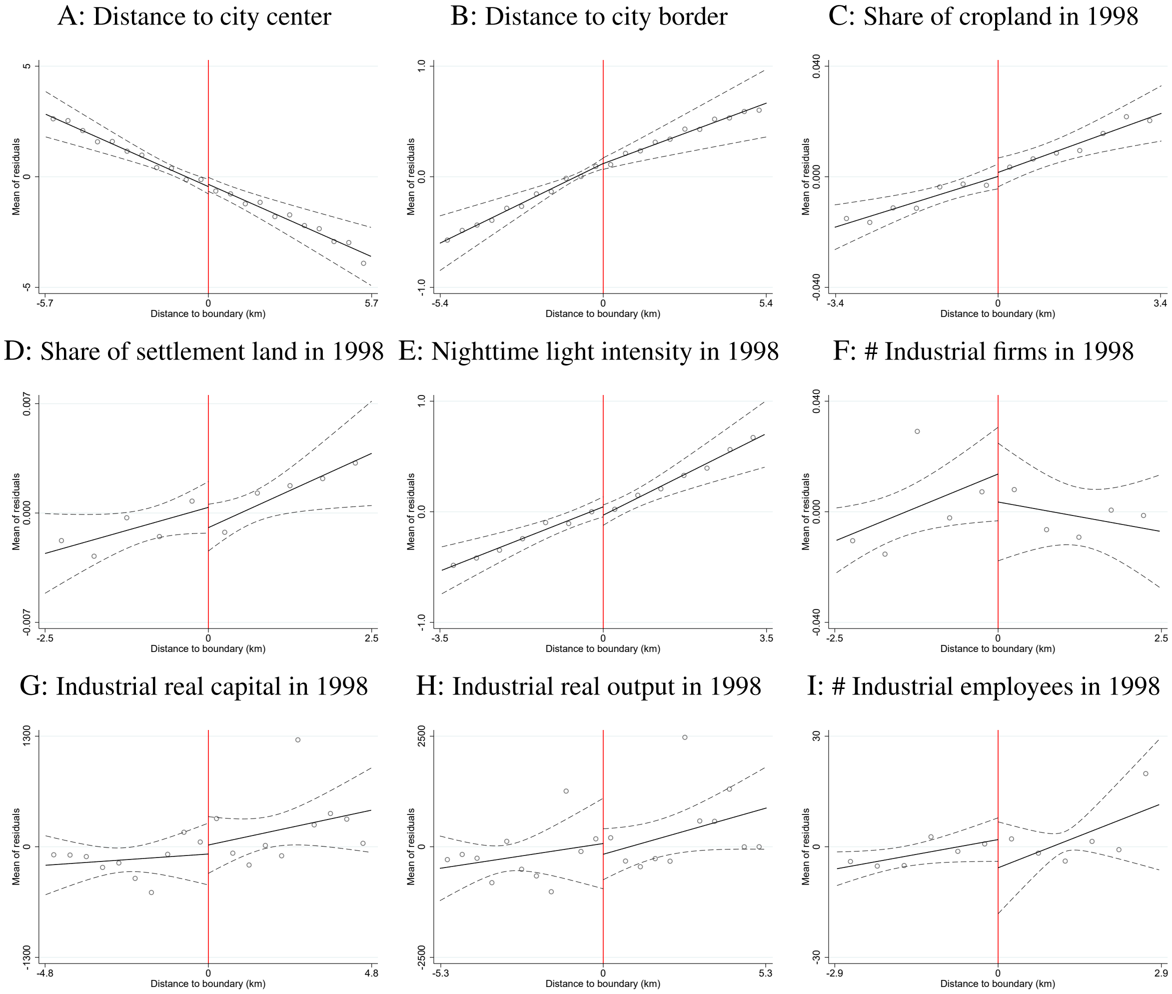
Notes: The unit of observation is 1 km × 1 km grid cells. This figure conducts a placebo test by permutating the timing of plant openings, using the specifications that control for grid cell fixed effects, plant × year fixed effects, grid cell-specific linear trends, and a set of geographic controls interacted with year fixed effects. *Presence of fires* is an indicator for observing at least one fire. The bars denote the distribution of coefficients from 1000 permutations, and the vertical lines denote the true coefficients. The *p*-values denote the relative position of the true coefficients among the placebo coefficients.

Figure A4: No Burning Zones: Balance Checks



Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This figure provides balance checks on preexisting factors on each side of the no burning zone boundaries, using the optimal bandwidth selected following the procedure proposed by Imbens and Kalyanaraman (2012). The circles denote the mean of outcomes within a distance bin, after partialling out the airport × boundary segment fixed effects. The lines are fitted separately for grid cells on each side of the boundaries, and the dashes are the associated 95% confidence intervals, where standard errors used to construct the confidence intervals are clustered at the airport level.

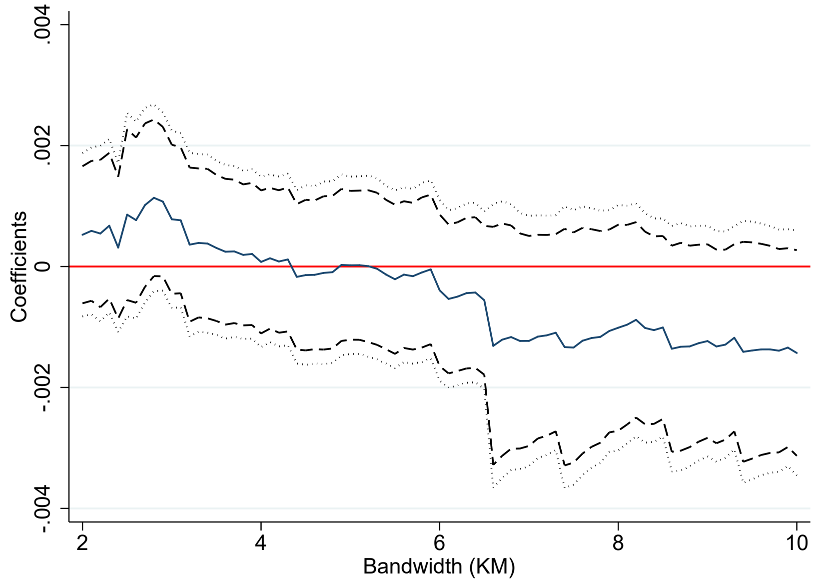
Figure A5: No Burning Zones: Balance Checks (Cont.)



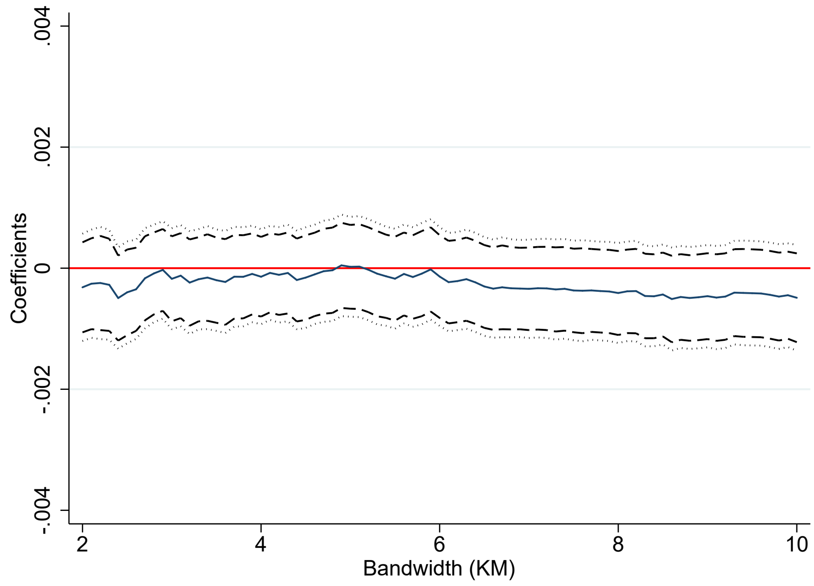
Notes: The unit of observation is 1 km $\times$ 1 km grid cells. The no burning zones are areas within 15 km of an airport. This figure provides balance checks on preexisting factors on each side of the no burning zone boundaries, using the optimal bandwidth selected following the procedure proposed by Imbens and Kalyanaraman (2012). The circles denote the mean of outcomes within a distance bin, after partialling out the airport $\times$ boundary segment fixed effects. The lines are fitted separately for grid cells on each side of the boundaries, and the dashes are the associated 95% confidence intervals, where standard errors used to construct the confidence intervals are clustered at the airport level.

Figure A6: No Burning Zones: Sensitivity to Bandwidth Choices

A: Number of fires

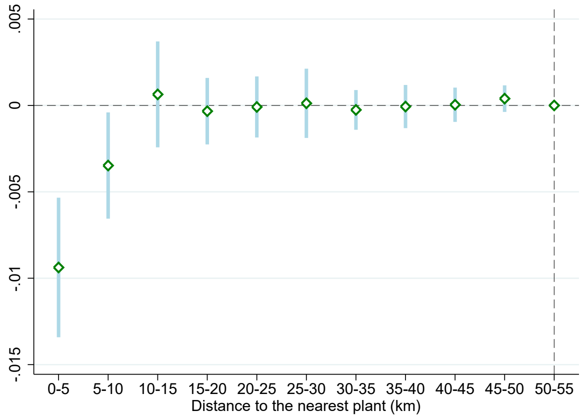


B: Presence of fires



Notes: The unit of observation is 1 km $\times$ 1 km grid cells. The no burning zones are areas within 15 km of an airport. This figure shows the robustness of the RD estimates to alternative bandwidths ranging from 2 km to 10 km in 0.1 km increments. *Presence of fires* is an indicator for observing at least one fire. The solid lines denote RD coefficients. The dashed and dotted lines denote the associated 90% and 95% confidence intervals, respectively, where the standard errors used to construct the confidence intervals are clustered at the airport level.

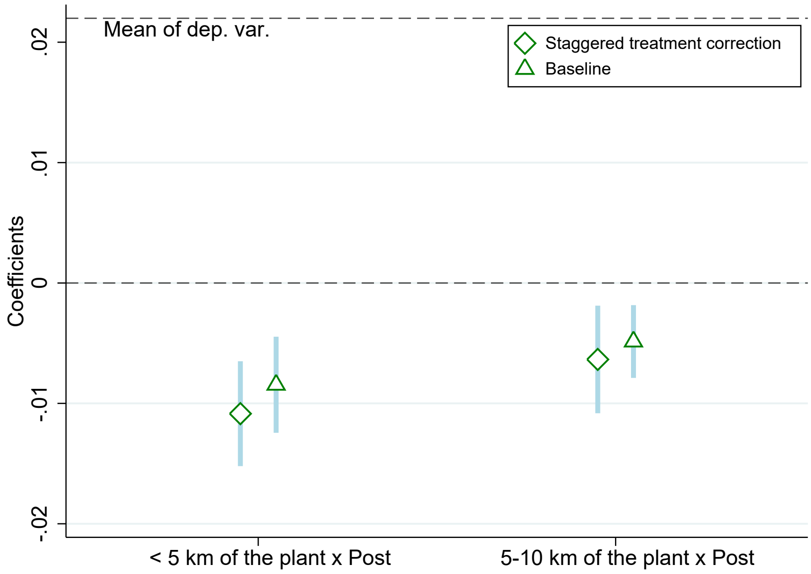
Figure A7: Biomass Power Plants: Determining Treatment and Control Group



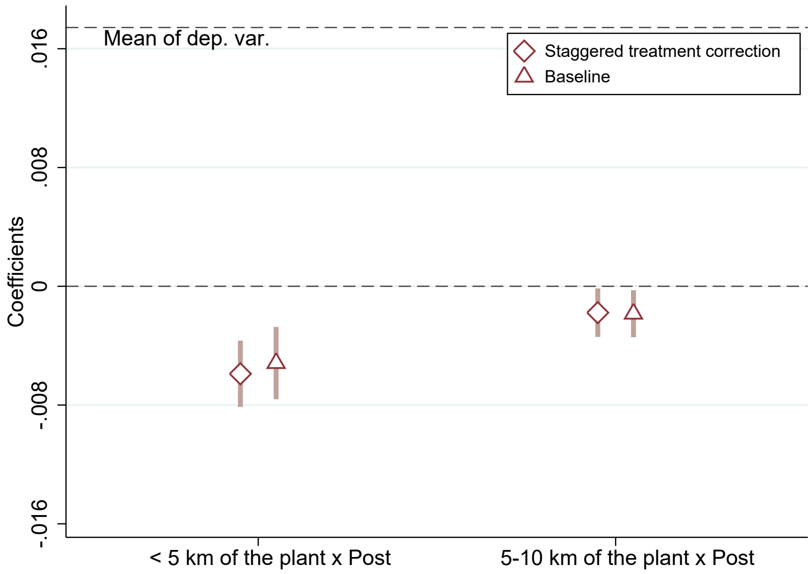
Notes: The unit of observation is 1 km × 1 km grid cells. The sample includes areas within 55 km of a plant. This figure plots the effects of a plant opening on the number of agricultural fires in 10 distance bins (each has a width of 5 km). The 50-55 km bin is omitted as the reference group. The diamonds denote coefficients and bars denote associated 95% confidence intervals, where standard errors used to construct the confidence intervals are clustered at the plant level.

Figure A8: Biomass Power Plants: Correcting for Staggered Treatment Timing

A: Number of fires

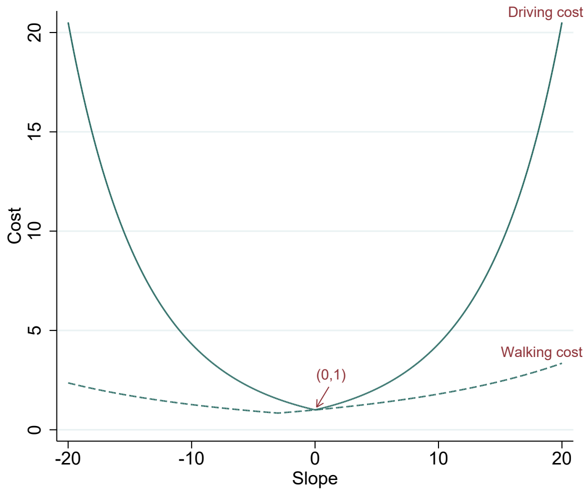


B: Presence of fires



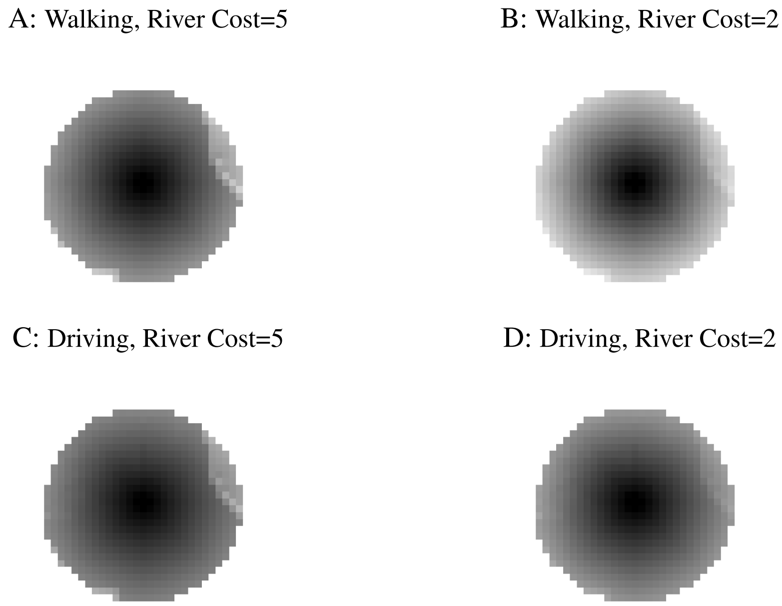
Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This figure shows the baseline estimates corrected for staggered treatment timing (denoted by the diamonds). For comparison, the baseline estimates are also plotted (denoted by the triangles). *Presence of fires* is an indicator for observing at least one fire. $<$ *5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. The bars denote the 95% confidence intervals, where standard errors used to construct the confidence intervals are clustered at the plant level. To correcting for the bias associated with staggered treatment timing, I follow Callaway and Sant'Anna (2020) by estimating equation (1) for each treatment cohort separately (treatment cohort is defined by the year of plant opening) and then average the cohort-specific estimates using the size of each treatment cohort as weights.

Figure A9: Cost Functions



Notes: The horizontal axis denotes the slope (in degrees) from one grid cell to another, which is calculated using the vertical movement (using elevation changes) and horizontal movement (using grid cell size, also compensating for diagonal movement). The positive slope indicates ascents and negative slopes indicate descents. The slope-dependent cost functions, which model the cost with movement by foot or by car on rugged terrain, are taken from Tobler (1993) and Pingel (2010) and are normalized to one for zero slope.

Figure A10: Least Cost Distance Gradient

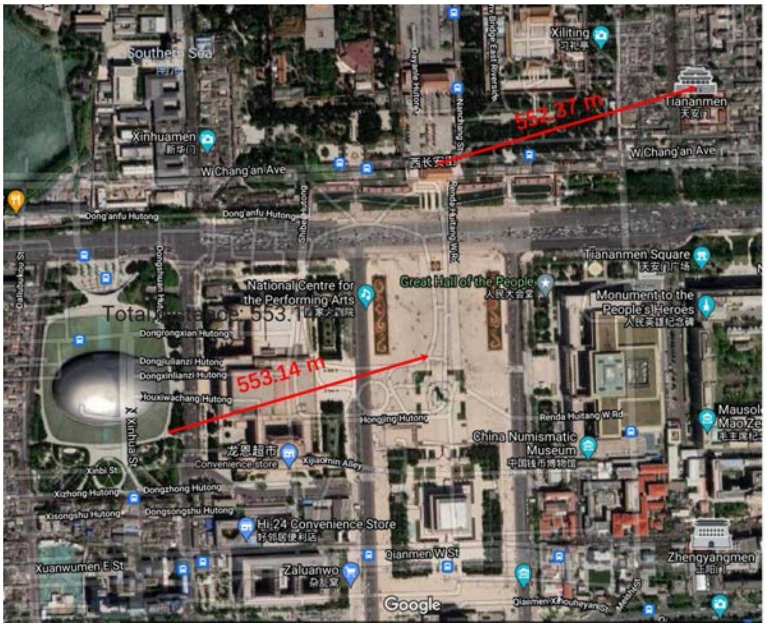


Notes: The size of the grid cells are 1 km $\times$ 1 km. This figure shows, for a typical plant located in the North China Plain, the least cost distance from each grid cell to the grid cell where the nearest plant is located, with lighter color denotes higher cost distance, in different scenarios. The cost distance is calculated by imposing penalties on movements on rugged terrain and movements across grid cells with rivers. Panel A and B model the walking cost on rugged terrain. Panel C and D model the driving cost on rugged terrain. Panel A and C assign a unit distance cost of

5 to grid cells with rivers. Panel B and D assign a unit distance cost of 2 to grid cells with rivers.

Figure A11: Illustration of Map Censoring

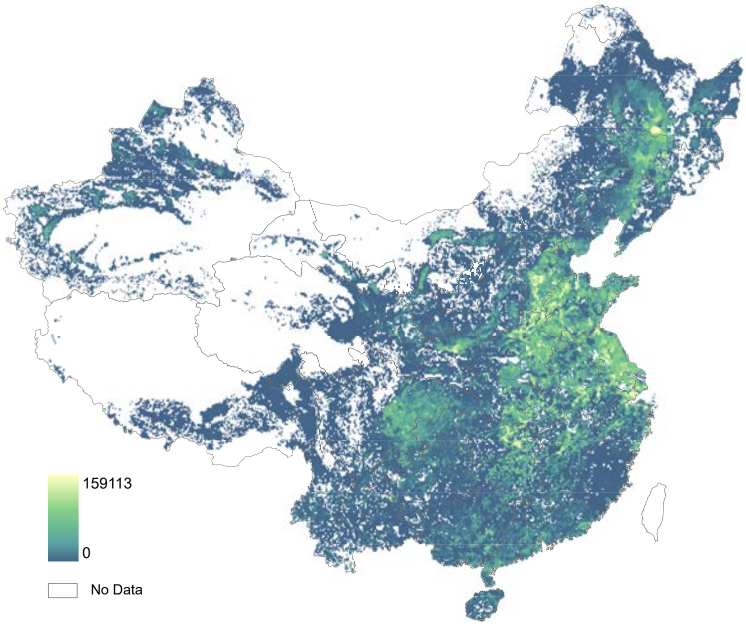
A: Place 1



B: Place 2

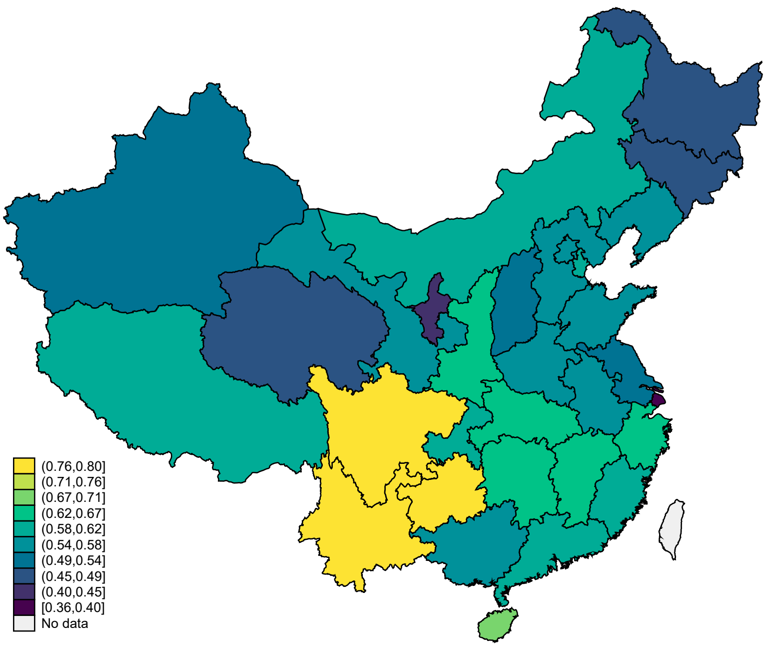
Notes: This figure shows the degree of map "censoring" in China by overlaying the "censored" street map shown in translucent color (in GCJ02 coordinate system) onto the satellite map (in WGS84 coordinate system), using Google Maps. "Censoring" means that the GCJ02 coordinate, which is used in Chinese street map, adds an offset up to several hundred meters to the WGS84 coordinate. Panel A shows a place at the center of Beijing while Panel B shows a place 8 km east of the place in Panel A. In each place, there are two red arrows with the starts denoting the locations of two nearby buildings (within 1 km) in the WGS84 satellite map and the ends denoting their mappings into the GCJ02 street map. The lengths of the arrows measured in meters are also labeled in red on the arrows. This figure shows two findings: (1) at two nearby locations within the same place (i.e., within Panel A or Panel B), the offsets (or the degree of "censoring") are virtually identical: the directions are parallel and lengths differ by less than 1 meter; (2) across two places that are several km away from each other (i.e., Panel A versus Panel B), the difference in the offsets (or the degree of ”censoring”) becomes slightly larger by about 10 meters (the directions are still quite parallel).

Figure A12: Crop Residues In 2000



Notes: This figure shows the amount of crop residues in 2000 (in tons), calculated using equation (A3) with the data on harvested area, yield, and straw-to-grain ratio for each crop in each 5 arcmin $\times$ 5 arcmin grid cell in China.

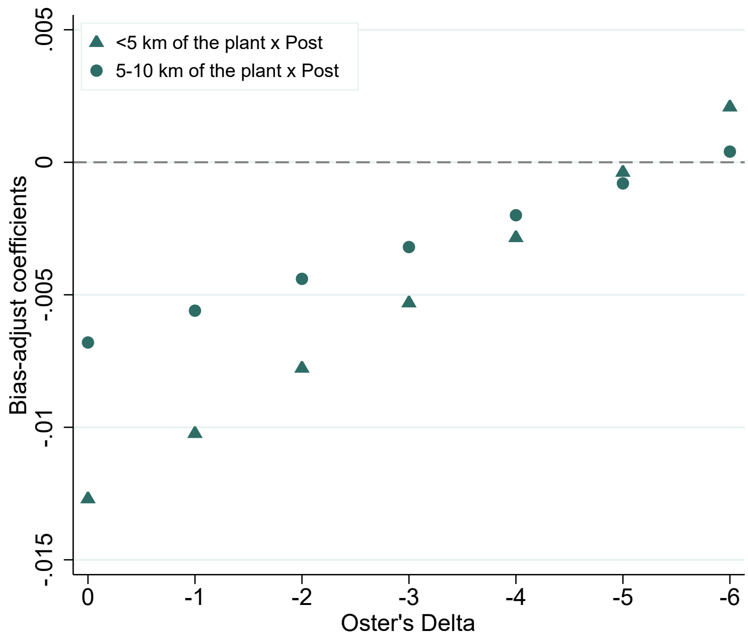
Figure A13: Biomass Power Plants: Expected Treatment Value



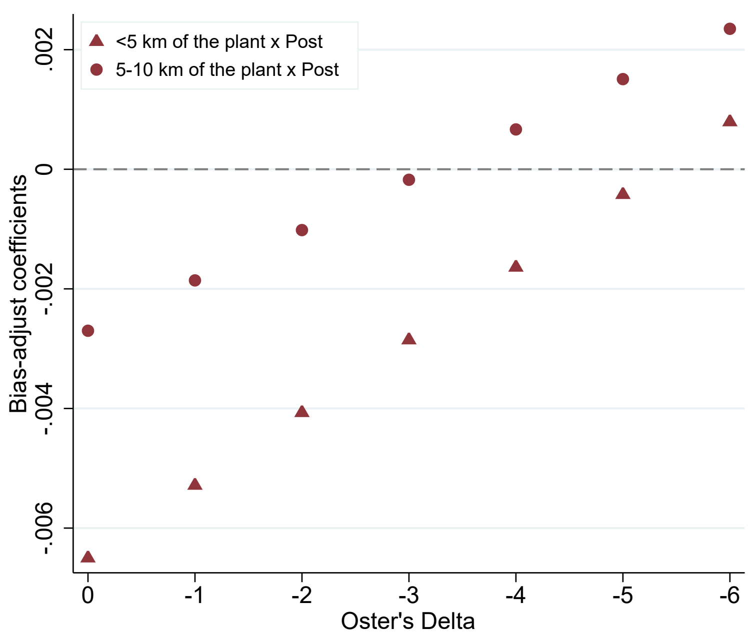
Notes: This figure shows the distribution of expected treatment value in the year of a plant opening, which is calculated using equation (A5) by assuming a uniform distribution of plant openings over the year and using the empirical distribution of crop residues over the year across provinces. Note that this value can be calculated for provinces without a plant opening as it is based on hypothetical distribution of plant openings over the year.

Figure A14: Biomass Power Plants: Correcting for Selection on Unobservables

A: Number of fires



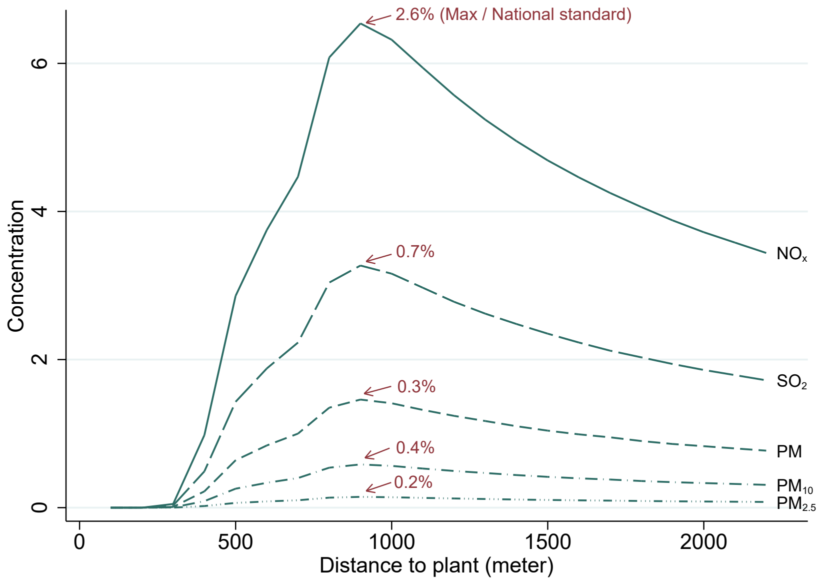
B: Presence of fires



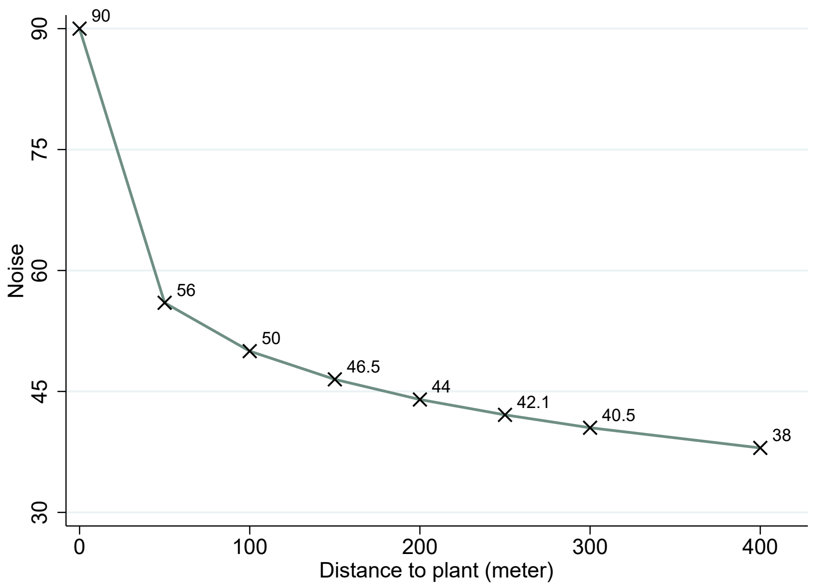
Notes: This figure shows the coefficients adjusted for selection on unobservables using the method proposed by Oster (2019). *Presence of fires* is an indicator for observing at least one fire. Delta denotes the degree of selection on unobservables relative to observables. Delta=0 denotes no selection on unobservables (the baseline results). The negative values of Delta denote that selection on unobservables is negatively related to selection on observables.

Figure A15: Air Pollutants and Noise Generated by A Typical Plant

A: Air pollutants per hour (μg/m^3)



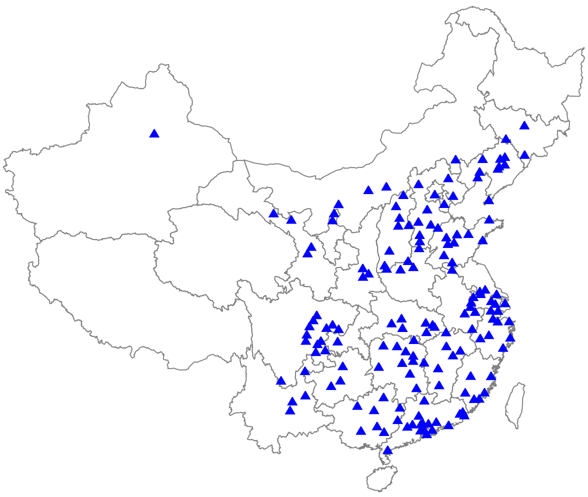
B: Noise (dB)



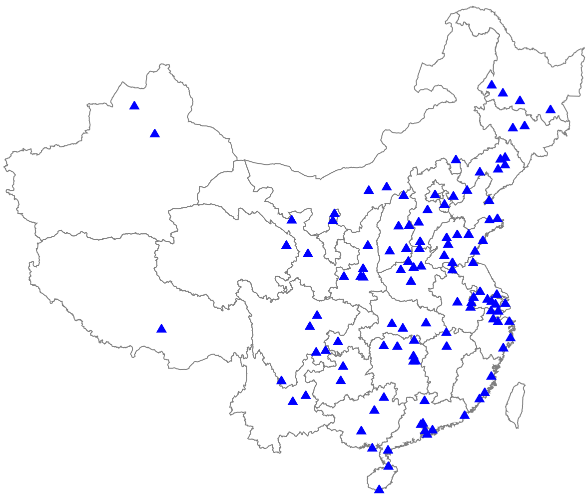
Notes: Panel A shows the amount of air pollutants generated by a typical plant in an hour and the dispersion with distance. Panel B shows the noise generated by a typical plant and the decay with distance.

Figure A16: Other City-level Environmental Regulations

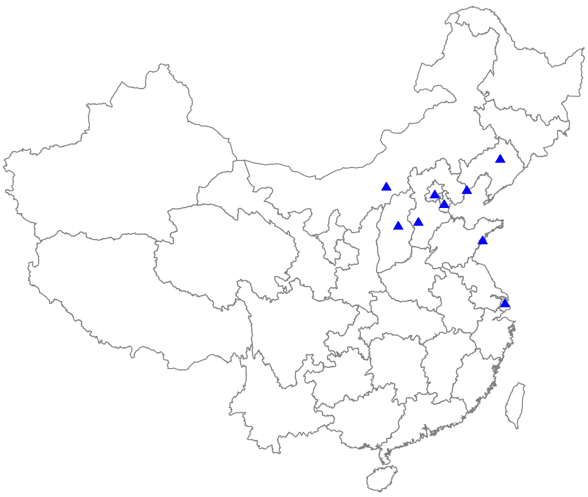
A: Two Control Zones (1998)



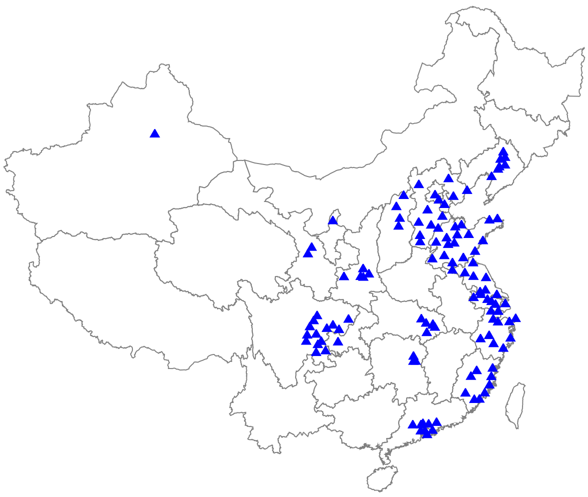
B: Key Cities for Air Pollution Control (2001)



C: Olympics Host and Neighbor Cities (2008)

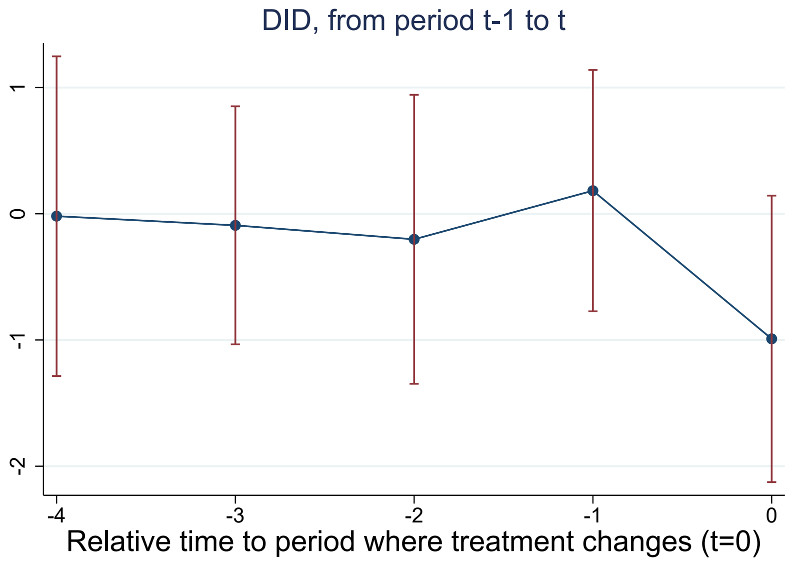


D: Key Regions for Air Pollution Control (2012)



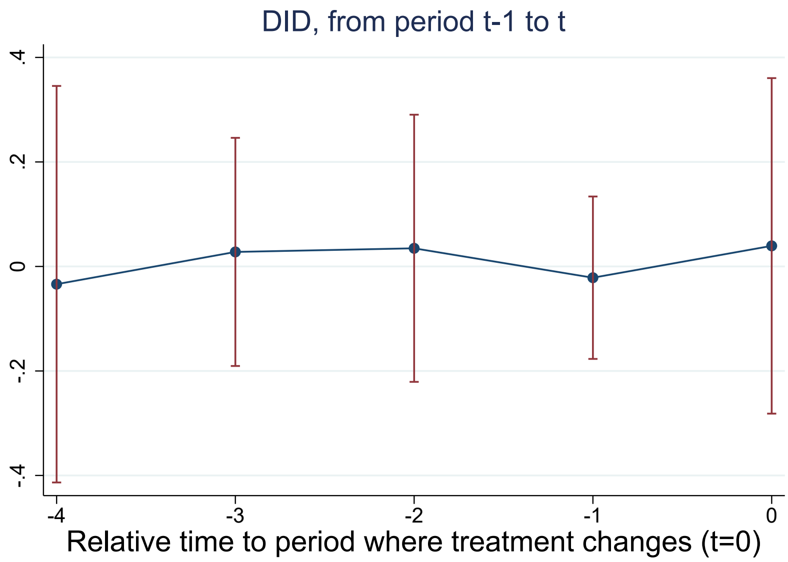
Notes: This figure shows the locations of the target cities of four confounding city-level environmental regulations. Panel A shows the locations of the Two Control Zones cities designated in 1998 with the goal of reducing acid rain and SO_2 pollution. Panel B shows the locations of the Key Cities for Air Pollution Control cities designated in 2001 with the goal of reducing acid rain, SO_2, and particulate matter. Panel C shows the locations of the host cities of the 2008 Beijing Olympic Games as well as a few neighbor cities, which faced radical air pollution regulation during 2007-2008. Panel D shows the locations of the Key Regions for Air Pollution Control cities designated in 2012 with the goal of reducing primary and secondary pollution and $PM_{2.5}$.

Figure A17: The Effect of Plant Openings on $PM_{2.5}$



Notes: This figure shows the effect of plant openings on $PM_{2.5}$ using the DID_M and DID_M^{pl} estimators from de Chaisemartin and d'Haultfoeuille (2020). The dependent variable is average ambient $PM_{2.5}$ concentration within 10 km of a plant. The year of a plant opening is normalized to $t = 0$. DID_M corresponds to $t = 0$, which compares the $t = -1$ to $t = 0$ outcome evolution, in units with treatment changes from $t = -1$ to $t = 0$, and in units without treatment changes from $t = -1$ to $t = 0$. DID_M^{pl} corresponds to $t = -1, -2, -3, -4$, which compares the $t-1$ to t outcome evolution, in units whose treatment status does not change from $t-1$ to $t = -1$. DID_M^{pl} is used to test the common trends assumption. The standard errors used to construct the 95% confidence intervals are calculated using a block bootstrap at the plant level.

Figure A18: The Effect of Plant Openings on Water Quality

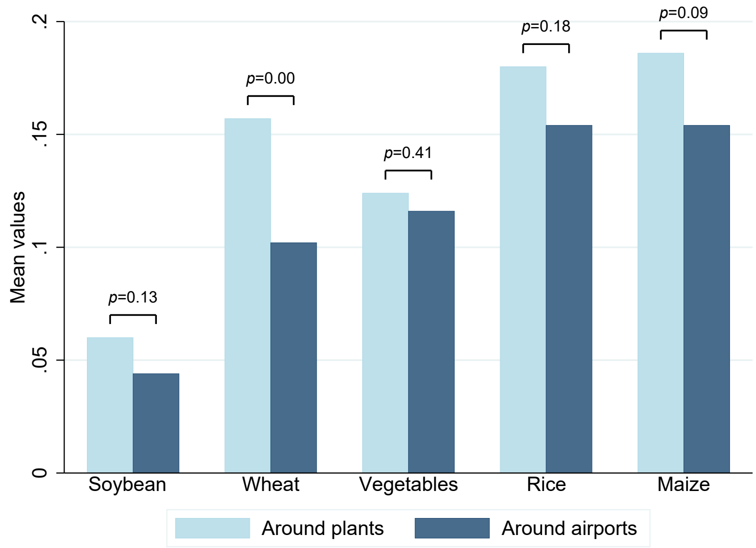


Notes: This figure shows the effect of nearby plant openings on water quality using the DID_M and DID_M^{pl} estimators from de Chaisemartin and d'Haultfoeuille (2020). The dependent variable is a water quality index ranging from 1 to 6 with lower values denoting better quality. The year of a nearby plant opening is normalized to $t = 0$. DID_M corresponds to $t = 0$, which compares the $t = -1$ to $t = 0$ outcome evolution, in units with treatment changes from $t = -1$ to $t = 0$, and in units without treatment changes from $t = -1$ to $t = 0$. DID_M^{pl} corresponds to $t = -1, -2, -3, -4$, which compares the $t-1$ to t outcome evolution, in units whose treatment status does not change from $t-1$ to $t = -1$.

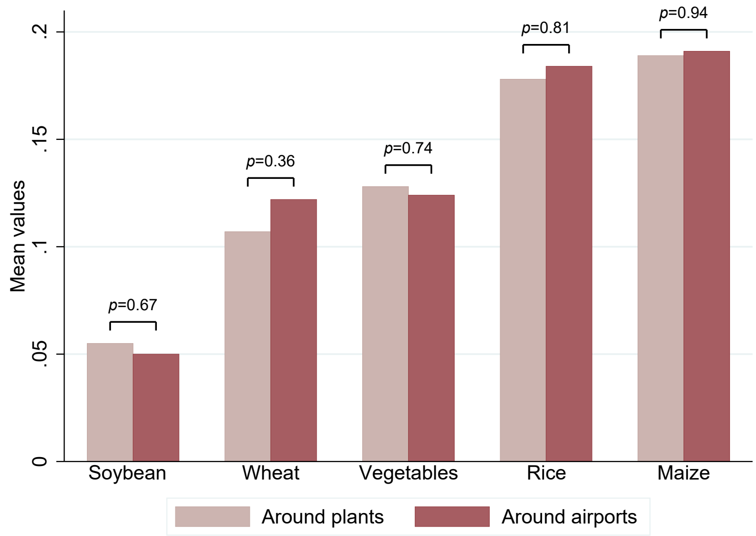
DID_M^{pl} is used to test the common trends assumption. The standard errors used to construct the 95% confidence intervals are calculated using a block bootstrap at the water quality monitoring station level.

Figure A19: Comparing Crop Share around Plants and Airports

A: Crop share in full sample



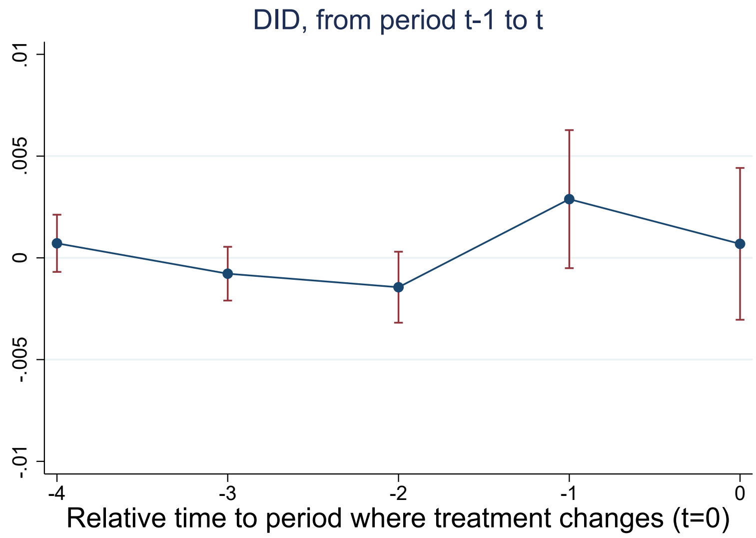
B: Crop share in matched sample



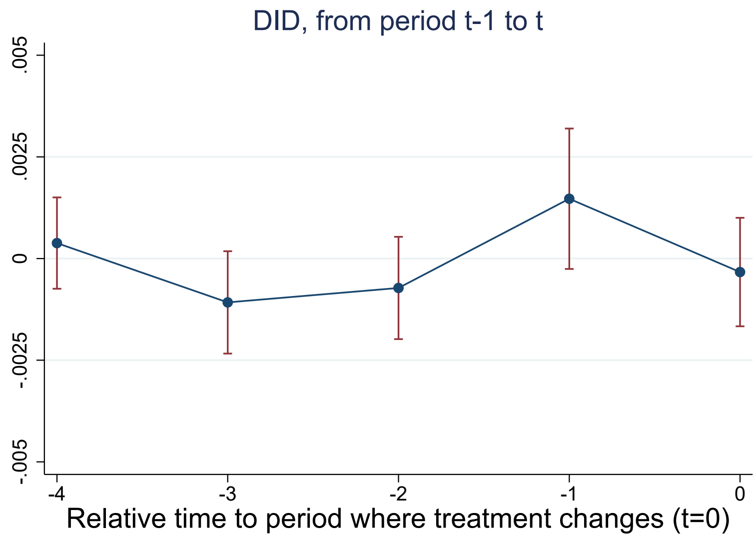
Notes: This figure shows the mean crop share for five major crops harvested around plants and airports, where crop share is calculated as the harvested area for a crop divided by the total harvested area for all crops around a plant (or an airport). The p-values for t-tests of equality of means between plants and airports are also reported on top of each pair of bars. Panel A uses the full sample, whereas Panel B uses a matched sample as described in Appendix I.

Figure A20: Difference-in-Differences Estimation around Airports

A: Fire Number



B: Fire Presence



Notes: This figure shows the effect of the no burning zones using the DID_M and DID_M^{pl} estimators from de Chaisemartin and d'Haultfoeuille (2020), focusing on areas within 10 km of the zone boundaries. The year of an airport opening is normalized to $t = 0$. DID_M corresponds to $t = 0$, which compares the $t = -1$ to $t = 0$ outcome evolution, in units with treatment changes from $t = -1$ to $t = 0$, and in units without treatment changes from $t = -1$ to $t = 0$. DID_M^{pl} corresponds to $t = -1, -2, -3, -4$, which compares the $t - 1$ to t outcome evolution, in units whose treatment status does not change from $t - 1$ to $t = -1$. DID_M^{pl} is used to test the common trends assumption. The standard errors used to construct the 95% confidence intervals are calculated using a block bootstrap at the airport level.

Table A1: Biomass Power Plants: Summary Statistics

Variable	**Obs**	**Mean**	**Stdev**	**P25**	**P50**	**P75**
Number of fires	1986576	0.022	0.267	0	0	0
Presence of fires	1986576	0.017	0.131	0	0	0
The year of opening of a plant	1986576	2010.355	3.331	2008	2010	2013
Distance to the nearest plant	1986576	9.619	3.405	7.212	10.200	12.490
<5 km of the plant	1986576	0.120	0.325	0	0	0
5-10 km of the plant	1986576	0.361	0.480	0	0	1
Longitude	1986576	116.977	6.028	113.883	117.085	119.807
Latitude	1986576	34.703	6.157	30.948	34.037	37.876
Ruggedness (in meter)	1986576	50.157	86.965	6.928	15.524	46.347
Slope (in percentage point)	1986576	1.431	2.511	0.194	0.438	1.311
Elevation (in meter)	1986576	189.982	343.232	20	53	189
Distance to level 1-3 rivers (in km)	1986576	45.611	46.910	8.600	30.293	72.144
Distance to level 4 rivers (in km)	1986576	51.668	45.088	15.448	38.899	75.978
Distance to level 5 rivers (in km)	1986576	10.788	13.742	2.720	6.991	13.789
Number of level 1-3 rivers (in km)	1986576	0.037	0.215	0	0	0
Number of level 4 rivers (in km)	1986576	0.014	0.122	0	0	0
Number of level 5 rivers (in km)	1986576	0.053	0.235	0	0	0
Distance to expressways (in km)	1986576	96.826	137.678	13.275	40.948	116.099
Distance to railways (in km)	1986576	22.938	29.924	4.262	11.439	30.017
Distance to national roads (in km)	1986576	14.902	16.443	3.611	9.655	21.622
Number of non-agricultural fires	1986576	0.016	0.822	0	0	0
Presence of non-agricultural fires	1986576	0.007	0.084	0	0	0
Share of cropland	1986576	0.734	0.367	0.5	1	1
Share of settlement land	1986576	0.076	0.208	0	0	0
Number of industrial firms	869127	0.203	2.191	0	0	0
Real capital of industrial firms (in 1000 *yuan*)	869127	3146.945	52792.234	0	0	0
Real output of industrial firms (in 1000 *yuan*)	744966	12247.771	178103.203	0	0	0
Number of employees of industrial firms	869127	42.323	503.809	0	0	0
Distance to city center (in km)	1986576	44.824	32.227	19.198	40.001	60.862
Distance to city border (in km)	1986576	15.663	12.740	5.850	12.941	22.384
Number of fragments created by rivers	1986576	1.104	0.344	1	1	1

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This table shows the summary statistics for main variables used in analyzing the effects of biomass power plants on agricultural fires. The sample contains areas within 15 km of a plant.

Table A2: No Burning Zones: Summary Statistics

Variable	**Obs**	**Mean**	**Stdev**	**P25**	**P50**	**P75**
Number of fires	1829849	0.011	0.140	0	0	0
Presence of fires	1829849	0.009	0.096	0	0	0
The opening year of an airport	1829849	1987.085	20.039	1974	1993	2003
Distance to boundary	1829849	-0.352	2.283	-2.349	-0.516	1.551
Inside	1829849	0.434	0.496	0	0	1
Longitude	1829849	110.434	11.327	104.697	111.986	118.899
Latitude	1829849	33.490	7.428	27.753	32.347	39.460
Ruggedness (in meter)	1829849	113.283	157.335	12.042	42.048	156.048
Slope (in percentage point)	1829849	3.252	4.553	0.341	1.202	4.476
Elevation (in meter)	1829849	729.846	980.319	43	324	1103
Distance to level 1-3 rivers (in km)	1829849	61.127	73.632	11.063	32.656	85.657
Distance to level 4 rivers (in km)	1829849	66.341	77.670	12.289	37.634	97.156
Distance to level 5 rivers (in km)	1829849	18.968	21.534	6.017	13.329	24.803
number of level 1-3 rivers (in km)	1829849	0.025	0.174	0	0	0
number of level 4 rivers (in km)	1829849	0.015	0.126	0	0	0
number of level 5 rivers (in km)	1829849	0.027	0.169	0	0	0
Distance to expressway (in km)	1829849	216.002	258.150	15.236	130.807	302.718
Distance to railway (in km)	1829849	82.617	161.555	5.807	15.073	87.549
Distance to national road (in km)	1829849	14.174	24.334	3.524	8.575	15.818
Share of cropland	1829849	0.468	0.426	0	0.417	1
Share of settlement land	1829849	0.069	0.204	0	0	0
Number of industrial firms	655155	0.231	2.299	0	0	0
Real capital of industrial firms (in 1000 *yuan*)	655155	3729.715	50608.727	0	0	0
Real output of industrial firms (in 1000 *yuan*)	561664	13868.006	179087.156	0	0	0
Number of employees of industrial firms	655155	47.307	524.265	0	0	0
Nighttime light intensity	1386478	11.255	17.380	0	4	12
Distance to city center (in km)	1829849	44.762	62.392	15.554	23.843	40.131
Distance to city border (in km)	1829849	25.095	28.686	6.901	17.130	32.819

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This table shows the summary statistics for main variables used in analyzing the effects of no burning zones on agricultural fires. The no burning zones are areas within 15 km of an airport. The sample include areas within 4 km of the zone boundaries.

Table A3: Biomass Power Plants: Testing the Enhanced Inspection Mechanism Using Mayors' Incentives

	(1)	(2)
Dep. var. =	Number of fires	Presence of fires
$<$ 5 km of the plant × Post	-0.0090**	-0.0057**
	(0.0037)	(0.0022)
5-10 km of the plant × Post	-0.0051*	-0.0016
	(0.0031)	(0.0013)
High incentive	0.0004	0.0017
	(0.0054)	(0.0042)
High incentive × Post	0.0083*	0.0064
	(0.0047)	(0.0042)
High incentive × $<$ 5 km of the plant	0.0092	0.0017
	(0.0061)	(0.0024)
High incentive × 5-10 km of the plant	0.0009	0.0005
	(0.0024)	(0.0014)
High incentive × $<$ 5 km of the plant × Post	-0.0043	-0.0005
	(0.0060)	(0.0026)
High incentive × 5-10 km of the plant × Post	-0.0024	-0.0013
	(0.0039)	(0.0015)
Grid cell FE	Yes	Yes
Plant × Year FE	Yes	Yes
Geo. controls × Year FE	Yes	Yes
Grid cell trends	Yes	Yes
Cluster level	Plant	Plant
Observations	1,847,320	1,847,320
R-squared	0.5292	0.2419
Mean of dep. var.	0.0220	0.0174

Notes: The unit of observation is 1 km × 1 km grid cells. This table shows no differential effects of plant openings in areas where the mayor has greater promotion incentive (younger than 57). *Presence of fires* is an indicator for observing at least one fire. *$<$ 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A4: Biomass Power Plants: Higher-Order Time Trends

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var. =	Number of fires				Presence of fires			
$<$ 5 km of the plant $\times$ Post	-0.0108***	-0.0107***	-0.0107***	-0.0107***	-0.0059***	-0.0058***	-0.0058***	-0.0058**
	(0.0031)	(0.0032)	(0.0034)	(0.0035)	(0.0020)	(0.0021)	(0.0022)	(0.0023)
5-10 km of the plant $\times$ Post	-0.0050***	-0.0050***	-0.0050***	-0.0050**	-0.0022**	-0.0022*	-0.0022*	-0.0022*
	(0.0017)	(0.0018)	(0.0018)	(0.0019)	(0.0011)	(0.0011)	(0.0012)	(0.0012)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Plant $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls $\times$ Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Quadratic	Cubic	Quartic	Quintic	Quadratic	Cubic	Quartic	Quintic
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant
Observations	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576
R-squared	0.6199	0.6200	0.6199	0.6199	0.2973	0.2974	0.2974	0.2974
Mean of dep. var.	0.0220	0.0220	0.0220	0.0220	0.0174	0.0174	0.0174	0.0220

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This table shows the robustness of the baseline results to higher-order grid cell-specific time trends. *Presence of fires* is an indicator for observing at least one fire. *$<$ 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A5: Biomass Power Plants: Logarithm Transformation of Fire Number

	(1)	(2)
Dep. var.=	log(1+ Number of fires)	log(1+ Number of fires)
$<$ 5 km of the plant $\times$ Post	-0.0046***	-0.0062***
	(0.0011)	(0.0012)
5-10 km of the plant $\times$ Post	-0.0020***	-0.0028***
	(0.0007)	(0.0008)
Grid cell FE	Yes	Yes
Plant $\times$ Year FE	Yes	Yes
Geo. controls $\times$ Year FE	No	Yes
Grid cell trends	No	Yes
Cluster level	Plant	Plant
Observations	1,986,576	1,986,576
R-squared	0.1889	0.2893
Economic magnitude ($<$ 5 km)	-21.4%	-28.8%
Economic magnitude (5-10 km)	-9.3%	-13.0%

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This table shows the robustness of the baseline results by using log(1 + fire number) as the dependent variable. *$<$ 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Following Ciacci and Sviatschi (2021), the economic magnitudes reported in the last two rows are calculated as follows:

$$\frac{\partial \log(y)}{\partial x} = \frac{\partial \log(1+y)}{\partial x} \frac{\partial \log(y)}{\partial \log(1+y)} = \beta \frac{1+y}{y} \approx \hat{\beta} \frac{1+\bar{y}}{\bar{y}}$$

Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A6: Biomass Power Plants: Topography-Based Distance Measure

Dep. var. =	(1) #Fires	(2) #Fires>0	(3) #Fires	(4) #Fires>0	(5) #Fires	(6) #Fires>0	(7) #Fires	(8) #Fires>0
Least cost dist. to the plant	0.0008***	0.0003**	0.0011***	0.0005***	0.0007***	0.0003**	0.0010***	0.0005***
× Post	(0.0002)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0001)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant	Plant	Plant
Cost function	Walking	Walking	Walking	Walking	Driving	Driving	Driving	Driving
River cost	5	5	2	2	5	5	2	2
Observations	1,985,200	1,985,200	1,985,200	1,985,200	1,985,200	1,985,200	1,985,200	1,985,200
R-squared	0.5197	0.2378	0.5197	0.2378	0.5197	0.2378	0.5197	0.2378
Mean of dep. var.	0.0220	0.0174	0.0220	0.0174	0.0220	0.0174	0.0220	0.0174

Notes: The unit of observation is 1 km × 1 km grid cells. This table tests the robustness of the baseline results to using a topography-based distance, or the least cost distance, from a grid cell to the nearest plant. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Columns (1)-(4) model the slope-dependent cost using a function derived from walking data and columns (5)-(8) model the slope-dependent cost using a function derived from driving data. Columns (1), (2), (5), and (6) assume a unit cost of 5 for grid cells with rivers. Columns (3), (4), (7), and (8) assume a unit cost of 2 for grid cells with rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A7: Biomass Power Plants: Controlling for Selection on Crop Residues

Dep. var.=	(1) Number of fires	(2) Presence of fires	(3) Number of fires	(4) Presence of fires
< 5 km of the plant × Post	-0.0086***	-0.0052***	-0.0129***	-0.0069***
	(0.0023)	(0.0014)	(0.0032)	(0.0016)
5-10 km of the plant × Post	-0.0046**	-0.0017*	-0.0069***	-0.0029***
	(0.0018)	(0.0010)	(0.0021)	(0.0011)
Grid cell FE	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes
Crop residues in 2000 × Year FE	Yes	Yes	Yes	Yes
Geo. controls × Year FE	No	No	Yes	Yes
Grid cell trends	No	No	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant
Observations	1,704,480	1,704,480	1,704,480	1,704,480
R-squared	0.2776	0.1620	0.5329	0.2382
Mean of dep. var.	0.0223	0.0177	0.0223	0.0177

Notes: The unit of observation is 1 km × 1 km grid cells. This table addresses the concern that the plants may select into areas with a high amount of crop residues—which may correlate with fires—by interacting the initial amount of crop residues with year fixed effects. *Presence of fires* is an indicator for observing at least one fire. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A8: Biomass Power Plants: Dropping Areas Close to a Plant

	(1)	(2)	(3)	(4)
	Drop areas within 1 km of a plant		Drop areas within 2 km of a plant	
Dep. var.=	Number of fires	Presence of fire	Number of fires	Presence of fire
$<$ 5 km of the plant × Post	-0.0131***	-0.0067***	-0.0132***	-0.0067***
	(0.0028)	(0.0015)	(0.0028)	(0.0015)
5-10 km of the plant × Post	-0.0068***	-0.0027***	-0.0068***	-0.0027***
	(0.0018)	(0.0009)	(0.0018)	(0.0009)
Grid cell FE	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant
Observations	1,977,184	1,977,184	1,948,016	1,948,016
R-squared	0.5197	0.2376	0.5210	0.2379
Mean of dep. var.	0.0220	0.0174	0.0220	0.0174

Notes: The unit of observation is 1 km × 1 km grid cells. This table shows robustness to dropping areas close to a plant. *Presence of fires* is an indicator for observing at least one fire. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A9: Biomass Power Plants: Alternative Sample Selection

	(1)	(2)	(3)	(4)
	All plants & grid cells		All plants & non-overlapping grid cells	
Dep. var.=	Number of fires	Presence of fire	Number of fires	Presence of fire
$<$ 5 km of the plant × Post	-0.0115***	-0.0057***	-0.0127***	-0.0064***
	(0.0019)	(0.0011)	(0.0027)	(0.0015)
5-10 km of the plant × Post	-0.0048***	-0.0016**	-0.0058***	-0.0020**
	(0.0012)	(0.0006)	(0.0016)	(0.0009)
Grid cell FE	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant
Observations	2,617,312	2,617,312	2,158,608	2,158,608
R-squared	0.4920	0.2349	0.5099	0.2357
Number of plants	249	249	249	249
Mean of dep. var.	0.0217	0.0173	0.0221	0.0177

Notes: The unit of observation is 1 km × 1 km grid cells. This table shows the robustness of the baseline results to including all plants. *Presence of fires* is an indicator for observing at least one fire. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Columns (1) and (2) include all plants and grid cells within 15 km. This means that a grid cell may be linked to more than one plant in a certain year: it can be treated by a plant but can also serve as control for another plant, if these two plants have overlapping 15 km buffers. Columns (3) and (4) include all plants and non-overlapping grid cells within 15 km. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A10: Biomass Power Plants: Alternative Treatment Timing Assumptions

	(1)	(2)	(3)	(4)	(5)	(6)
	Exclude the year of opening		Exclude the years in [-1, 1]		”Expected” treatment value	
Dep. var. =	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0
$<$ 5 km of the plant × Post	-0.0121***	-0.0062***	-0.0120***	-0.0074***	-0.0115***	-0.0059***
	(0.0036)	(0.0017)	(0.0038)	(0.0025)	(0.0039)	(0.0017)
5-10 km of the plant × Post	-0.0075***	-0.0031***	-0.0087***	-0.0045***	-0.0072***	-0.0031***
	(0.0023)	(0.0011)	(0.0033)	(0.0017)	(0.0023)	(0.0011)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant
Observations	1,862,415	1,862,415	1,614,093	1,614,093	1,986,576	1,986,576
R-squared	0.4827	0.2432	0.4075	0.2619	0.5196	0.2378
Mean of dep. var.	0.0217	0.0173	0.0208	0.0168	0.0220	0.0174

Notes: The unit of observation is 1 km × 1 km grid cells. This table shows the robustness of the baseline results to excluding years around the plant opening or assigning an expected treatment value to the year of a plant opening. *$<$ 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Columns (1) and (2) exclude observations in the year of opening. Columns (3) and (4) exclude observations in a three-year window around the year of opening. Columns (5) and (6) replace the value of *Post* with the expected treatment value, which is calculated using equation (A5), in the year of opening. This is generated by assuming a uniform distribution of plant openings over the year and using the empirical distribution of crop residues over the year. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A11: Biomass Power Plants: Effect on Farmers' Income

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var. =	Household income calculated at 2010 constant price (unit: *yuan*)					
Opening	-2,385	626.1	-3,055			
	(2,182)	(2,457)	(3,648)			
Opening × Farmer	3,693***	4,759***	3,973***	4,234***	1,864*	
	(1,360)	(1,531)	(1,252)	(1,377)	(946.9)	
Farmer	-4,510***	-5,252***	-4,473***	-4,830***	-3,355***	-5,095**
	(1,313)	(1,262)	(1,346)	(1,319)	(1,226)	(2,039)
3 waves before opening × Farmer						-64.65
						(1,687)
2 waves before opening × Farmer						1,158
						(3,916)
0 wave after opening × Farmer						3,386*
						(1,823)
1 waves after opening × Farmer						1,022
						(1,371)
2 waves after opening × Farmer						4,454**
						(1,811)
3 waves after opening × Farmer						753.2
						(1,778)
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Survey wave FE	Yes	Yes	Yes	Yes	Yes	Yes
City × Survey wave FE	No	No	No	Yes	Yes	Yes
Farmer-specific wave trends	No	No	No	No	Yes	Yes
Cluster level	City	City	City	City	City	City
Observations	28,900	19,940	19,364	28,900	28,900	28,900
R-squared	0.117	0.078	0.074	0.134	0.130	0.138
Mean of farmers' income	30,676	29,097	29,142	30,676	30,675	30,676

Notes: The data structure is household-city-survey wave. This table examines effect of plant openings on farmers' income using 4 waves of the China Family Panel Studies (CFPS) dataset. The dependent variable denotes household income at 2010 constant price. Opening is a dummy variable equal to 1 if there are any plant openings in a city by a certain wave. Farmer is a dummy variable equal to 1 if the household lives in the rural areas. Column (2) compares cities with plant openings in wave 2 to other cities that never have any plant openings. Column (3) compares cities with plant openings in wave 3 to other cities that never have any plant openings. Standard errors in parentheses are clustered at the city level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A12: Biomass Power Plants: Changes in Small Firms Around the Plants

Dep. var. =	(1) ΔNumber of firms	(2) ΔNumber of employees	(3) ΔNumber of firms	(4) ΔNumber of employees
< 5 km of the plant	-0.0423	0.0696	-0.0411	0.0711
	(0.0358)	(0.0703)	(0.0366)	(0.0698)
5-10 km of the plant	-0.0165	0.0432	-0.0156	0.0438
	(0.0153)	(0.0282)	(0.0155)	(0.0284)
Plant FE	Yes	Yes	Yes	Yes
Geo. controls	No	No	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant
Observations	19,560	19,560	19,560	19,560
R-squared	0.0786	0.0177	0.0821	0.0194
Number of plants	30	30	30	30
Mean of dep. var. (in levels)	1.2152	59.1870	1.2152	59.1870

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This table reexamines the structural transformation mechanism using the 2004 and 2008 economic census data covering the universe of industrial firms including the small firms, by estimating a long-difference model. The dependent variable measures changes from 2004 to 2008 in a grid cell. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A13: Biomass Power Plants: Testing Farmer Adaption to the Plants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Crop share					Value of production	
Dep. var. =	ΔWheat	ΔRice	ΔMaize	ΔVegetables	ΔSoybean	ΔTotal	ΔPer hectare
< 5 km of the plant	-0.0100	0.0095	0.0250	-0.0227	-0.0021	0.5002	0.0000
	(0.0117)	(0.0134)	(0.0196)	(0.0217)	(0.0043)	(0.7059)	(0.0001)
5-10 km of the plant	-0.0059	0.0235*	-0.0050	0.0004	-0.0063	0.7539	0.0001
	(0.0088)	(0.0120)	(0.0106)	(0.0121)	(0.0042)	(0.5109)	(0.0001)
Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant	Plant	Plant
Observations	586	586	586	586	586	586	586
R-squared	0.5034	0.2429	0.5794	0.3253	0.4029	0.4631	0.4124
Number of plants	60	60	60	60	60	60	60
Mean of dep. var. (in levels)	0.1406	0.1913	0.2064	0.1333	0.0394	8.8563	0.0021

Notes: The unit of observation is 5 arcmin $\times$ 5 arcmin grid cells . This table tests farmer adaption to plant openings using the 2005 and 2010 gridded agricultural production datasets, by estimating a long-difference model. The dependent variable measures changes from 2005 to 2010 for the top 5 major crops' shares and the value of production, both total and per hectare, for all crops. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A14: Biomass Power Plants: Controlling for Potential Fire Bans under Transmission Towers

	(1)	(2)	(3)	(4)
	Excluding plants with differential reduction in fires for each half of the catchment area		Testing heterogeneity by lightning intensity	
Dep. var.=	Number of fires	Presence of fires	Number of fires	Presence of fires
$<$ 5 km of the plant × Post	-0.0129***	-0.0064***	-0.0116***	-0.0063***
	(0.0030)	(0.0016)	(0.0029)	(0.0018)
5-10 km of the plant × Post	-0.0071***	-0.0028***	-0.0068***	-0.0026**
	(0.0019)	(0.0010)	(0.0021)	(0.0010)
Lightning intensity × Post			0.0006	-0.0001
			(0.0020)	(0.0009)
Lightning intensity × $<$ 5 km of the plant × Post			-0.0020	-0.0023*
			(0.0019)	(0.0012)
Lightning intensity × 5-10 km of the plant × Post			0.0009	0.0001
			(0.0014)	(0.0008)
Observations	1,861,440	1,861,440	1,506,528	1,506,528
R-squared	0.5261	0.2384	0.3379	0.2390
Grid cell FE	Yes	Yes	Yes	Yes
Plant × Year FE	Yes	Yes	Yes	Yes
Geo. controls × Year FE	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes
Cluster level	Plant	Plant	Plant	Plant
Mean of dep. var.	0.0221	0.0174	0.0205	0.0164

Notes: The unit of observation is 1 km × 1 km grid cells. This table tests whether the baseline reduction in fires is driven by potential fire bans under transmission towers. *Presence of fires* is an indicator for observing at least one fire. *< 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. *Lightning intensity* is a time-invariant variable denoting the frequency of lightning strike. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A15: Biomass Power Plants: Other City-Level Environmental Regulations

	(1)	(2)	(3)	(4)	(5)
	Panel A: Dep. var. =Number of fires				
$<$ 5 km of the plant $\times$ Post	-0.0127***	-0.0126***	-0.0127***	-0.0126***	-0.0126***
	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0028)
5-10 km of the plant $\times$ Post	-0.0068***	-0.0067***	-0.0068***	-0.0068***	-0.0068***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0018)
Observations	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576
R-squared	0.5196	0.5196	0.5196	0.5196	0.5196
Mean of dep. var.	0.0220	0.0220	0.0220	0.0220	0.0220
	Panel B: Dep. var. =Presence of fires				
$<$ 5 km of the plant $\times$ Post	-0.0065***	-0.0064***	-0.0065***	-0.0065***	-0.0064***
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)
5-10 km of the plant $\times$ Post	-0.0027***	-0.0027***	-0.0027***	-0.0027***	-0.0027***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Observations	1,986,576	1,986,576	1,986,576	1,986,576	1,986,576
R-squared	0.2378	0.2378	0.2378	0.2378	0.2378
Mean of dep. var.	0.0174	0.0174	0.0174	0.0174	0.0174
Grid cell FE	Yes	Yes	Yes	Yes	Yes
Plant $\times$ Year FE	Yes	Yes	Yes	Yes	Yes
Geo. controls $\times$ Year FE	Yes	Yes	Yes	Yes	Yes
Grid cell trends	Yes	Yes	Yes	Yes	Yes
Distance to TCZ city center $\times$ Year FE	Yes	No	No	No	No
Distance to KCAPC city center $\times$ Year FE	No	Yes	No	No	No
Distance to TCZ city center $\times$ Post06	No	No	Yes	No	No
Distance to Olympic host & neighbor city center $\times$ Year07-08	No	No	No	Yes	No
Distance to KRAPC city center $\times$ Post12	No	No	No	No	Yes
Cluster level	Plant	Plant	Plant	Plant	Plant

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This table tests the robustness of the baseline results to controlling for other city-level environmental regulations. *Presence of fires* is an indicator for observing at least one fire. *$<$ 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. The regulations include: (1) the Two Control Zones (TCZ) policy initiated in 1998 to control for acid rain and SO_2 pollution, with 175 cities as target cities; (2) the Key Cities for Air Pollution Control (KCAPC) policy initiated in 2001 to control for acid rain, SO_2, and particulate matter, with 113 cities as target cities; (3) the TCZ policy after 2005, which linked the enforcement to the performance evaluation system of local leaders; (4) the regulation during the Beijing 2008 Olympic Games, which imposed radical air quality requirement in all host cities and some neighbor cities from 2007 to 2008; (5) the Key Regions for Air Pollution Control (KRAPC) policy initiated in 2012 to target $PM_{2.5}$ and secondary pollution formed by chemical or photochemical reactions, with 117 cities as target cities. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A16: Biomass Power Plants: Effect on Air Pollution

	(1)	(2)
Dep. var. =	$PM_{2.5}$	$PM_{2.5}$
Post	-1.0951*	-1.0265*
	(0.5798)	(0.5660)
Plant FE	Yes	Yes
Year FE	Yes	Yes
Plant time trends	Yes	Yes
Weather controls	No	Yes
Cluster level	Plant	Plant
Observations	3,040	3,004
R-squared	0.9999	0.9999
Mean of dep. var. ($\mu g/m^3$)	54.9	54.9

Notes: The unit of observation is plants. This table tests the effect of plant openings on air pollution. The dependent variable is a satellite-based $PM_{2.5}$ concentration measure averaged across the 10 km buffer of a plant. *Post* is an indicator for years after the plant opening. Weather controls include precipitation, temperature, wind speed, wind direction, and dew point temperature. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A17: Biomass Power Plants: Effect on Water Quality

	(1)	(2)
Dep. var. =	Water quality index	
Presence of plants within 20 km	0.0960	0.1005
	(0.0950)	(0.1018)
Station FE	Yes	Yes
Year FE	Yes	Yes
Station time trends	Yes	Yes
Weather controls	No	Yes
Cluster level	Station	Station
Observations	3,234	3,106
R-squared	0.9357	0.9388
Mean of dep. var.	3.4938	3.5106

Notes: The unit of observation is monitoring stations. This table tests the effect of nearby plant openings on water quality. *Presence of plants within 20 km* is an indicator for any plant openings within 20 km of a water quality monitoring station. *Water quality index* is an integer ranging from 1 to 6, with lower values denoting higher water quality. Weather controls include precipitation, temperature, wind speed, wind direction, and dew point temperature. Standard errors in parentheses are clustered at the water quality monitoring station level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A18: Biomass Power Plants: Effect on Individual Water Pollutants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var. =	DO	CODMn	COD	NH	Petroleum	Phenol	Hg	Pb
Presence of plants within 20 km	-0.0737	0.2185	0.1509	-0.0796	-0.0171	-0.0019	0.0031	0.0007
	(0.1803)	(0.9638)	(0.9009)	(0.2801)	(0.0278)	(0.0018)	(0.0048)	(0.0009)
Station FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Station trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Station	Station	Station	Station	Station	Station	Station	Station
Observations	3,091	3,057	3,024	2,958	2,876	2,845	2,795	2,841
R-squared	0.8975	0.9176	0.9057	0.9445	0.8055	0.4899	0.6095	0.8669
Mean of dep. var.	7.2390	6.8604	5.9097	2.2543	0.1115	0.0054	0.0330	0.0058

Notes: The unit of observation is water quality monitoring stations. This table tests the effect of nearby plant openings on individual water pollutants. DO stands for dissolved oxygen, CODMn stands for permanganate index, COD stands for chemical oxygen demand, NH stands for ammonia nitrogen, Hg stands for mercury, and Pb stands for lead. *Presence of plants within 20 km* is an indicator for any plant openings within 20 km of a water quality monitoring station. Weather controls include precipitation, temperature, wind speed, wind direction, and dew point temperature. Standard errors in parentheses are clustered at the water quality monitoring station level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A19: Biomass Power Plants: Overlapping Catchment Areas

	(1)	(2)
Dep. var.=	Number of fires	Presence of fire
$<$ 5 km of the plant $\times$ Post	-0.0116***	-0.0056***
	(0.0024)	(0.0014)
5-10 km of the plant $\times$ Post	-0.0059***	-0.0020**
	(0.0016)	(0.0008)
Overlap	0.0003	0.0001
	(0.0005)	(0.0003)
Overlap $\times$ Post	0.0003	0.0003
	(0.0006)	(0.0004)
Overlap $\times$ $<$ 5 km of the plant	-0.0001	-0.0003
	(0.0005)	(0.0004)
Overlap $\times$ 5-10 km of the plant	-0.0006	-0.0001
	(0.0005)	(0.0003)
Overlap $\times$ $<$ 5 km of the plant $\times$ Post	0.0009*	0.0005
	(0.0005)	(0.0004)
Overlap $\times$ 5-10 km of the plant $\times$ Post	0.0002	0.0000
	(0.0005)	(0.0004)
Observations	2,617,312	2,617,312
R-squared	0.4933	0.2357
Grid cell FE	Yes	Yes
Plant $\times$ Year FE	Yes	Yes
Geo. controls $\times$ Year FE	Yes	Yes
Grid cell trends	Yes	Yes
Cluster level	Plant	Plant
Mean of dep. var.	0.0217	0.0173

Notes: The unit of observation is 1 km $\times$ 1 km grid cells . This table tests whether areas falling into overlapping catchment areas of more than one plant are different from other areas. *Presence of fires* is an indicator for observing at least one fire. *$<$ 5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. *Overlap* is an indicator equal to 1 if a grid cell around a certain plant also falls in the catchment areas of other plants that are active. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A20: Biomass Power Plants: Matched Sample

	(1)	(2)
Dep. var. =	Number of fires	Presence of fires
$<$ 5 km of the plant $\times$ Post	-0.0092***	-0.0053***
	(0.0029)	(0.0013)
5-10 km of the plant $\times$ Post	-0.0052***	-0.0021**
	(0.0016)	(0.0009)
Grid cell FE	Yes	Yes
Plant $\times$ Year FE	Yes	Yes
Geo. controls $\times$ Year FE	Yes	Yes
Grid cell trends	Yes	Yes
Cluster level	Plant	Plant
Observations	1,505,872	1,505,872
R-squared	0.5658	0.2403
Mean of dep. var.	0.0183	0.0141

Notes: The unit of observation is 1 km $\times$ 1 km grid cells . This table repeats the baseline DID estimation for biomass power plants using a matched sample such that the distribution of the top 5 crop shares is similar to that around the airports. *Presence of fires* is an indicator for observing at least one fire. $<$ *5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A21: Biomass Power Plants: Trimmed Sample

	(1)	(2)
Dep. var. =	Number of fires	Presence of fires
$<$ 5 km of the plant $\times$ Post	-0.0130***	-0.0071***
	(0.0026)	(0.0017)
5-10 km of the plant $\times$ Post	-0.0070***	-0.0027**
	(0.0020)	(0.0011)
Observations	1,546,960	1,546,960
R-squared	0.3147	0.2314
Grid cell FE	Yes	Yes
Plant $\times$ Year FE	Yes	Yes
Geo. controls $\times$ Year FE	Yes	Yes
Grid cell trends	Yes	Yes
Cluster level	Plant	Plant
Mean of dep. var.	0.0220	0.0181

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. This table repeats the baseline DID estimation focusing on a sample of plants that are placed at least 30 km away from the airports. *Presence of fires* is an indicator for observing at least one fire. $<$ *5 km of the plant* and *5-10 km of the plant* are indicators for grid cells within 5 km and 5-10 km of the plant, respectively. *Post* is an indicator for years after the plant opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers, and number of rivers. Standard errors in parentheses are clustered at the plant level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A22: No Burning Zones: Balance Checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
DV =	Ruggedness				Slope				Elevation			
Inside	0.0034	0.0066	0.0680	0.0843	-0.0099	-0.0101	-0.0044	-0.0109	0.5534	-0.2947	-1.5618	-1.4489
	(0.6606)	(0.8218)	(0.9999)	(0.6605)	(0.0214)	(0.0247)	(0.0292)	(0.0212)	(0.9036)	(1.3408)	(1.8152)	(1.6029)
MDV.	119.8	119.9	120.4	119.9	3.449	3.454	3.467	3.450	817.1	818.7	820.5	819.2
BW (km)	2	4	6	1.987	2	4	6	2.112	2	4	6	5.121
DV =	Distance to level 3+ rivers				Distance to level 4 rivers				Distance to level 5 rivers			
Inside	0.0024	-0.0055	-0.0027	-0.0061	0.0080	0.0045	0.0123	0.0043	-0.0003	0.0078	0.0198	0.0145
	(0.0194)	(0.0111)	(0.0108)	(0.0108)	(0.0188)	(0.0108)	(0.0101)	(0.0094)	(0.0179)	(0.0112)	(0.0152)	(0.0103)
MDV	62.10	62.04	62.01	62.03	68.49	68.44	68.37	68.40	19.71	19.69	19.69	19.68
BW (km)	2	4	6	5.094	2	4	6	4.854	2	4	6	4.371
DV =	# Level 3+ rivers				# Level 4 rivers				# Level 5 rivers			
Inside	0.0016	-0.0004	-0.0004	-0.0011	0.0013	0.0006	-0.0005	0.0004	-0.0007	-0.0008	-0.0004	-0.0007
	(0.0015)	(0.0014)	(0.0017)	(0.0013)	(0.0014)	(0.0011)	(0.0013)	(0.0011)	(0.0019)	(0.0021)	(0.0022)	(0.0019)
MDV	0.0225	0.0229	0.0238	0.0227	0.0137	0.0137	0.0135	0.0136	0.0277	0.0280	0.0276	0.0281
BW (km)	2	4	6	3.337	2	4	6	4.070	2	4	6	2.546
DV =	Distance to expressway				Distance to railway				Distance to national road			
Inside	0.0112	0.0014	0.0002	0.0040	0.0015	-0.0020	-0.0044	-0.0100	-0.0162	-0.0036	-0.0036	-0.0137
	(0.0206)	(0.0114)	(0.0100)	(0.0113)	(0.0195)	(0.0107)	(0.0127)	(0.0152)	(0.0186)	(0.0113)	(0.0139)	(0.0129)
MDV	227.8	227.9	227.9	228	84.51	84.67	84.75	84.78	17.19	17.26	17.39	17.32
BW (km)	2	4	6	6.968	2	4	6	6.916	2	4	6	5.068
DV =	Distance to city center				Distance to city border				Share of cropland in 1998			
Inside	-0.2365	-0.0639	0.0886	0.1027	-0.0316*	-0.0162	0.0047	-0.0007	0.0025	0.0019	0.0026	0.0015
	(0.1657)	(0.1094)	(0.1466)	(0.1269)	(0.0187)	(0.0110)	(0.0138)	(0.0119)	(0.0029)	(0.0031)	(0.0035)	(0.0031)
MDV.	64.22	64.35	64.69	64.58	26.69	26.64	26.57	26.60	0.456	0.456	0.455	0.456
BW (km)	2	4	6	5.675	2	4	6	5.411	2	4	6	3.418
DV =	Share of settlement land in 1998				Nighttime light intensity in 1998				# Industrial firms in 1998			
Inside	-0.0019*	-0.0015	-0.0008	-0.0013	-0.0357	-0.0798	-0.1038	-0.0753*	-0.0039	-0.0194	-0.0061	-0.0102
	(0.0010)	(0.0011)	(0.0013)	(0.0009)	(0.0456)	(0.0511)	(0.0802)	(0.0433)	(0.0160)	(0.0141)	(0.0126)	(0.0158)
MDV	0.0299	0.0303	0.0302	0.0299	5.058	5.027	4.968	5.047	0.0894	0.0917	0.0914	0.0877
BW (km)	2	4	6	2.500	2	4	6	3.458	2	4	6	2.474
DV =	Industrial real capital in 1998				Industrial real output in 1998				# Industrial employees in 1998			
Inside	-0.5598	-75.2837	125.3249	106.0889	-384.4716	-793.8926	-151.0150	-243.1472	-0.8705	-3.1505	0.5069	-7.7295
	(440.9480)	(287.0483)	(242.9399)	(270.7347)	(752.4502)	(706.9266)	(668.0807)	(687.9686)	(5.1991)	(3.7758)	(3.2658)	(6.3049)
MDV	1734	1773	1783	1775	3404	3463	3450	3453	24.51	25.08	25.15	25.26
BW (km)	2	4	6	4.784	2	4	6	5.328	2	4	6	2.861

Notes: The unit of observation is 1 km $\times$ 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table provides balance checks on preexisting variables on each side of the no burning zone boundaries, using the baseline RD specification. *Inside* is an indicator for areas within the no burning zones. Each coefficient comes from an independent regression. All regressions control for airport $\times$ boundary segment fixed effects. The optimal bandwidth in columns (4), (8), and (12) is selected following the procedure proposed by Imbens and Kalyanaraman (2012). Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A23: No Burning Zones: Additional Donut Hole Specifications

	(1)	(2)	(3)	(4)
	Drop areas within 2 km of the boundary		Drop areas within 3 km of the boundary	
Dep. var. =	Number of fires	Presence of fires	Number of fires	Presence of fires
Inside	-0.0009	0.0002	0.0002	0.0001
	(0.0012)	(0.0006)	(0.0011)	(0.0006)
RD polynomial	Linear	Linear	Linear	Linear
Airport × Segment FE	Yes	Yes	Yes	Yes
Cluster level	Airport	Airport	Airport	Airport
Observations	1,624,555	1,716,501	1,400,735	1,722,591
R-squared	0.0231	0.0285	0.0260	0.0281
Mean of dep. var.	0.0105	0.0090	0.0106	0.0090
Bandwidth (km)	3.5543	3.7561	3.0655	3.7713

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table tests the robustness of the baseline results to excluding areas close (within 2 or 3 km) to the no burning zone boundaries. *Presence of fires* is an indicator for observing at least one fire. *Inside* is an indicator for areas within the no burning zones. The optimal bandwidth is selected following the procedure proposed by Imbens and Kalyanaraman (2012). Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A24: No Burning Zones: Difference-in-Discontinuities Analysis Using Alternative Bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var. =	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0
Inside × Post	0.0014	0.0006	0.0017	0.0005	0.0000	-0.0002	-0.0019	-0.0006
	(0.0013)	(0.0009)	(0.0013)	(0.0007)	(0.0007)	(0.0006)	(0.0017)	(0.0007)
RD polynomial	Linear	Linear	Linear	Linear	Linear	Linear	Linear	Linear
RD polynomial × Post	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Airport × Segment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Airport × Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport
Observations	555,504	555,504	832,112	832,112	1,386,992	1,386,992	1,664,992	1,664,992
R-squared	0.1774	0.1453	0.1698	0.1427	0.1653	0.1424	0.1888	0.1430
Mean of dep. var.	0.0067	0.0060	0.0068	0.0060	0.0068	0.0060	0.0069	0.0060
Bandwidth (km)	2	2	3	3	5	5	6	6

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table checks the robustness of the difference-in-discontinuities design to alternative bandwidths. *Inside* is an indicator for areas within the no burning zones. *Post* is an indicator for years after the airport opening. Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A25: No Burning Zones: Two-Dimensional RD Design Using Alternative Bandwidths

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var. =	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0	#Fires	#Fires>0
Inside	0.0003	0.0001	-0.0000	0.0002	0.0004	0.0003	0.0007	0.0005
	(0.0006)	(0.0003)	(0.0006)	(0.0003)	(0.0006)	(0.0004)	(0.0006)	(0.0004)
RD polynomial in lon. and lat.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Airport × Segment FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster level	Airport	Airport	Airport	Airport	Airport	Airport	Airport	Airport
Observations	914,902	914,902	1,372,493	1,372,493	2,285,908	2,285,908	2,743,315	2,743,315
R-squared	0.0339	0.0373	0.0274	0.0335	0.0240	0.0296	0.0226	0.0286
Mean of dep. var.	0.0115	0.0097	0.0113	0.0094	0.0109	0.0092	0.0109	0.0092
Bandwidth (km)	2	2	3	3	5	5	6	6

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table checks the robustness of the two-dimensional RD design to alternative bandwidths. *Inside* is an indicator for areas within the no burning zones. Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A26: No Burning Zones: Heterogeneity by Mayors' Promotion Incentives

	(1)	(2)	(3)	(4)
	High incentive	Low incentive	High incentive	Low incentive
Dep. var. =	Number of fires		Presence of fires	
	Panel A: Bandwidth = 2 km			
Inside	-0.0001	0.0037*	-0.0009	0.0014
	(0.0009)	(0.0019)	(0.0006)	(0.0013)
Observations	662,844	99,589	662,844	99,589
Mean of dep. var.	0.0130	0.0119	0.0109	0.0098
	Panel B: Bandwidth = 4 km			
Inside	0.0003	0.0006	-0.0001	0.0004
	(0.0009)	(0.0014)	(0.0005)	(0.0010)
Observations	1,325,024	198,551	1,325,024	198,551
Mean of dep. var.	0.0124	0.0113	0.0104	0.0093
	Panel C: Bandwidth = 6 km			
Inside	-0.0001	-0.0004	-0.0000	0.0004
	(0.0008)	(0.0018)	(0.0005)	(0.0010)
Observations	1,986,902	297,233	1,986,902	297,233
Mean of dep. var.	0.0123	0.0111	0.0104	0.0091
	Panel D: Optimal bandwidth (km)			
Inside	0.0007	0.0001	-0.0002	0.0003
	(0.0011)	(0.0017)	(0.0005)	(0.0011)
Observations	820,569	237,684	1,171,523	234,767
Mean of dep. var.	0.0129	0.0110	0.0105	0.0091
Bandwidth	2.4760	4.7933	3.5332	4.7316
RD polynomial	Linear	Linear	Linear	Linear
Airport × Segment FE	Yes	Yes	Yes	Yes
Cluster level	Airport	Airport	Airport	Airport

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table checks the heterogeneity of the baseline results by mayors' promotion incentives. The subsamples are created by whether the mayor's age is younger than 57 (*High incentive*) or not (*Low incentive*), as local leaders older than 57 have very low probability of being promoted. *Presence of fires* is an indicator for observing at least one fire. *Inside* is an indicator for areas within the no burning zones. The optimal bandwidth in Panel D is selected using the procedure proposed by Imbens and Kalyanaraman (2012). Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A27: No Burning Zones: Heterogeneity by Timing of Burning

	(1)	(2)	(3)	(4)
	Daytime	Nighttime	Daytime	Nighttime
Dep. var. =	Number of fires		Presence of fires	
	Panel A: Bandwidth = 2 km			
Inside	0.0004	0.0001	-0.0004	0.0001
	(0.0006)	(0.0002)	(0.0004)	(0.0001)
Observations	914,902	914,902	914,902	914,902
Mean of dep. var.	0.0105	0.0010	0.0089	0.0009
	Panel B: Bandwidth = 4 km			
Inside	-0.0000	0.0001	-0.0002	0.0001
	(0.0006)	(0.0002)	(0.0004)	(0.0001)
Observations	1,829,849	1,829,849	1,829,849	1,829,849
Mean of dep. var.	0.0100	0.0010	0.0085	0.0008
	Panel C: Bandwidth = 6 km			
Inside	-0.0004	0.0000	-0.0002	0.0001
	(0.0007)	(0.0002)	(0.0004)	(0.0001)
Observations	2,743,315	2,743,315	2,743,315	2,743,315
Mean of dep. var.	0.0100	0.0010	0.0085	0.0008
	Panel D: Optimal bandwidth (km)			
Inside	0.0004	0.0001	-0.0002	0.0001
	(0.0006)	(0.0002)	(0.0004)	(0.0001)
Observations	1,481,591	1,226,262	1,957,101	1,707,071
Mean of dep. var.	0.0102	0.0010	0.0085	0.0009
Bandwidth	3.2393	2.6811	4.2806	3.7334
RD polynomial	Linear	Linear	Linear	Linear
Airport × Segment FE	Yes	Yes	Yes	Yes
Cluster level	Airport	Airport	Airport	Airport

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table conducts RD estimation around the no burning zone boundaries separately for fires observed in daytime and nighttime. *Presence of fires* is an indicator for observing at least one fire. *Inside* is an indicator for areas within the no burning zones. The optimal bandwidth in Panel D is selected using the procedure proposed by Imbens and Kalyanaraman (2012). Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A28: No Burning Zones: Matched Sample

	(1)	(2)
Dep. var. =	Number of fires	Presence of fires
Inside	0.0018	0.0001
	(0.0012)	(0.0006)
RD polynomial	Yes	Yes
Airport × Segment FE	Yes	Yes
Cluster level	Airport	Airport
Observations	730,386	1,028,066
R-squared	0.0286	0.0334
Mean of dep. var.	0.0156	0.0125
Bandwidth (km)	2.6355	3.7071

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table repeats the baseline RD estimation around the no burning zone boundaries using a matched sample such that the distribution of the top 5 crop shares is similar to that around the plants. *Presence of fires* is an indicator for observing at least one fire. *Inside* is an indicator for areas within the no burning zones. The optimal bandwidth is selected following the procedure proposed by Imbens and Kalyanaraman (2012). Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A29: No Burning Zones: Trimmed Sample

	(1)	(2)
Dep. var. =	Number of fires	Presence of fires
Inside	0.0005	0.0002
	(0.0008)	(0.0004)
Observations	1,312,975	2,078,715
R-squared	0.0223	0.0256
RD polynomial	Yes	Yes
Airport × Segment FE	Yes	Yes
Cluster level	Airport	Airport
Mean of dep. var.	0.0096	0.0080
Bandwidth (km)	3.6509	5.7769

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table repeats the baseline RD estimation focusing on a sample of airports that are placed at least 30 km away from the plants. *Presence of fires* is an indicator for observing at least one fire. *Inside* is an indicator for areas within the no burning zones. The optimal bandwidth is selected following the procedure proposed by Imbens and Kalyanaraman (2012). Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Table A30: No Burning Zones: Difference-in-Differences Estimation

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.=	Number of fires			Presence of fires		
	Panel A: Sample is within 10 km of the zone boundaries					
Inside x Post	0.0043	0.0038	0.0018	0.0008	0.0004	0.0003
	(0.0031)	(0.0031)	(0.0017)	(0.0006)	(0.0006)	(0.0006)
Observations	2,774,720	2,774,720	2,774,720	2,774,720	2,774,720	2,774,720
R-squared	0.3157	0.3158	0.6363	0.1433	0.1436	0.2181
Mean of dep. var.	0.0071	0.0071	0.0071	0.0059	0.0059	0.0059
	Panel B: Sample is within 15 km of the zone boundaries					
Inside x Post	0.0043	0.0038	0.0017	0.0012	0.0007	0.0002
	(0.0028)	(0.0028)	(0.0015)	(0.0007)	(0.0007)	(0.0006)
Observations	3,987,408	3,987,408	3,987,408	3,987,408	3,987,408	3,987,408
R-squared	0.3023	0.3024	0.5884	0.1419	0.1421	0.2154
Mean of dep. var.	0.0068	0.0068	0.0068	0.0058	0.0058	0.0058
Grid FE	Yes	Yes	Yes	Yes	Yes	Yes
Airport x Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Geo. controls x Year FE	No	Yes	Yes	No	Yes	Yes
Grid cell trends	No	No	Yes	No	No	Yes
Cluster level	Airport	Airport	Airport	Airport	Airport	Airport

Notes: The unit of observation is 1 km × 1 km grid cells. The no burning zones are areas within 15 km of an airport. This table estimates a difference-in-differences model using areas within 10 or 15 km of of the zone boundaries. *Inside* is an indicator equal to 1 for areas within the zones. *Post* is an indicator equal to 1 for years after an airport opening. Geographic controls include longitude, latitude, slope, ruggedness, elevation, distance to rivers and number of rivers. *Presence of Fires* is an indicator equal to 1 if at least one fire is observed. Standard errors in parentheses are clustered at the airport level. * denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.

Chapter 3

Go with the Politician

Yongwei Nian
Chunyang Wang

Abstract

Chinese local leaders are frequently moved across prefectures. By combining local leader rotation data and comprehensive firm land parcel purchase data across prefectures from 2006 to 2016, this paper examines how firm-politician connections affect resource allocation and finds that a firm headquartered in a leader's previous work prefecture purchases three times more land parcels in that leader's new governing prefecture than the prefecture-year mean, at half the unit prices. Identification is from within firm-year variation in various prefectures through exogenous politician rotation. Land usage efficiency is lower for these follower firms land parcels. Land allocation distortion is also economically sizable.

JEL classification: D73, O10, P26
Keywords: Official rotation; Land Market; Corruption; China

China has been a growth superstar for the past four decades and transformed itself from a poor to middle-income country. At the same time, China ranked 11^{th} in the *Economist*'s crony capitalism index in 2016.[1] According to its president in 2017, reducing distortion from cronyism to establish close but clear business-government relations is a major task facing the Chinese government.[2] From Chinese leaders' perspective, cronyism entails unfair economic inequality, where politically connected businessmen obtain wealth through favoritism, and such inequality is often viewed as the main factor contributing to social unrest and threatening the communist party's rule. Moreover, China is experiencing an economic slowdown. Reducing the distortion from cronyism will lift economic productivity and therefore growth (Allen, Qian, and Qian, 2005; Hsieh and Klenow, 2009).

Studying cronyism and its impact on resource allocation is vital, but the literature on connections between firms and politicians, pioneered by Faccio (2006),[3] has endogeneity issues such as mutual choice; that is, a firm choosing to have a politician seated on its board and a politician choosing to sit on its board are based on the two sides' characteristics, according to her definition of political connections (Schoenherr, 2019). Faccio (2006) notes that her original work demonstrates correlation but likely not causation because, obviously, political connections emerge endogenously. For example, a firm might obtain more government resources from a politician because they share the same development ideology. Moreover, the omitted variables problem might be more severe in developing economies such as China, where for example, it is common for firms to bribe to form connections, which is possibly a substitute for a politician sitting on a firm's board as a method of connecting, making such identification even more difficult.

In this paper, we target the above endogeneity issue using a unique dataset where a firm purchases land parcels across different prefectures and a rotation policy where a local government leader is moved from one prefecture to another. We define a firm's political connections with a particular politician from a social connection perspective, and our

[1] Please see https://www.economist.com/international/2016/05/07/the-party-winds-down.

[2] Please see the Chinese president's target to establish a "new business-government relation" guideline at http://theory.people.com.cn/n1/2017/0906/c413700-29519441.html

[3] Subsequent research studies firms' political connections in firm financing (Khwaja and Mian, 2005; Li et al., 2008), government contracts (Goldman, Rocholl and So, 2013; Tahoun, 2014; Brogaard, Denes and Duchin, 2015), government bailouts (Faccio, Masulis and McConnell, 2006) and other contexts.

independent variable, political connections for a firm in a prefecture, equals one if the prefecture in which a firm's headquarters is located is the same as a local leader's previous work prefecture. Thus, a firm will have different political connection values for different prefectures in each year. The value varies over time as a politician is rotated from one prefecture to another. Such a definition does not necessarily indicate a real political connection between a firm and a politician, but the probability of that firm having a connection with that official is high for the following reasons. First, firm management teams and the politician are in the same prefecture. According to social connection studies by Zipf (1949), Verbrugge (1983), Marmaros and Sacerdote (2006), and recently Bailey et al. (2018), geographic proximity plays the most important role in friendship formation and social connections. Similarly, Jia, Kudamatsu, and Seim (2015), using various anecdotal evidence from China, measure political connections between two officials by whether they have previously worked in the same prefecture. Second, both a publicly listed firm and a politician in its headquarters prefecture share similar objectives. Local economic growth is the main determinant of a local leader's future career path (Blanchard and Shleifer, 2001; Li and Zhou, 2005). Public companies are among the largest firms in the local economy. They have strong incentives to interact with each other. For example, firms can invest heavily in politically key years for the local politician to be promoted since local GDP growth is a primary promotion indicator.

Land is our investigation target for the following reasons. First, it is important to both firms and local governments. Land is a key factor in production, especially at present when Chinese land or related residential real estate prices in major prefectures are at world highs (see Figure 1)[4]. In our dataset, the majority of listed firms purchase at least one land parcel, and these land buyers on average purchase 20 land parcels. Therefore, land purchases are frequent at both the extensive and intensive margins. For local governments, land is the sole production factor under their control, following two major reforms initiated by the central government, including financial reform, which abolished local government's control over state banks' local branches, and household registration, or *hukou*, reform which increased workers' mobility, as evidenced by the movement of large numbers of migrant workers since the late 1990s. Moreover, land has

[4] In 2021, the price of an apartment in Beijing, Shanghai, and Shenzhen reached approximately 10,000 US dollars per square meter, comparable to that in New York and San Francisco (www.numbeo.com).

become the most important local government revenue source since 1994, as seen in Figure 2. Revenue from selling land amounts to approximately 60% of local government revenue in recent decades. Second, the Greek definition of cronyism largely has a "long term" connotation. Compared to other industrial policies, such as cash subsidies, which usually take the form of a lump sum payment, we view land as much more suitable than other factors. Third, China initiated the largest anti-corruption campaign in the history of communist rule in 2013. Corruption cases are mostly closely related to land deals. Indeed, "land" was a commonly repeated word in corruption news reports. For example, the party secretary of Kunming Prefecture from 2007 to 2011 was well connected to land deals and was the prefecture leader of Suqian Prefecture in Jiangsu Province before 2007. After he moved to govern Kunming Prefecture, many real estate and construction firms from Suqian invested heavily in Kunming.[5]

By employing data from the turnover of over one thousand prefecture leaders and data from over one million land sale transactions from 2006 to 2016, we find that politically connected firms purchase three times more land parcels in a leader's new governing prefecture than the prefecture-year mean. This effect significantly increases for firms with higher connection intensity, where connection intensity is defined as the number of years the current leader governed his or her previous work region or how recently the current leader left his or her previously governed region. These land parcels had 51% lower unit prices, using the average unit prices of similar nearby land parcels as a benchmark, and these land parcels show lower subsequent land usage efficiency as measured by nighttime light intensity. We also explore dynamic aspects of connections and show that the "go with the politician" effect is quite persistent over a politician's tenure.

We control for a series of flexible fixed effects, including prefecture-by-year fixed effects, firm-by-year fixed effects, and origin province-by-destination province fixed effects.[6] If a local leader moves from prefecture A to B, then A is called the origin prefecture, while B is called the destination prefecture. Despite controlling for highly flexible fixed effects, we also examine firms in a prefecture that would shortly become connected, but none of the pre-connection trends are significant for our various specifications, and the "go with the politician" effect

[5] Please see https://news.sina.cn/gn/2015-03-24/detail-iavxeafs2097176.d.html.

[6] A province is composed of twelve prefectures on average.

instantly vanishes after the leader is again rotated.

The observation of land purchase flow following a politician's move does not necessarily indicate corruption. This could be explained, in economic theory, from an information asymmetry perspective where the politician grants parcels to these firms through familiarity. In this case, under information asymmetry theory (Stiglitz and Weiss, 1981), we should expect these firms to be better performers and land parcels to be used more efficiently. However, we observe the opposite: by using light intensity as a measure of land usage efficiency, we find that land parcels purchased by connected firms have lower light intensity growth after the purchase than parcels purchased by unconnected firms, suggesting that information asymmetry channels do not apply in this instance. Quantitatively, as shown later in the calculation section, a land parcel purchased by a connected firm had 1.2% lower economic output than it would have if it had been purchased by an unconnected firm. The Chinese central government requires firms to start construction immediately after purchase, with delays no longer than two years, or otherwise face the risk of being taken over by the local government. Moreover, there is a year-limit regulation imposed by the central government on firms selling constructed apartments to prevent hoarding beyond the year limit to gain from future price appreciation. The central government's land market regulator, the Ministry of Land and Resources, issued a report in 2010 stating that there were 2,208 cases where land parcels had not been developed by the central government's deadline of a two-year development window.[7] Because the local government is the main year-limit regulation implementer, a close relationship with the local government is beneficial for firms to evade regulation by the central government.

Our paper contributes to the literature on political connections, starting from works by Fisman (2001) and Faccio (2006) and the general literature on social connections (Bailey et al., 2018). We use geographic proximity as an indicator of business-politician connections to explore their impact on business resource allocation. To the best of our knowledge, this has been less explored because of data limitations since a good identification of such business-politician relations requires the rotation of politicians and geographic variation within firms. Schoenherr (2019) studies the effect of political connections on resource allocation using a

[7] This website lists the detailed locations and buyers of land parcels violating China's land regulation: http://www.yicai.com/show_topic/388496/

similar methodology. He uses the replacement of state-owned enterprises' (SOEs') CEOs by the newly elected Korean president as a within-private-firm variation across these SOEs because a private firm obtains contracts from a number of SOEs. In contrast to his work, our data have staggered yearly variation as local politicians are moved across different years. Shi et al. (2020) also study official rotation in China. Using aggregate data, they find that investment increased 3% after leader rotation. However, our work employs within-firm analysis, and the confounding factors are reduced significantly.

In terms of the research theme, the work by Chen and Kung (2019) is closest to ours. Using the same land dataset, they find that "princeling" firms obtain land parcels at discount prices. Our paper differs from theirs by mainly studying the geographical distribution of a firm's activity and allocative efficiency. Moreover, our geographic political connection measure has a general application, even in other countries.

The remainder of the paper proceeds as follows. Section I introduces China's institutional background and land market. Section II describes the data and methodology. Section III presents the results. Section IV reports the robustness checks. Section V shows further results, including heterogeneity analysis, anti-corruption campaigns as a natural experiment, locally headquartered firms, and hometown firm analysis. Section VI concludes the paper.

I. Institutional Background

A. *Importance of Local Government Leaders, Including in Allocating Land*

We are studying politicians at the local level in this paper. There are several reasons for us to focus on local governments. Considering China's size, it is obvious that the central government's connection with thousands of publicly listed firms is hardly believed to be extensive, simply due to leaders' limited attention. Local government is better positioned to study politician-business relations. Local government is important in many developing countries, such as China, Brazil, and India. In China, local government typically absorbs up to 70% of fiscal expenditure (Xu, 2011). Blanchard and Shleifer (2001) and Xu (2011) pioneer "tournament" theory, which explains local government's key role in China's growth, finding that competition among local government officials for career promotion has driven China's superstar growth. In particular, China has an authoritarian political structure, and local government officials are held to a high degree of accountability. Part of the reason is that the central government uses its centralized political power to promote local politicians, and it is easy for the central government to set GDP growth as its main target.[8] Therefore, local leaders work hard to fulfill GDP goals, as formally tested using provincial GDP and local leader promotion data in Li and Zhou (2005).

The importance of local leaders can also be seen from China's political structure. China adopts M-form (multidivisional form) central and local government relations, defined by Qian, Roland, and Xu (2006), where local government consists of "self-contained units" and complementary tasks are grouped together, in contrast to a U-form (unitary form) organization. They show that such a political structure facilitates regional coordination and political experimentation. For example, assume that a prefecture has two bureaus, a Department of Land and Resources and a Department of Housing and Urban-Rural Development, which are two key departments related to the allocation and use of land. Vertically, they are under the leadership of their respective ministers. Under the U-form political structure, the two departments lean towards the two ministries more than their located prefecture leader. However, with the M-form structure as in China, regional leaders, instead of departmental upper management, are empowered by the central government.

Why is it important for firms to connect to local leaders? The political arrangement enables prefecture leaders to easily coordinate each department in pursuit of regional policy targets. Xu (2011) uses special economic zone (SEZ) data as an example. The central government experimented with regional growth by designating Shenzhen as one of the early SEZs. If the transport department in Shenzhen did not coordinate well with the Shenzhen leader for some reason, for example, key people within the Communist Party disagreeing with the idea of SEZs were in the same faction as the transport department leader, it is highly possible for an SEZ to fail, considering the importance of infrastructure in the early stages of development. Well-executed coordination between dozens of government departments in Shenzhen, under the

[8] Another strand of literature emphasizes political connections with higher ranked politicians as the main factor for political promotion; see, for example, Shin, Adolph, and Liu (2012).

Shenzhen party secretary, played a key role in Shenzhen's success and, in turn, China's success (Xu, 2011).

B. *Land Market in China*

Land is central in the history of Chinese communist rule. At present, according to the Chinese constitution, all land is owned by the government. Before 1949, when the People's Republic of China was established, land was the most important factor in the Communist Party's success against its opponents. The main reason was that the communist leader Mao Zedong promised and implemented a policy of seizing land from landlords and distributing it to farmers, which made his recruitment for the People's Liberty Army a considerable success. Before the 1980s, individuals could not own or rent land, and collectivism led to the greatest famine from 1959 to 1961. After Deng Xiaoping initiated the Opening and Reform Policy in 1978, the household responsibility system, where individual farmers obtained rights to use land, began with an initial experiment in Anhui Province before gradually being implemented across the country. However, urban areas had to wait until 1987 for public auctions of land parcels to commence in China, first in the Shenzhen Special Economic Zone. Although the word "purchase" is used in this paper, it actually refers to buying land use rights for a certain number of years: 70 years for residential land; 50 years for industrial land; and 40 years for commercial land because the Chinese Constitution continues to endow all land ownership to the central government.[9]

The government is the sole seller in the primary land market in China and sets the land quota for each prefecture.[10] Central government leaders are sensitive about strategic food security and restrict land supply due to the shortage of arable land in China because many leaders in China experienced, first hand, the Great Famine in the 1960s. Local government instead has significant discretionary power over designating land types for different usages, such as industrial land use and residential land use.

China's land market is arguably the world's most corrupt (Zhu, 2012) because of the higher

[9] Please see http://www.gov.cn/zhengce/2020-12/25/content_5574204.htm.

[10] Total land area for sale in each prefecture is under the central government's control. Appendix Figure A1 presents indirect evidence that land area development is insensitive to local leader tenure, where local leaders should have a strong incentive to sell more land parcels immediately before their turnover.

number of regulations on land usage coupled with the weak institutional environment in China, although other countries also have extensive zoning regulations on housing and related land use. Corruption in land and housing pressured the central government to require that residential and commercial land be sold publicly beginning in 2004. However, corrupt land sales remained widespread. Regulations are omnipresent, from a building's floor-to-area ratio to the percentage of green space. Violation of any regulation leads to unsaleable building units because real estate developers need to satisfy all the regulations to obtain housing sale permits. In many prefectures, buyers are required to pay social security to the local government where the house is located, for at least one year, to be eligible to purchase. Local governments are also required to publicly auction some industrial land, but the selection of bidders is left to their discretion, as they have their own industrial policy and utilize soft information better than the central government. Therefore, while industrial land might be publicly auctioned, there are strict limitations on which firms may purchase the land. Therefore, implicitly industrial land is still sold through private negotiations. One piece of evidence is that most industrial land auctions have only one bidder, in contrast to the multiple bidders observed in the residential land market.

II. Data and Methodology

In this paper, we mainly use three datasets on publicly listed firms, land transactions, and local government officials. Tables A1 and A2 show the summary statistics of our main variables.

Data on publicly listed companies are from WIND (2016), which is the main provider of financial information in China, akin to WRDS in the United States. We manually collect publicly listed companies' subsidiaries' names from annual reports and match their land parcel data with land buyers' names. Changes in firm headquarters location for publicly listed firms are practically nonexistent in our Chinese data, although such location changes might be common in other countries (Klier and Testa, 2002).[11] One perspective is that due to China's weak institutional environment, firms usually gain some shadow support from the local government, with such backing frequently being on the boundary between legal and illegal. A

[11] Please see the report below on how difficult it was for one of the largest firms in China, Sany Corporation, the sixth-largest heavy equipment manufacturer in the world, to move its headquarters from Changsha to Beijing in 2012. Even in 2018 and 2019, some outlets still show Changsha as the headquarters. http://finance.ifeng.com/business/gs/20121122/7333448.shtml

prefecture losing the presence of a key firm's headquarters is often viewed very negatively in China. Concerns about revenge may make a firm reluctant to move.

Data on land transactions are manually collected from a government disclosure website (Land China, 2016) on which the central government has required all land transactions to be listed since 2006 to increase transparency. The data before 2006 are not complete and feature many missing parcels. Thus, our data period starts in 2006. The webpage provides the exact location of each land parcel, its sale price, unit price, floor-to-area ratio, buyer's name, land usage type, and land quality. Before selling a land parcel, a local government can increase the land's quality by providing electricity and water access.

Data on local political leaders, prefecture- or provincial-level party secretaries, mayors, and governors are from CSMAR (2017).[12] The party secretary is the de facto boss, and theoretically, the mayor or governor is responsible for the implementation of the region's policies. We collect each leader's name, age, education level, and where they go after leaving their position. Appendix Figure A2 shows the local leader turnover frequency. Appendix Table A3 lists local leader turnover frequency by province in our sample period and each province's GDP per capita in 2016, its economic development level. Except for the four directly controlled municipalities of Beijing, Chongqing, Shanghai, and Tianjin, provinces have fairly comparable levels of local leader turnover numbers, with the most in Guangdong and the least in Hainan, which are highly correlated with the number of prefectures in a province.

We use satellite data on nighttime light intensity obtained from the National Oceanic and Atmospheric Administration (NOAA, 2013) to serve as a proxy for land usage efficiency. The light intensity is observed between 8:30 pm and 10:00 pm by weather satellites that circle the earth 14 times a day (Henderson, Storeygard, and Weil, 2012). High-quality data that remove cloud coverage and fires are available between 1992 and 2013. Data after 2013 are not used due to scale and measure changes in the satellites, which is consistent with previous literature (Donaldson and Storeygard, 2016). The light intensity scale of these data is measured from 0 to 63, with a higher number indicating more intensity and more economic activity. Considering China's GDP per capita, approximately US$5,000 during our sample period, achieving light

[12] CSMAR has provided these political leaders' data in its database only since December 2017. Before that, when this research was ongoing, we manually collected data from prefectural and provincial statistical yearbooks for various years. We double-checked to ensure that the politician turnover data are accurate.

intensity above 63 is very rare. This is consistent with Henderson, Storeygard, and Weil's (2012) assertion that nighttime light intensity serves as a better proxy for development in developing countries. The use of nighttime light intensity data not only makes regional analysis possible at a very fine level but also alleviates our concern over data manipulation by the Chinese government.

We then match the listed companies' and their subsidiaries' names with land parcel buyers' names. We find that 2,188 out of 3,271 publicly listed companies acquire land in our data period. The average number of land parcels obtained by each land buyer is 20, averaging approximately two pieces of land per year. In addition, as displayed in Appendix Table A2, non-local public firms acquire approximately 87% of all land, which demonstrates the importance of non-local firms in a region's land market. Land types include industrial land, commercial land, residential land, and public land. Commercial land has similar properties to residential land, and both are developed by real estate developers, unlike industrial land, which is mostly sold to industrial firms.[13] 14% of land purchased is public land slated for public project buildings. These parcels are sold to firms for specific purposes, and construction takes the form of a public private partnership (PPP).

We then construct a firm by prefecture by year panel, using all possible combinations of all public firms and all prefectures from 2006 to 2016 as our main sample. Note that even if a firm with headquarters in a politician's previous work location does not seemingly overlap with the politician's time in that location as its IPO happened after the politician moved, we still include that firm in our data because it is highly likely that that firm was a major market player in that prefecture during that politician's tenure despite not being listed.[14]

Even though a firm's time-varying characteristics are less relevant for our analysis because they are absorbed by our firm-by-year fixed effects, one might still be curious about what different characteristics connected and unconnected firms have. Appendix Table A4 shows group differences between connected and unconnected firms by comparing key firm

[13] Chinese households prefer to live in residential-commercial mixed communities, as indicated by the higher prices there. Real estate developers bundle the sales of apartments with amenities that primarily consist of large shopping malls nearby that are often developed by the same residential developer.

[14] Under this circumstance, the connection with that leader can be either stronger or weaker. For example, it is highly possible that that politician played an important role in that firm being listed because there are large regulatory listing barriers in China (Piotroski and Zhang, 2014). The opposite side is that a firm might not have been important enough at that time to have connections with a politician.

characteristics commonly appearing in the literature, such as total factor productivity (TFP),[15] return on assets (ROA), return on equity (ROE), firm age, SOE status, debt over assets, and ln(Assets). Column (3) shows the differences between these two types of firms. Connected firms, on average, have higher TFP, ROA, and ROE; are younger; are more likely to be SOEs; have lower leverage; and are larger in terms of assets.

Below is our baseline empirical specification:

$$(1) \qquad Y_{fct} = \beta_0 + \beta_1 Connection_{fct} + \mu_{ft} + \lambda_{ct} + \delta_{p(f)p(c)} + \varepsilon_{fct}$$

where Y_{fct} indicates the number, area, and value of land parcels bought by firm f in prefecture c in year t, divided by the average number, area, and value of land parcels bought by all firms in prefecture c in year t, respectively. For ease of notation, we call these three dependent variables *Normalized Number*, *Normalized Area*, and *Normalized Value*. We use such normalized measures of land purchases to ease the interpretation of the coefficients. Specifically, with these measures, the coefficients could be interpreted as percent changes from the prefecture-year mean. $Connection_{fct}$ is a dummy variable that equals 1 if firm f has its headquarters in the previous work prefecture of the prefectural leader in prefecture c at year t. We also include province-level appointment cases in our variable construction. If a provincial leader in province B previously worked in prefecture A, then firms from A will switch their $Connection_{fct}$ dummy from 0 to 1 in the provincial capital prefecture of B. We use the capital prefecture rather than all the prefectures in that province as the baseline because provincial leaders' work and residential areas are still in the capital prefecture.[16] To be specific about our *Connection* variable, when the politician's turnover year is related to the departing and arriving politicians, the $Connection_{fct}$ variable relates to the politician with the longest tenure during the transition year. For example, if the incoming politician takes his or her position in October, then $Connection_{fct}$ is equal to 1 for the departing politician's connected firms for that

[15] Please see Appendix B for its calculation.

[16] If a land parcel is sold to a connected firm with help from a provincial politician who is one level directly above the prefecture party secretary, then the politician is most likely to choose the land and the prefecture party secretary he or she is familiar with. That is why we use provincial capital prefecture here where provincial leaders are located. In Appendix Table A5, we compare the effects of connections in capital prefectures to the effects of connections in non-capital prefectures for provincial leaders. In contrast to capital prefectures, the "go with the politician" effects in these non-capital prefectures are small and statistically insignificant.

prefecture-year. μ_{ft} denotes firm-by-year fixed effects. λ_{ct} denotes prefecture-by-year fixed effects. $\delta_{p(f)p(c)}$ denotes origin province $p(f)$ by destination province $p(c)$ fixed effects. We cluster the standard errors at the prefecture level, that is, we allow for serial correlation within a prefecture.

The coefficient on *Connection* is a difference-in-differences estimate of the effect of politicians and firms being previously located in the same place on land purchases. For example, if prefecture A had only two leaders in our sample period who moved from prefectures B and C, respectively, then the coefficient on *Connection* will be the weighted average of the difference between land purchases in prefecture A, obtained by firms headquartered in prefecture B when the leader is from prefecture B, relative to the latter period when the leader is from prefecture C, and the difference between land purchases in prefecture A, obtained by firms headquartered in prefecture C when the leader is from prefecture C, relative to the period when the leader is from prefecture B. Firms from other prefectures are used to estimate the various fixed effects.

Furthermore, we use $Connection\_Intensity_{fct}$ to measure the intensity of firm-politician connections. We assign a value of 5 to $Connection\_Intensity_{fct}$ when firm f is headquartered in the prefecture where a politician in prefecture c in year t has his or her last workplace. The value of $Connection\_Intensity_{fct}$ declines when the connection between a firm and a politician is remote. For example, if a politician sequentially moves from prefecture A to B, C, D, E, and F, where F is his or her current work prefecture, then firms headquartered in E enjoy a $Connection\_Intensity_{fct}$ value of 5 in prefecture F, while firms headquartered in A have a $Connection\_Intensity_{fct}$ value of 1. We still keep the unconnected firms' value at 0. Alternatively, we also consider another connection intensity measure indicated by the years a politician and a firm had a coworking relationship, as time is an important determinant in the development of trust and relationship formation. We estimate the following equation to test whether the effect increases with intensity:

$$(2) \quad Y_{fct} = \beta_0 + \beta_1 Connection_{fct} + \beta_2 Connection\_Intensity_{fct} + \mu_{ft} + \lambda_{ct} + \delta_{p(f)p(c)} + \varepsilon_{fct}$$

We expect β_1 to be positive in Equation (1), indicating that a connected firm will purchase more land parcels. Furthermore, we expect β_2 to be positive in Equation (2), implying that the "go with the politician" effect becomes more pronounced with higher connection intensity.

We also include a series of fixed effects. We have firm-by-year fixed effects to control for firm-specific time-varying characteristics, such as financial constraints, investment styles, whether a former politician was sitting on a firm's board, and the CEO's previous government work experience. We also include prefecture-by-year fixed effects to control for destination prefecture-specific, time-varying characteristics, such as government policy and economic prosperity. When only prefecture party secretaries are considered, the prefecture-by-year fixed effects absorb prefecture leader characteristics such as tenure length, age, education, and experience, as shown later in our robustness checks. The origin province-by-destination province fixed effects alleviate our concern that some similarities between regions might drive a firm's non-headquarters land purchase and the appointment of leaders between regions, for example, a similar industrial structure or similar culture. The ideal controls would be original prefecture-by-destination prefecture fixed effects. However, it is not only computationally impossible with 110,889 fixed effects from 333 prefectures in our data but also has an over-controlling issue that subsumes the essential effects we want to estimate. We include origin prefecture-by-destination prefecture fixed effects within a province in our robustness checks, as 52% of rotations happen within a province.

A key assumption for our identification is that local leader rotation is exogenous. The decision to move a local leader comes from an upper-level government, hardly relying on a single firm's characteristics. The existing literature reveals that the GDP growth rate and a local leader's political connections with upper-level governments are major determinants of prefecture leader turnover (Li and Zhou, 2005; Shih, Adolph, and Liu, 2012). These two factors can largely be absorbed by our two flexible fixed effect controls: origin prefecture-by-year fixed effects for the former factor and destination prefecture-by-year fixed effects, especially when only the prefecture party secretary is considered, for the latter factor.[17] The effect of $Connection_{fct}$ on firm outcomes Y_{fct} is therefore identified through plausibly exogenous

[17] The origin prefecture-by-year fixed effects are automatically controlled by firm-by-year fixed effects, since firms do not change their headquarters in our data.

matching between firms and officials caused by turnover.[18] Moreover, we will examine the dynamics of land purchases to rule out further potential endogeneity issues. Specifically, some underlying factor within a prefecture-firm pair might drive both being a connected firm and the purchase of more land parcels. That underlying factor is likely to be gradual, while changes in *Connection* occur instantly after leader turnover. We would thus expect firms to purchase more land parcels before becoming connected and after having been connected if that underlying factor is an important reason for increased land purchases. To test whether there is such a pattern in the data, we focus on firms that are about to become connected or have been connected until recently. If these firms are no different in these years than in other years in which they are unconnected, then this suggests that connected firms purchase more land parcels because of being connected rather than due to some underlying factor.

III. Results

Our results in Table 1 show that after a politician's move, geographically connected firms from the politician's previous workplace purchase significantly more land in the politician's new governing prefecture. Our results are robust to different dependent variables, such as *Normalized Number*, *Normalized Area*, and *Normalized Value*. Columns (1)-(3) include no fixed effects, and Columns (4)-(6) control for prefecture-by-year fixed effects, firm-by-year fixed effects, and origin province-by-destination province fixed effects. Robust standard errors are clustered at the prefecture level.[19]

The results are highly statistically significant, all at the 1% level, indicating that switching to a connected firm after a politician moves in a prefecture will result in that firm purchasing more land in that prefecture. With the addition of three groups of flexible fixed effects, the coefficients drop by approximately half from the regressions without controls in the first three columns. The larger coefficient for land size regression of 3.36 compared to the coefficient for

[18] In Appendix Figure A3, we show that the average cumulative abnormal return obtained by conducting standard event studies for local firms before the politician turnover event dates was approximately 0 but dropped sharply afterwards. This also provides rough evidence that politician turnover leads firm performance rather than vice versa. The data on the dates of politician turnover are manually collected from Baidu (2017). The link is: https://www.baidu.com (assessed March 19, 2018).

[19] In Appendix Table A6, we adopt other clustering levels, such as two-way clustering at the prefecture and year levels and province-level clustering. Given the small number of clusters at the province level, we also apply the wild bootstrap approach suggested by Cameron, Gelbach, and Miller (2008) and Roodman et al. (2019). The results are comparable with the baseline estimate.

land value regression of 3.00 also provides rough evidence on price favoritism for connected firms.

The coefficients are also economically sizable. To see this, consider the *Normalized Value* coefficient of 3.00. By construction, the mean of the dependent variable in a prefecture-year is 1, and the coefficient 3.00 implies a 300% change in the dependent variable from the prefecture-year mean. Specifically, if a firm switches from unconnected to connected, the value of land parcels bought by that firm in a prefecture-year will increase by three times the average value of land parcels bought by all firms in that prefecture-year.

In the next step, we address the concern that unobserved changes may simultaneously increase a firm's likelihood of becoming connected in a particular prefecture and year, as well as increase a firm's land purchases in that prefecture and year. Therefore, we first investigate whether firms that are about to become connected are systematically different from other firms. In Table 2, we add three pre-trend dummy variables that indicate firms that will become connected firms next year and within the next two or three years, respectively. We see that connected firms are no different shortly before becoming connected firms than other unconnected firms. Moreover, the coefficients on the *Connection* variable barely change across the columns compared to Table 1. The finding from Table 2 supports our interpretation that a firm purchases more land in a prefecture and year because it becomes connected in that prefecture and year, rather than because of some underlying gradual change that simultaneously increases land purchases and the chances of becoming connected.

To provide further support, we investigate the dynamics of the effect of *Connection* on land purchases in detail. We estimate a specification with many dummy variables for future, current, and past connected firms. The effect can be seen in Figure 3, where the dynamic effects of being connected, together with pre-trends (the three years before a firm has connections) and post-trends (the three years after a firm loses connections), are examined.[20] We define post-trends after a firm switches from being connected to unconnected as soon as a connected politician moves to a new workplace or simply retires. Not only does the figure reveal significant increases in land purchases for our defined connected firms after a politician's move, but the

[20] The omitted category consists of: (1) the fourth and all prior years before a firm has connections; (2) the fourth and all subsequent years after a firm loses connections.

pre-trends and post-trends are also insignificant, as shown by the overlap with the zero line between the confidence intervals. Connected firms are no different from other unconnected firms shortly before becoming connected or shortly after ceasing to be connected. During the period when a firm is connected, Figure 3 shows that the land purchased gradually increases with the incumbent politician's tenure. This might be because a politician's de facto power increases with tenure, giving that politician more discretion in allocating land to favored firms.

We now extend the dummy independent variable *Connection* to specify the extent of connection or *Connection Intensity*. We use two indicators. *Connection Intensity* (*Co-Working Sequence*) takes values from 0 to 5, where 5 indicates when a firm's headquarters is located in a local leader's most recent work region, immediately before his or her move, and 1 indicates the farthest case in which a firm's headquarters is located in the politician's initial work region. The other indicator is *Connection Intensity* (*Year*), which denotes how many years the politician had previously worked in a firm's headquarters prefecture. We then adopt Equation (2) to estimate whether the "go with the politician" effect increases with connection intensity. The results are shown in Table 3. The coefficients on these two connection intensity measures are positive and statistically significant, implying a higher "go with the politician" effect at higher intensity. We also show the intensity effects using separate dummies for connection intensity in Appendix Figure A4 and find a roughly similar pattern.

This paper emphasizes the political administrative boundary and prefecture party secretary, and the importance of the latter is addressed in the Institutional Background section. We now turn to economic geography. Specifically, we create a dummy variable *Connection (Adjacent Prefectures)* to denote firms in a local leader's previous work prefecture's neighboring prefectures. However, we find that the coefficients are insignificant and negative when we run the baseline regression with *Connection* replaced by this new connection measure, as shown in Appendix Table A7, confirming our political argument. The negative sign might be from the competitive relationship among leaders governing adjacent prefectures, as they usually vie for possible higher ranked provincial positions (Xu, 2011).

Table 4 shows that the unit land price paid by connected firms is significantly lower than that paid by other firms, suggesting favoritism by the local politician towards connected firms. Here, we compare a land parcel with other land parcels of the same type sold in the same year within a two kilometer radius. In this way, we could control for an important confounding factor: the location attractiveness of a land parcel. As connected firms are likely to obtain land parcels at better locations, failure to control for location attractiveness would underestimate their price discounts. In Column (1), we use the dummy independent variable *Connection* and find that connected firms pay significantly lower prices. In terms of economic significance, we find that being a connected firm reduces the cost per unit area of land for a firm by 51%, which is a very large effect. In Columns (2) and (3), we include both *Connection* and *Connection Intensity* to estimate whether the effect strengthens as connection intensity increases. These intensity effects are negative but statistically weak. We also show the intensity effects using separate dummies for connection intensity in Appendix Figure A5 and almost all these dummies show negative coefficients.

Table 5 demonstrates that land usage efficiency, as measured by nighttime light intensity, is lower after land purchases for connected firms, as seen from the significantly negative sign for the interaction term of *Post Deal* and *Connection*, where *Post Deal* is a dummy variable, equaling 1 after land purchase and 0 otherwise, and the dependent variable is *Normalized Light Intensity*, denoting the light intensity of a parcel sold to a firm in a prefecture and year divided by the average light intensity of all parcels sold to all firms in that prefecture and year. Many sold land parcels are smaller than the 1 km by 1 km grid cell of the light data (Donaldson and Storeygard, 2016). We view the proxy as still valid because a land parcel purchased by a publicly listed firm usually plays a key economic role in its matched grid cell, and others are very likely to be its economic extensions, such as wide roads and restaurants. We corroborate our measure of land usage efficiency in Appendix Table A8 by showing that the results are similar when: (1) we consider only land parcels that are larger than the median size and larger than the grid cell; (2) we consider only land parcels that are far away from the prefecture center so that there is much less light influence from neighboring land parcels;[21] (3) we consider only grid cells containing one land parcel to rule out confounding other parcel influences.

We use nighttime light intensity as a measure of land usage efficiency not only because light intensity has been frequently used as a regional economic activity measure but also because of

[21] The data on the prefecture center are manually collected from Google (2017). The link is: https://www.google.com/maps/ (accessed September 4, 2017).

particular institutional regulatory features in China. Keeping land undeveloped for more than two years is illegal in China. Connecting to an incumbent politician might shelter firms from this land regulation. Another regulation relating to light intensity concerns the housing timeframe selling restriction. Real estate developers are required to finish their selling within certain years to prevent developers from hoarding for future price appreciation gain. Again, connecting to a politician shelters firms from this regulation, as the local government is the implementer of these central government regulations. Both of the regulations constrain firms' profit maximization. Therefore, firms have incentives to evade them. The lower light growth result for connected firms suggests that investment by connected firms in the politician's new governing region is slower, and this largely rules out the information flow perspective.

Quantitatively, the coefficient on *Post Deal* × *Connection* is -0.04 in Column (1).[22] Our specification suggests that switching to a connected buyer will reduce the nighttime light intensity by 4% of the mean nighttime light intensity of land parcels sold to all firms in that prefecture and year. Henderson, Storeygard, and Weil (2012) find that the relation between nighttime light intensity and GDP is linear and that the estimated elasticity is "roughly 0.3". Assuming the elasticity to be 0.3, a decrease in nighttime light intensity by 4% means a reduction in the economic output of that land parcel by 1.2%. Our rough estimate suggests that the allocative distortion is quantitatively large. In Columns (2) and (3), we add the interaction between *Post Deal* and *Connection Intensity* to test whether the effect strengthens as connection intensity increases. We find negative coefficients on these interactions, although their significance is weak. We also show the intensity effects using separate dummies for connection intensity in Appendix Figure A6 and find all these dummies display negative coefficients.

However, the fact that a land parcel purchased by a connected firm has slower light growth can also be because such a firm might face higher management costs in operating on that land parcel (Giroud, 2013). The higher management costs are possibly driven by the longer distance between the firm's headquarters and the purchased land parcel location. We then test whether a longer distance between a firm's headquarters and a land parcel purchased by that firm implies

[22] The coefficient on *Post Deal* is negative. In China, local governments usually remove prior economic development activities on a land parcel before selling. One of the commonly seen forms is an "urban village", which is composed of crowded multi-story buildings in the urban area and usually heavily populated by the poor and transient. Their removal for new city developments is likely to result in vanishing nighttime light intensity for some years.

slower light intensity growth. We use *Remote* to denote a land parcel located outside of a firm's headquarters province, which is a dummy variable and equal to 1 if yes and 0 if no. We view this measure as preferable to simply using geographical distance because different provinces are highly likely to have different cultures, and the larger cultural difference also potentially increases a firm's management cost. By running the regression of a land parcel's nighttime light intensity on *Post Deal* and its interaction with *Remote*, we find that the coefficient on the interaction term is not significant, as displayed in Appendix Table A9, which suggests that the higher management cost factor might not be the reason for connected firms' lower light intensity growth.

In Appendix Table A10, we examine the effect of being connected on the extensive margin of our results, that is, whether being connected in a prefecture will lead a firm to enter that market. Here, a market is defined as a prefecture and year. Entry is defined as whether a firm purchases land, and it is a dummy variable equal to 1 after a firm purchases a land parcel until politician turnover in that prefecture and 0 otherwise. Our results are still highly significant. The appointment of a connected politician will increase the probability of entry into that governed prefecture by 3 percentage points, with the mean level of entry being 0.3%. We also investigate the connection intensity dimensions and find a similar pattern to Table 3. Specifically, the probability of entry is significantly increasing with connection intensity. Appendix Figure A7 further confirms this by showing the intensity effects using separate dummies for connection intensity.

IV. Robustness

Table 6 reports the results of various robustness tests. We only use *Normalized Value*, the value of land parcels bought by a firm in a prefecture and year divided by the average value of land parcels bought by all firms in that prefecture and year, as a dependent variable to save space, but our results are robust to using the other two measures of land purchases.

Column (1) displays the results of the baseline regressions, excluding politicians with terms of office shorter than one year. Such cases are, politically, for transition purposes. Related politicians know such transitions in advance, and it is highly possible that connected firms also know this information through their relationships. The coefficient barely changes.

Column (2) uses the subsample composed of only prefecture party secretaries. The coefficient on *Connection* is larger, as the party secretary is the de facto boss in his or her governed region. Considering the limited number of listed companies and their large influence on the local economy, the party secretary has a strong incentive to interact with these companies. Since in this case there is only one leader in a prefecture, the prefecture-by-year fixed effects also absorb the leader-by-year fixed effects, which control for the leader's characteristics such as sex, age, and education. The results of this subsample analysis also alleviate our concern that leader characteristics might drive our results. For example, a local leader close to retirement might be motivated to cater to firms by granting more land to them to obtain an after-retirement benefit.

Column (3) excludes locally born politicians whose hometown is the same as their governed prefecture. Persson and Zhuravskaya (2016) find behavioral differences between locally born politicians and others in China. For example, locally born politicians provide more public goods, and their rise to power has a higher probability of being supported by local elites. Since they have a higher probability of catering to local firms, excluding them might lead to a larger coefficient. However, locally born politicians might also have greater power to influence local firms because of their already intimate relationships, so the benefits for them from luring non-local firms are higher. The latter effect might dominate, as we obtain a reduced coefficient of 2.72 after excluding locally born politicians.

Column (4) considers only cross-province politician rotation. On the one hand, politicians might have had connections with firms within a province in the past. A within-province move for politicians will have a lower "go with the politician" effect. On the other hand, politicians who move outside a province might need to cater more to local firms since their connected firms have more information or management costs because the geographic distance is greater. Our results seem to support the latter argument by showing smaller coefficients.

Column (5) considers a subsample that excludes the listed company's subsidiaries. Subsidiaries might be different from listed companies because they are independent legal entities and possibly have somewhat different objective functions than parent companies. The estimated coefficient barely changes.

In Column (6), we exclude the firm's local headquarters' land purchases, which is another important category expected to behave differently from other companies. On the one hand, newly appointed local leaders might favor local firms to obtain support, which will increase our coefficient after dropping these local firms. On the other hand, local firms might collude with the incumbent local political actors to curb the newly appointed politician's influence, which will reduce our coefficient after eliminating local firms. The effects seem to cancel each other out, as we observe only a slightly changed coefficient.

In Column (7), we consider connected firms as only the firms headquartered in a politician's most recent work prefecture or only the firms with a *Connection Intensity* score of 5. Unsurprisingly, we see that the coefficient dramatically increases to 8.49.

Considering that within-province rotation occurs more frequently, accounting for 52% of all rotations, in Column (8), we add the origin prefecture-by-destination prefecture fixed effects within a province to the baseline regression. The coefficient slightly decreases to 2.71 and is still significant at the 1% level. In Appendix Table A11, we also consider four kinds of regressions with some variations in our flexible fixed effects. Our results are still highly significant, and the magnitudes of key coefficients are similar to those in the baseline regression. For example, when the dependent variable is *Normalized Value,* coefficients range from 2.90 to 3.08, comparable to the 3.00 in the baseline regression. We view the stability of these estimates as strong evidence that the "go with the politician" effect or the relationship between our variable of interest, *Connection*, and land purchases is unlikely to result from problems such as omitted variable bias.

In Column (9), we consider only provincial leaders. The effect is only 50% of the baseline result. We conduct further analysis regarding provincial and prefectural leaders in Appendix Table A12. Our results are still robust in each case, although the effect of connections with provincial leaders is approximately one-third of the effect of connections with prefectural leaders, as shown in Column (4). This is consistent with our Institutional Background information on how important the prefecture party secretary is. Additionally, prefectural leaders are responsible for the capital prefecture's land allocation, as shown by the seller's name in our data, which is the prefectural government. Although provincial leaders are more powerful, they might also be too powerful to be interested in interacting with individual firms.

In Appendix Table A13, we rule out politicians who have worked in Beijing, where the central government is located. This group of politicians might be special because they are often sent down to local governments to gain experience. In Appendix Table A14, we exclude firms headquartered in Beijing and Shanghai, not only for political reasons, as these two municipalities usually have special status politicians, but also for economic reasons, as a large fraction of listed firms are headquartered in these two locations. In both tables, our results are robust, with significant and comparable coefficients to our baseline regressions in Table 1.

V. Further Results

A. *Heterogeneity Analysis*

In this subsection, we turn to heterogeneity analyses of our baseline results by differentiating key relevant prefecture and firm characteristics. Table 7 shows our subsample analysis results. Groupings are categorized by each characteristic's median value if the characteristic is a continuous variable. We only report the results for the normalized value as dependent variable, but we have robust results for the alternative dependent variables.

Columns (1) and (2) differentiate prefectures according to their provinces' institutional quality—only available at the provincial level—constructed as a marketization index created by Fan, Wang, and Zhu (2011). Consistent with our corruption mechanism, the coefficient of 3.61 for the low institutional quality regions is almost twice as large as that for the high institutional quality regions of 1.84, and the group difference is also highly significant, indicating that the "go with the politician" effect is much stronger in more corrupt regions, which is consistent with our mechanism.

Columns (3) and (4) differentiate firms according to their state ownership status. Interestingly, we find that private firms display a stronger "go with the politician" effect than SOEs. Private firms might have a stronger political connection with the local government if they bribe local politicians or build reciprocal relationships with them. We conjecture that this is mainly due to the "grasp the large, let go of the small" SOE reform initiated by the central government in the late 1990s (Hsieh and Song, 2015). Following this reform, even a small local SOE might be part of a conglomerate that is most likely controlled by the central government. To improve the performance of central government-controlled SOEs, the central government uses a profitability measure to promote or demote central government-controlled SOE leaders (Yang, Wang, and Nie, 2013). Such central government-controlled SOE leaders usually have higher political ranks than the prefecture party secretary. For profitability or political faction considerations, SOEs might not follow a local politician's instructions to achieve their political goals, which might explain the smaller coefficient for SOEs.

Columns (5) and (6) show that small firms have a significantly greater coefficient than large firms. On average, there are fewer than ten publicly listed firms in each prefecture. Thus, "small" is a relative term here, and the small firms are, on average, still much larger than the majority of the firms registered in a prefecture. Usually, large firms are more geographically diverse, which makes them less responsive to local politicians' instructions. For example, using Vietnamese and Chinese data, Bai et al. (2017) find that larger firms have lower bribery expenditures compared to their revenues.

In Appendix Table A15, we further differentiate types of land use between industrial land and commercial/residential land. The latter two types are combined in this analysis, as they often share similar characteristics, such as comparable prices, and they are often bundled to be sold together. We see that the "go with the politician" effect for the subsample consisting of industrial land is much smaller than the effect for the subsample consisting of residential and commercial land, which is understandable because industrial land associated with industrial firms has a direct impact on the industrial sector and plays an important role in agglomeration economics. Regional GDP competition or internal GDP assessment from the guidelines of the upper-level government incentivizes politicians to promote growth and, consequently, less distortion, as shown by the smaller coefficients. The findings for different land types also support our favoritism channel because residential and commercial land is believed to suffer from more corruption than industrial land (Zhu, 2012).

We further explore firm heterogeneity by interacting firm characteristics with the *Connection* variable. If a firm experienced faster growth as measured by higher sales growth or had better performance as measured by higher ROA under a local leader's regime, then it is more likely that that firm would have a closer connection with that leader and the "go with the politician" effect is stronger in that leader's new governing region. However, as shown in Appendix Table A16, we do not find that the interaction terms are significant. This also does not support an alternative argument that an extortionary local leader coerced a local firm, usually a well performing firm, as it has more resources, in return for a future good land deal. A poorly performing firm would likely have strong incentive to bribe to increase their alternative profit (Lui, 1985) and not likely to be extorted to connect.

We also examine local leader's heterogeneity effect. The "go with the politician" effect found in this paper could also be driven by politicians with lower chances of promotion or politicians who are myopic due to a shorter term duration. We test both conjectures in Appendix Table A17. For the effect of promotion chances, we interact *Connection* with leader age, as younger leaders are more likely to be promoted. For myopia, we interact *Connection* with leaders' term length. The coefficients on the interaction terms in both cases are insignificant. The fact that the "go with the politician" effects are not stronger in these two cases might be from connected firms' concern that they would not enjoy future regulatory benefits from a politician who would soon leave office or a politician with a lower chance of promotion. The successor has a lower probability to follow a non-promoted politician's suggestion to continue to favor a previously connected firm such as granting regulatory benefits, which leaves connected firms less willing to purchase land. The two cases also show that our baseline results are not driven by some special groups of politicians.

B. *Anti-corruption Campaign Since 2013*

Here, we exploit a large-scale nationwide anti-corruption campaign initiated by the central government in 2013 to test its effect on our analysis of the "go with the politician" phenomenon. The anti-corruption campaign began after the conclusion of the 18th National Congress of the Communist Party of China in 2013. As of 2016, more than 100,000 people had been indicted for corruption.[23] This large-scale event has been exploited to a large extent, particularly in studying corporate matters, e.g., Chen and Kung (2019). To investigate whether the magnitude

[23] Please see https://www.economist.com/china/2015/10/22/robber-barons-beware.

of our baseline estimate changes after the anti-corruption campaign began, we add an interaction term of *Connection* and a year dummy denoting whether that year is 2013 or later. The negative coefficient on the interaction term in Columns (1)-(3) of Table 8 demonstrates that the “go with the politician” effect became weaker after 2013, which is consistent with our favoritism mechanism.

We also investigate whether connected firms generated lower stock returns when their connected local leaders were investigated for corruption. We collect the profiles of arrested local leaders from the CSMAR database (CSMAR, 2017) and identify 141 politicians who were arrested for corruption. We use the day when a politician was taken away for investigation as the event day. In China, the government’s announcement of a corruption investigation is often viewed as an official statement of a politician’s corruption. We also manually search Wikipedia and some official websites, such as xinhuanet.com, to cross-check the event dates. Then, for each of the 141 politicians, we identify his or her work prefectures and those firms with land purchase records in these prefectures and calculate their cumulative abnormal returns (CARs) around the investigation time of the politician. We adopt the Fama-French three-factor model to compute CARs (Fama and French, 1993; MacKinlay, 1997). The event window is [-1, 1] with 0 denoting the event date. To account for possible information delay or leakage, we use alternative event windows such as [-10, 10]. The estimation window is [-140, -20], lasting for 120 trading days, which is typical in the event study literature.

Appendix Table A18 demonstrates that connected firms suffer a larger drop in market value around the time of investigation of corrupt politicians than unconnected firms. Specifically, the results show that over the three trading days around the corruption investigation, the firm value of connected firms declined by 0.6% more than that of unconnected firms. We also add specifications with firm-by-investigation year fixed effects, which would effectively control for the financial conditions and other firm characteristics that would affect a firm’s stock returns in an investigation year. We essentially compare the CARs for the same firm in the same year but different exact dates when having purchased land from its connected and unconnected corrupt leaders. A firm during the investigation of a connected leader experienced 0.5% lower stock returns for the event window of [-1, 1], further supporting our favoritism mechanism. For the alternative event window [-10, 10], the CARs are much larger. Specifically, when a prefecture leader was investigated for corruption, the connected firms had 2.4% lower returns than unconnected firms and 2.3% lower returns than unconnected cases for the same firm and year.

C. *Locally Headquartered Firms*

Another important category of firms consists of locally headquartered firms. Since they foster business-government relationships, they should also be our investigation targets, especially because connected firms were previously locally headquartered firms. Appendix Table A19 reports the results of regressions similar to our baseline regressions except that *Connection* is replaced by *Local,* which takes a value of 1 if the firm is headquartered in that prefecture and 0 otherwise. We first test the effects on land purchases. The coefficients on *Local* are significantly positive, and the magnitudes are comparable to the coefficients in our previous baseline regressions, suggesting that local firms might be as important as connected firms to a local leader.

We also run a similar regression for the discounted price argument, with the connected firm dummy replaced by the local firm dummy. With this independent variable, we consider the unit land price paid by a firm divided by the average unit price of other land of the same type within a 2 km radius sold in the same year. The coefficient on *Local* is statistically insignificant, suggesting that connected firms play a more dominant role than local firms in obtaining corruption benefits. Both a local firm and a new politician might attempt to foster connections. It might not be optimal for local politicians to grant price discounts during initial encounters.

D. *Hometown Effect*

We do not find that firms from the leader's hometown exhibit such a "go with the politician" pattern. Hodler and Raschky (2014) provide suggestive evidence that leaders favor their hometowns. However, the recent existing studies on China have mixed findings. For example, Fisman et al. (2018) show positive effects of hometown ties in academia whereas Fisman et al. (2020) show negative effects in politician selection. We repeat our baseline regression with our *Connection* variable replaced by *Hometown*, indicating whether a firm's headquarters is in the same place as the prefecture leader's hometown. The insignificant coefficients on *Hometown,* as shown in Appendix Table A20, distinguish our results from the hometown regional favoritism argument. Our results might be due to most leaders being in middle age. In fact, most local leaders left their hometowns for college at age 18, then worked where the central government assigned them, usually not their hometowns. Therefore, these individuals have little connection with hometown firms. Our working experience measure makes more sense for connections because both the local leader and the then-local firms are geographically close and have related economic objectives.

E. *The Aggregate Consequences of "Go with the Politician"*

The existing literature surveyed by Xu (2011) argues that the central government in China implicitly sets a rule to promote local government leaders who can achieve higher local growth. However, it is inevitable that politicians have incentives to enrich themselves, as seen in reports of significant fortunes made by politicians, many of whom were linked through firms, during the recent anti-corruption campaign. Appendix Table A21 shows our test of the aggregate "go with the politician" effect. The dependent variables are whether a politician would be promoted in the next period and the logarithm of the next period prefectural GDP. Our results show that allocating more land to connected firms does not generate higher prefectural GDP, thus supporting our previous favoritism channel. Moreover, these politicians are less likely to be promoted, which suggests that they might gain private fortunes through these firms instead.

VI. Conclusion

In this paper, we exploit an important institutional feature of China's political system, frequent local leader turnover in different prefectures generating within-firm location variation in connections to the local leader, to identify the effects of firm-politician relationships on resource allocation. We find that geographically connected firms, defined as firms headquartered in the prefecture that is the politician's previous work location, purchase more land in the politician's new governing prefecture. Using the surrounding average land price as a benchmark, the average price paid by connected firms is lower than that paid by unconnected firms. These land parcels are also less efficiently developed by these firms, as measured by slower growth in nighttime light intensity. Our baseline regression controls for prefecture-by-year, firm-by-year, and original province-by-destination province fixed effects, largely alleviating endogeneity concerns.

Although it is often claimed that a rotation policy is a key method for reducing corruption, our findings suggest that further monitoring of local officials is needed because there might be corruption spillovers occurring from one region to another. Broadly speaking, this paper complements the literature on the personnel economics of the state, which lacks studies on local government leaders as policy implementers according to the review article by Finan, Olken, and Pande (2015).

China's soaring housing prices and ghost cities have received considerable attention (Glaeser et al., 2017). This paper suggests that a firm's political connections might shelter it from housing construction regulations, thereby reducing the actual housing supply and in turn contributing to high housing prices. Selling housing units outside mandated time frames by demanding a high sale price leaves many constructed housing units unoccupied and partially contributes to the emergence of ghost towns.

Although politicians' geographical mobility might not be as intensive in other countries as in this paper's setting, such cases are not unusual. Hodler and Raschky (2014) study how leaders enriched their hometowns across the world. The government ownership and distribution of land studied in this paper is also common in other countries. Even in the United States, 28% of land is owned by the government (Vincent, Hanson, and Bjelopera, 2014). The misuse of land by governments often attracts widespread attention. For example, after India's 1990s liberalization, land was sold well below market prices in the name of public-private partnerships. We believe that our study has implications for other countries.

REFERENCES

Allen, Franklin, Jun Qian, and Meijun Qian. 2005. "Law, Finance, and Economic Growth in China." *Journal of Financial Economics* 77 (1): 57-116.

Bai, Jie, Seema Jayachandran, Edmund J. Malesky, and Benjamin A. Olken. 2017. "Firm Growth and Corruption: Empirical Evidence from Vietnam." *The Economic Journal* 129 (618): 651-677.

Baidu. 2018. "Baidu Search Engine." https://www.baidu.com (assessed March 19, 2018).

Bailey, Michael, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. 2018. "Social Connectedness: Measurement, Determinants, and Effects." *Journal of Economic Perspectives* 32 (3): 259-80.

Blanchard, Olivier, and Andrei Shleifer. 2001. "Federalism with and without Political Centralization: China versus Russia." *IMF Economic Review* 48 (1): 171-179.

Brogaard, Jonathan, Matthew Denes, and Ran Duchin. 2015. "Political Connections, Incentives and Innovation: Evidence from Contract-Level Data." University of Washington Working Paper.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90 (3): 414-427.

Chen, Ting, and James Kai-sing Kung. 2019. "Busting the 'Princelings': The Campaign against Corruption in China's Primary Land Market." *The Quarterly Journal of Economics* 134 (1): 185–226.

CSMAR. 2017. "China Stock Market & Accounting Research Database." https://www.gtadata.com/csmar.html?v=#/index (accessed March 3, 2018).

Donaldson, Dave, and Adam Storeygard. 2016. "The View from Above: Applications of Satellite Data in Economics." *Journal of Economic Perspectives* 30 (4): 171-98.

Faccio, Mara. 2006. "Politically Connected Firms." *American Economic Review* 96 (1): 369-386.

Faccio, Mara, Ronald W. Masulis, and John J. McConnell. 2006. "Political Connections and Corporate Bailouts." *Journal of Finance* 61 (6): 2597-2635.

Fama, Eugene F. and Kenneth R. French. 1993. "Common Risk Factors on Stocks and Bonds." *Journal of Financial Economics* 33 (1): 3–56.

Fan, Gang, Xiaolu Wang, and Hengpeng Zhu. 2011. *NERI Index of Marketization in China's Provinces* (in Chinese). Beijing: Economic Science Press.

Finan, Frederico, Benjamin A. Olken, and Rohini Pande. 2015. "The Personnel Economics of the State." NBER Working Paper No. 21825.

Fisman, Raymond. 2001. "Estimating the Value of Political Connections." *American Economic Review* 91 (4): 1095-1102.

Fisman, Raymond, Jing Shi, Yongxiang Wang, and Rong Xu. 2018. "Social Ties and Favoritism in Chinese Science." *Journal of Political Economy* 126 (3): 1134-1171.

Fisman, Raymond, Jing Shi, Yongxiang Wang, and Weixing Wu. 2020. "Social Ties and the Selection of China's Political Elite." *American Economic Review* 110 (6): 1752-1781.

Giroud, Xavier. 2013. "Proximity and Investment: Evidence from Plant-level Data." *The Quarterly Journal of Economics* 128 (2): 861-915.

Glaeser, Edward, Wei Huang, Yueran Ma, and Andrei Shleifer. 2017. "A Real Estate Boom with Chinese Characteristics." *Journal of Economic Perspectives* 31 (1): 93-116.

Goldman, Eitan, Joerg Rocholl, and Jongil So. 2013. "Politically Connected Boards of Directors and the Allocation of Procurement Contracts." *Review of Finance* 15 (5): 1617-1648.

Google. 2017. "Google Maps." https://www.google.com/maps/ (accessed September 4, 2017).

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994-1028.

Hodler, Roland, and Paul A. Raschky. 2014. "Regional Favoritism." *The Quarterly Journal of Economics* 129 (2): 995-1033.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." The *Quarterly Journal of Economics* 124, no. 4: 1403-1448.

Hsieh, Chang-Tai, and Zheng Michael Song. 2015. "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China." *Brookings Papers on Economic Activity*.

Jia, Ruixue, Masayuki Kudamatsu, and David Seim. 2015. "Political Selection in China: The Complementary Roles of Connections and Performance." *Journal of the European Economic Association* 13 (4): 631-668.

Khwaja, Asim Ijaz, and Atif Mian. 2005. "Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market." *The Quarterly Journal of Economics* 120 (4): 1371-1411.

Klier, Thomas., and William Testa. 2002. "Location Trends of Large Company Headquarters During the 1990s," Federal Reserve Bank of Chicago Economic Perspectives 26 (2): 12-26.

Land China. 2016. "Announcement on Land Transfer." https://www.landchina.com/ (accessed July 5, 2017).

Li, Hongbin, Lingsheng Meng, Qian Wang, and Li-An Zhou. 2008. "Political Connections, Financing and Firm Performance: Evidence from Chinese Private Firms." *Journal of Development Economics* 87 (2): 283-299.

Li, Hongbin, and Li-An Zhou. 2005. "Political Turnover and Economic Performance: The Incentive Role of Personnel Control in China." *Journal of Public Economics* 89 (9): 1743-1762.

Lui, Francis T. 1985. "An Equilibrium Queuing Model of Bribery." *Journal of Political Economy* 93 (4): 760-781.

MacKinlay, A. Craig. 1997. "Event Studies in Economics and Finance." *Journal of Economic Literature* 35 (1): 13-39.

Marmaros, David, and Bruce Sacerdote. 2006. "How Do Friendships Form?" *The Quarterly Journal of Economics* 121 (1): 79–119.

Ministry of Land and Resources of China. 1999-2016. *China Land and Resources Statistical Yearbook.* Beijing: Geological Publishing House.

National Bureau of Statistics of China. 1999-2016. *China Statistical Yearbook*. Beijing: China Statistics Press.

National Bureau of Statistics of China. 2017. *China City Statistical Yearbook*. Beijing: China Statistics Press.

National Oceanic and Atmospheric Administration (NOAA). 2013. "Version 4 DMSP-OLS Nighttime Lights Time Series." NOAA Earth Observation Group. https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html (accessed July 10, 2017).

Persson, Petra, and Ekaterina Zhuravskaya. 2016. "The Limits of Career Concerns in Federalism: Evidence from China." *Journal of the European Economic Association* 14 (2): 338-374.

Piotroski, Joseph D., and Tianyu Zhang. 2014. "Politicians and the IPO Decision: The Impact of Impending Political Promotions on IPO Activity in China." *Journal of Financial Economics* 111 (1): 111-136.

Qian, Yingyi, Gerard Roland, and Chenggang Xu. 2006. "Coordination and Experimentation in M-form and U-form Organizations." *Journal of Political Economy* 114 (2): 366-402.

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb. 2019. "Fast and Wild: Bootstrap Inference in Stata Using Boottest." *The Stata Journal* 19 (1): 4-60.

Schoenherr, David. 2019. "Political Connections and Allocative Distortions." *Journal of Finance,* 74(2), 543-586.

Shi, Xiangyu, Tianyang Xi, Xiaobo Zhang, and Yifan Zhang. 2020. "Moving 'Umbrella': Bureaucratic Transfers, Collusion, and Rent-seeking in China." Working Paper Series of the China Center for Economic Research.

Shih, Victor, Christopher Adolph, and Mingxing Liu. 2012. "Getting Ahead in the Communist Party: Explaining the Advancement of Central Committee Members in China." *American Political Science Review* 106 (1): 166-187.

Stiglitz, Joseph E., and Andrew Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *American Economic Review* 71 (3): 393–410.

Tahoun, Ahmed. 2014. "The Role of Stock Ownership by US Members of Congress on the Market for Political Favors." *Journal of Financial Economics*, 111(1): 86-110.

Verbrugge, Lois M. 1983. “A Research Note on Adult Friendship Contact: A Dyadic Perspective.” *Social Forces* 62(1): 78–83.

Vincent, Carol Hardy, Laura A. Hanson, and Jerome P. Bjelopera. 2014. “Federal Land Ownership: Overview and Data.” Washington, DC, USA: Congressional Research Service.

WIND. 2016. “Wind Financial Terminal.” https://www.wind.com.cn/ (accessed August 10, 2017).

Wu, Jing, Joseph Gyourko, and Yongheng Deng. 2012. “Evaluating Conditions in Major Chinese Housing Markets: Dataset.” *Regional Science and Urban Economics* 42 (3): 531-543. http://real-faculty.wharton.upenn.edu/gyourko/chinese-residential-land-price-indexes/ (accessed April 4, 2021).

Xu, Chenggang. 2011. “The Fundamental Institutions of China’s Reforms and Development.” *Journal of Economic Literature* 49 (4): 1076-1151.

Yang, Ruilong, Yuan Wang, and Huihua Nie. 2013. “The Political Promotion for Quasi-Government Officers: Evidence from Central State-owned Enterprises in China.” MPRA Paper No. 50317, University Library of Munich, Germany.

Zhu, Jiangnan. 2012. “The Shadow of the Skyscrapers: Real Estate Corruption in China.” *Journal of Contemporary China* 21 (74): 243-260.

Zipf, George Kingsley. 1949. *Human Behavior and the Principle of Least Effort. Cambridge*. MA: Addison-Wesley Press.

Figures and Tables

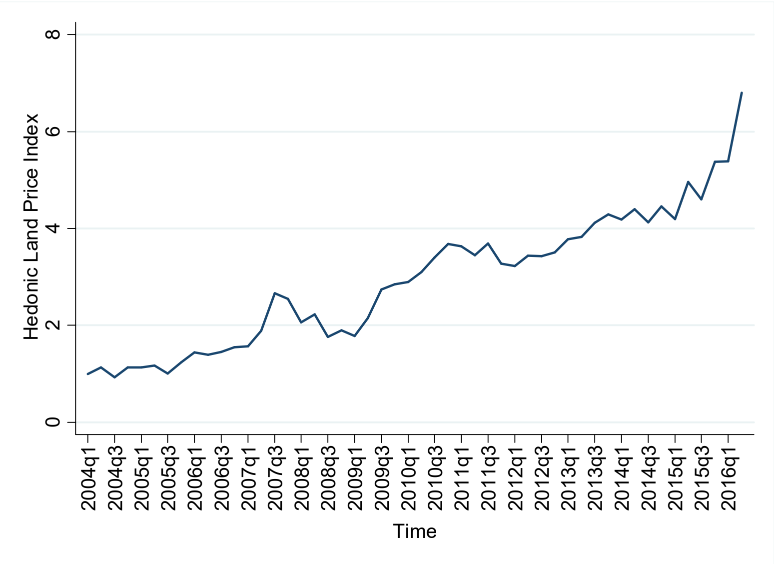


FIGURE 1: HEDONIC LAND PRICE INDEX (2004-2016)

Notes: The hedonic land price data are from Wu, Gyourko, and Deng (2012).

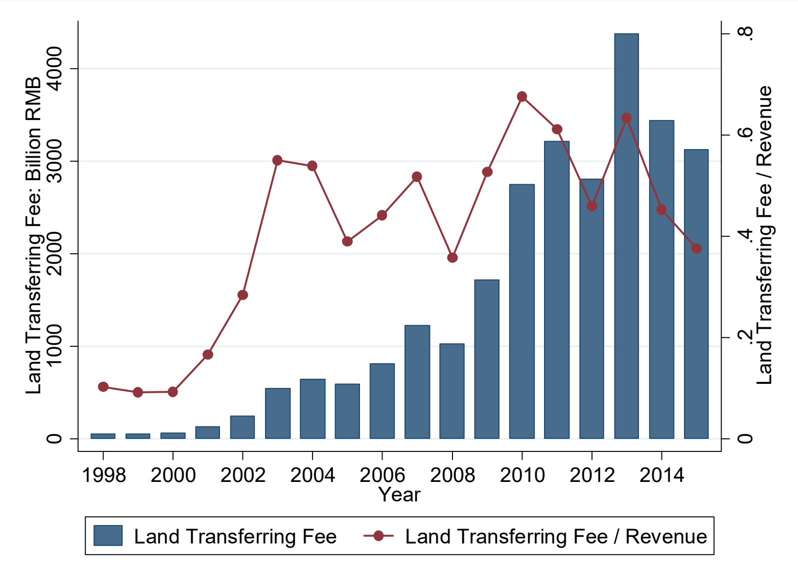


FIGURE 2: LOCAL GOVERNMENT REVENUE STRUCTURE (1998-2015)

Notes: Data are from the National Bureau of Statistics of China (1999-2016) and the Ministry of Land and Resources of China (1999-2016).

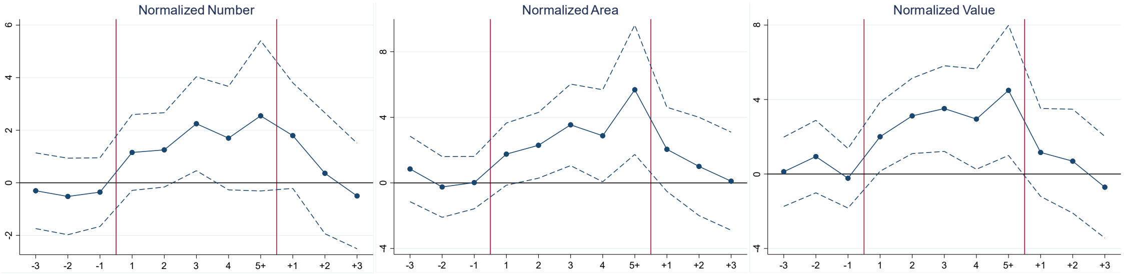


FIGURE 3: THE DYNAMICS OF A FIRM BEING CONNECTED AND UNCONNECTED

Notes: The figure plots coefficient estimates of dummy variables accounting for each of the three years before a firm becomes connected (-3, -2, -1), each of the first four years of being connected (1-4), the fifth and all subsequent years (5+) of being connected, and each of the three years after a firm loses connections when the politician rotates out of that prefecture (+1, +2, +3). The omitted category consists of: (1) the fourth and all prior years before a firm has connections; (2) the fourth and all subsequent years after a firm loses connections. The blue solid lines plot coefficient estimates for each dummy variable, and the dashed lines indicate the upper and lower limits of the 95% confidence intervals. The two vertical lines indicate the beginning and the end of connection, respectively. These estimates stem from baseline regressions with Normalized Number, Normalized Area, and Normalized Value as dependent variables, indicating the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. For each subfigure, one of these dependent variables is regressed on the aforementioned 11 dummy variables and the full set of prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects, using prefecture-firm-year-level land purchase data between 2006 and 2016. Standard errors used to construct the confidence intervals are adjusted for prefecture clustering.

TABLE 1—BASELINE RESULTS

Dependent variable	(1) Normalized Number	(2) Normalized Area	(3) Normalized Value	(4) Normalized Number	(5) Normalized Area	(6) Normalized Value
Connection	5.06	6.66	5.87	2.37	3.36	3.00
	(0.25)	(0.37)	(0.34)	(0.25)	(0.33)	(0.31)
Prefecture-year FEs	No	No	No	Yes	Yes	Yes
Firm-year FEs	No	No	No	Yes	Yes	Yes
Origin-destination province FEs	No	No	No	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113	9,547,267	9,547,267	9,451,113
R-squared	0.00	0.00	0.00	0.19	0.06	0.06

Notes: This table reports the estimated effect of whether a firm and a politician in a prefecture-year have a connection on the firm's land purchases in that prefecture-year. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Columns (1)-(3) do not control for any fixed effects. Columns (4)-(6) control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

TABLE 2—PRE-CONNECTION TRENDS

Dependent variable	(1) Normalized Number	(2) Normalized Area	(3) Normalized Value
Indicator for Three Years before Connection	-0.20	0.22	0.01
	(0.29)	(0.42)	(0.39)
Indicator for Two Years before Connection	0.02	0.10	0.57
	(0.31)	(0.37)	(0.39)
Indicator for One Year before Connection	0.05	0.01	-0.22
	(0.29)	(0.32)	(0.33)
Connection	2.37	3.36	3.00
	(0.25)	(0.33)	(0.31)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
p(All pre-trends = zero)	0.88	0.96	0.36
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06

Notes: This table reports the estimated effect of whether a firm and a politician in a prefecture-year have a connection on the firm's land purchases in that prefecture-year, by adding three pre-trend indicators to the baseline regression. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Indicators for one, two, and three years before connection are dummy variables, taking the value of 1 for firms that will become connected firms next year and within the next two or three years, and 0 otherwise, respectively. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

TABLE 3—CONNECTION INTENSITY AND LAND PURCHASES

	(1)	(2)	(3)
Dependent variable	Normalized Number	Normalized Area	Normalized Value
Panel A: Connection Intensity (Co-Working Sequence)			
Connection	-0.39	-0.29	-0.13
	(0.40)	(0.51)	(0.47)
Connection Intensity (Co-Working Sequence)	1.21	1.60	1.38
	(0.18)	(0.23)	(0.21)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06
Panel B: Connection Intensity (Year)			
Connection	1.14	1.95	1.78
	(0.31)	(0.38)	(0.38)
Connection Intensity (Year)	0.39	0.44	0.38
	(0.08)	(0.10)	(0.10)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06

Notes: This table reports the estimated effect of the intensity of connection between a firm and a politician in a prefecture-year on the firm's land purchases in that prefecture-year. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Connection Intensity (Co-Working Sequence) takes values from 0 to 5. Five indicates the closest case where a firm's headquarters is located in a local leader's most recent work prefecture, immediately before the leader's move. One indicates the farthest case where a firm's headquarters is located in a local leader's initial work prefecture. Connection Intensity (Year) denotes how many years a local leader previously worked in the firm's headquarters prefecture. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

TABLE 4—CONNECTION AND LAND PRICE

	(1)	(2)	(3)
Dependent variable	Unit Price / Average Unit Price within a 2 KM Radius		
Connection	-0.51	-0.36	-0.46
	(0.13)	(0.17)	(0.16)
Connection Intensity (Co-Working Sequence)		-0.07	
		(0.05)	
Connection Intensity (Year)			-0.02
			(0.03)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	13,227	13,227	13,227
R-squared	0.45	0.45	0.45

Notes: This table reports the estimated effect of Connection or Connection Intensity on land parcel price a firm pays. The unit of analysis is land parcel. The dependent variable is the unit price of a land parcel over average unit price of other similar land parcels (same type and sold in the same year) within a 2 km radius of that land parcel. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Connection Intensity (Co-Working Sequence) takes values from 0 to 5. Five indicates the closest case where a firm's headquarters is located in a local leader's most recent work prefecture, immediately before the leader's move. One indicates the farthest case where a firm's headquarters is located in a local leader's initial work prefecture. Connection Intensity (Year) denotes how many years a local leader previously worked in the firm's headquarters prefecture. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

TABLE 5—CONNECTION AND LAND USAGE EFFICIENCY

	(1)	(2)	(3)
Dependent variable		Normalized Light Intensity	
Post Deal	-0.053	-0.053	-0.053
	(0.011)	(0.011)	(0.011)
Post Deal × Connection	-0.043	-0.042	-0.034
	(0.012)	(0.019)	(0.016)
Post Deal × Connection Intensity (Co-Working Sequence)		-0.000	
		(0.004)	
Post Deal × Connection Intensity (Year)			-0.002
			(0.002)
Land parcel FEs	Yes	Yes	Yes
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Observations	262,119	262,119	262,119
R-squared	0.916	0.916	0.916

Notes: This table reports the estimated effect of Connection or Connection Intensity on land usage efficiency. The unit of analysis is land parcel. Normalized Light Intensity is the light intensity of a parcel sold to a firm in a prefecture-year divided by the average light intensity of all parcels sold to all firms in that prefecture-year. Post Deal is a dummy variable, taking the value of 1 after a land parcel's sale, and 0 otherwise. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Connection Intensity (Co-Working Sequence) takes values from 0 to 5. Five indicates the closest case where a firm's headquarters is located in a local leader's most recent work prefecture, immediately before the leader's move. One indicates the farthest case where a firm's headquarters is located in a local leader's initial work prefecture. Connection Intensity (Year) denotes how many years a local leader previously worked in the firm's headquarters prefecture. All regressions control for prefecture-by-year, firm-by-year, and land parcel fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

TABLE 6—ROBUSTNESS CHECKS

Dependent variable	Normalized value				
	Exclude Term < 1 Year	Only Prefecture Party Secretaries	Exclude Locally Born Leaders	Only Cross-Province Rotation	Exclude Subsidiary Firms' Land Purchase
	(1)	(2)	(3)	(4)	(5)
Connection	3.00	3.64	2.72	2.30	3.01
	(0.31)	(0.36)	(0.30)	(0.31)	(0.31)
Prefecture-year FEs	Yes	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes	Yes	Yes
Observations	9,451,113	9,451,113	9,451,113	9,451,113	9,414,019
R-squared	0.06	0.06	0.06	0.06	0.07
	Exclude Local Firms' Land Purchase	Immediate Working History	Origin-Destination Prefecture FEs within Provinces	Provincial Leaders	
	(6)	(7)	(8)	(9)	
Connection	2.92	8.49	2.71	1.51	
	(0.31)	(0.99)	(0.24)	(0.47)	
Prefecture-year FEs	Yes	Yes	Yes	Yes	
Firm-year FEs	Yes	Yes	Yes	Yes	
Origin-destination province FEs	Yes	Yes	Yes	Yes	
Origin-destination prefecture FEs within provinces	No	No	Yes	No	
Observations	9,352,415	9,451,113	9,450,975	9,451,113	
R-squared	0.06	0.06	0.07	0.06	

Notes: This table reports the estimated effect of whether a firm and a politician in a prefecture-year have a connection on the firm's land purchases in that prefecture-year. Normalized Value denotes the value of land parcels bought by a firm in a prefecture-year divided by the average value of land parcels bought by all firms in that prefecture-year. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Column (1) excludes local politicians with terms less than one year. Column (2) considers only prefecture party secretaries. Column (3) excludes locally born politicians. Column (4) considers only cross-province politician rotations. Column (5) excludes listed companies' subsidiary firms' land purchases. Column (6) excludes local firms' land purchases. Column (7) defines connections only for firms headquartered in the politician's most recent work prefecture. Column (8) includes origin prefecture-by-destination prefecture fixed effects within a province. Column (9) considers only connections with provincial leaders. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

TABLE 7—HETEROGENEITY ANALYSIS: INSTITUTIONAL QUALITY, SOE STATUS, AND FIRM SIZE

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Normalized Value					
	Low Institutional Quality	High Institutional Quality	Non-SOE	SOE	Small Firms	Large Firms
Connection	3.61	1.84	3.79	2.36	4.49	1.35
	(0.41)	(0.42)	(0.53)	(0.33)	(0.51)	(0.24)
Prefecture-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes	Yes	Yes	Yes
p(Equal coefficients)	0.00		0.02		0.00	
Observations	5,250,466	4,200,647	3,278,770	6,172,343	4,729,916	4,721,196
R-squared	0.05	0.09	0.09	0.02	0.07	0.01

Notes: This table reports the estimated effect of whether a firm and a politician in a prefecture-year have a connection on the firm's land purchases in that prefecture-year. Normalized Value denotes the value of land parcels bought by a firm in a prefecture-year divided by the average value of land parcels bought by all firms in that prefecture-year. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Columns (1) and (2) differentiate provinces according to their institutional quality. Columns (3) and (4) differentiate firms according to their ownership status. Columns (5) and (6) differentiate firms according to their asset sizes. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

TABLE 8—THE ANTI-CORRUPTION CAMPAIGN'S EFFECTS

	(1)	(2)	(3)
Dependent variable	Normalized Number	Normalized Area	Normalized Value
Connection	2.83	4.17	3.68
	(0.30)	(0.41)	(0.38)
Connection × 1(Year>=2013)	-1.06	-1.84	-1.54
	(0.30)	(0.42)	(0.40)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06

Notes: This table tests whether the anti-corruption campaign decreased the effect of a firm's connection with a politician in a prefecture-year on the firm's land purchases in that prefecture-year. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. 1(Year>=2013) is a dummy variable and is equal to 1 if the year is or after 2013 and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

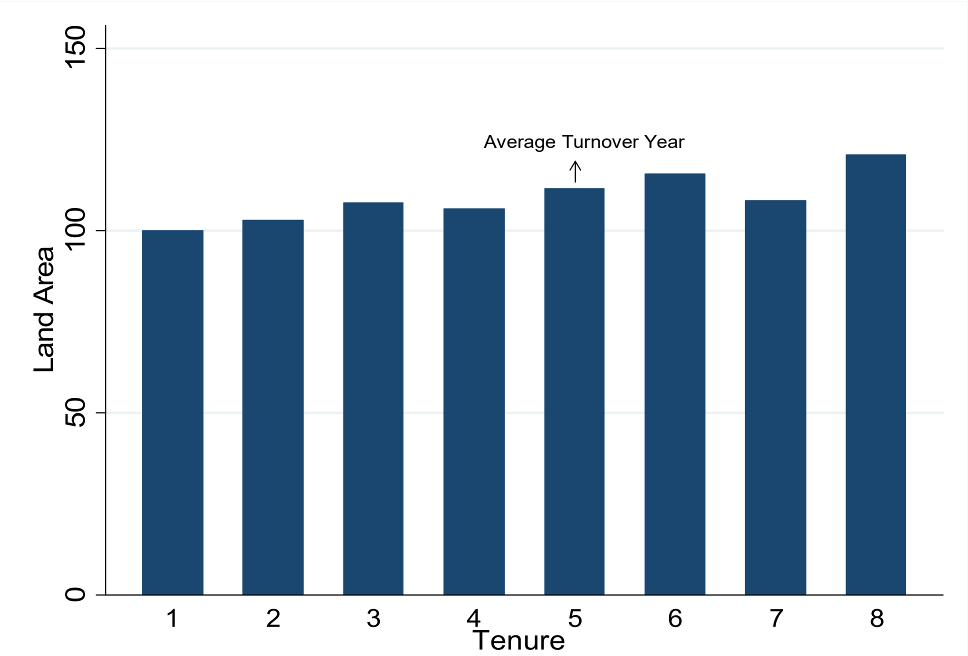
Online Appendix

Go with the Politician

Yongwei Nian
Chunyang Wang

Appendix A: Figures and Tables

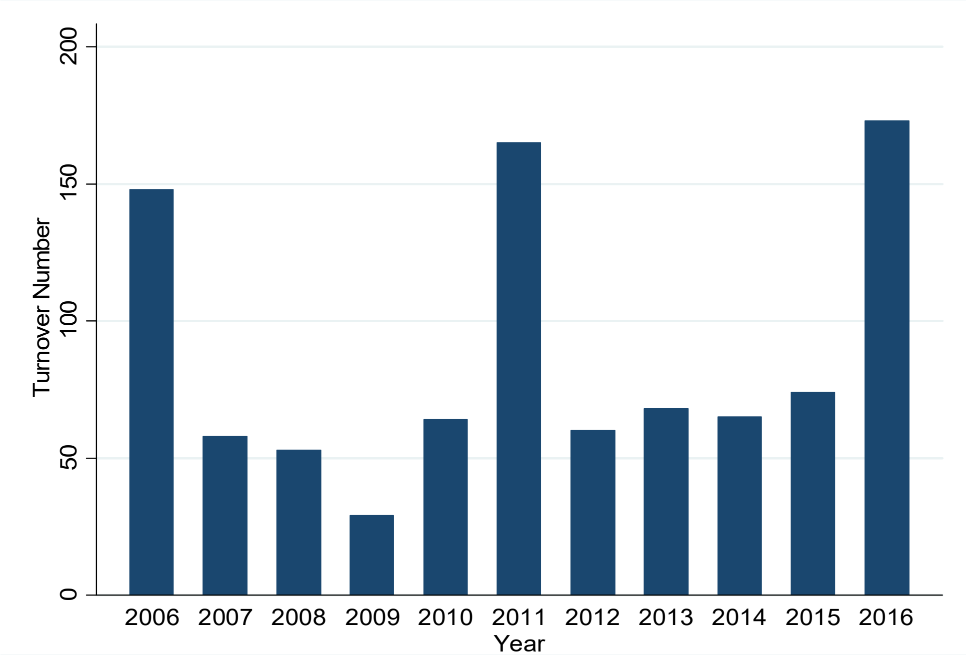
Figure A1: Land Area by Leader's Tenure



Sources: Land China (2016), CSMAR (2017), and authors' own calculation.

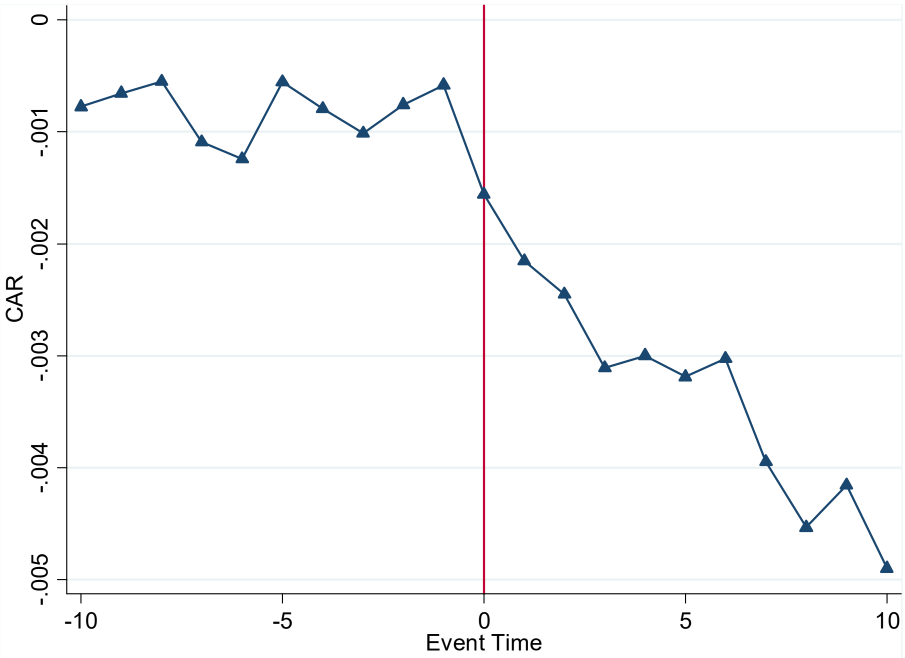
Notes: This figure displays land area sold by prefecture leader's tenure. Total land area sold in leader's first year is normalized to 100.

Figure A2: Leader Turnover Number Distribution by Year



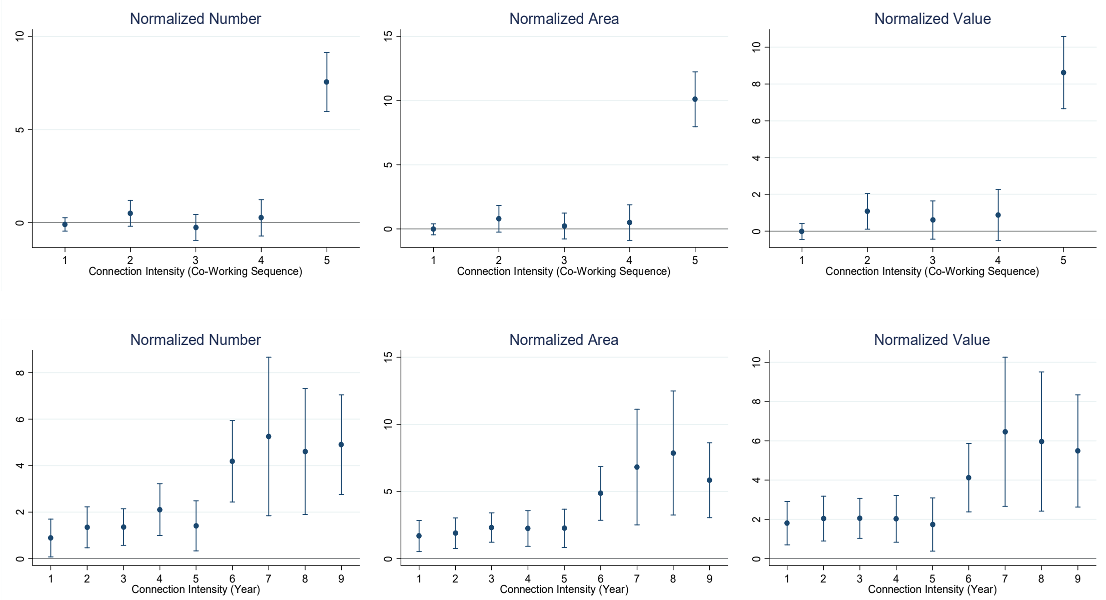
Sources: CSMAR (2017) and authors' own calculation.

Figure A3: Average Cumulative Abnormal Return around Politician Turnover



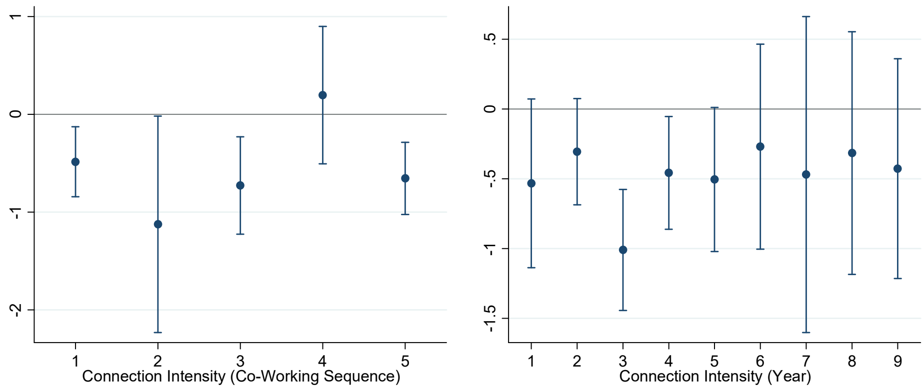
Notes: This figure proves the unexpectedness of politician turnover by showing the insignificant market reaction preceding politician turnover. The triangles denote average cumulative abnormal return (CAR) of firms headquartered in a prefecture around the leave of a local leader, estimated using the Fama-French three-factor model. The day on which a politician's leave is announced is normalized as 0. The estimation window is [-140, -20].

Figure A4: Intensity Effects on Land Purchases without Parametric Restrictions



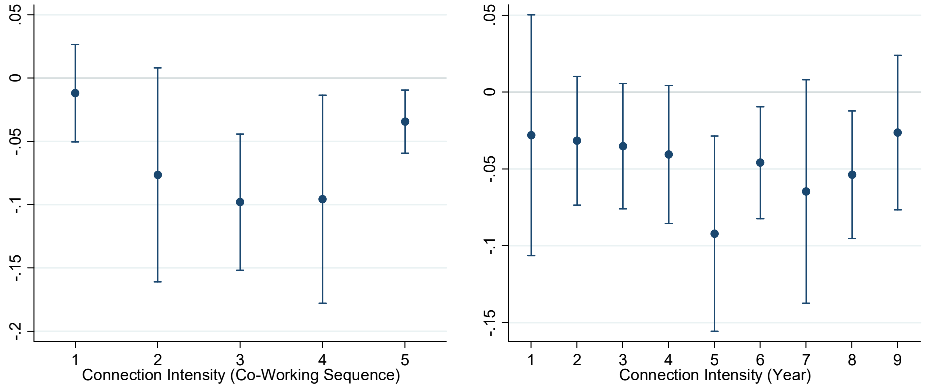
Notes: This figure shows the estimated effect of the intensity of connection between a firm and a politician in a prefecture-year on the firm's land purchases in that prefecture-year without parametric restrictions. Each subfigure is created by estimating a variant of Equation (1), with Connection replaced by a set of separate dummies for different values of Connection Intensity (0 is omitted), and plotting the coefficients and associated 95% standard errors on the dummies. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection Intensity (Co-Working Sequence) takes values from 0 to 5. Five indicates the closest case where a firm's headquarters is located in a local leader's most recent work prefecture, immediately before the leader's move. One indicates the farthest case where a firm's headquarters is located in a local leader's initial work prefecture. Connection Intensity (Year) denotes how many years a local leader previously worked in the firm's headquarters prefecture.

Figure A5: Intensity Effects on Land Price without Parametric Restrictions



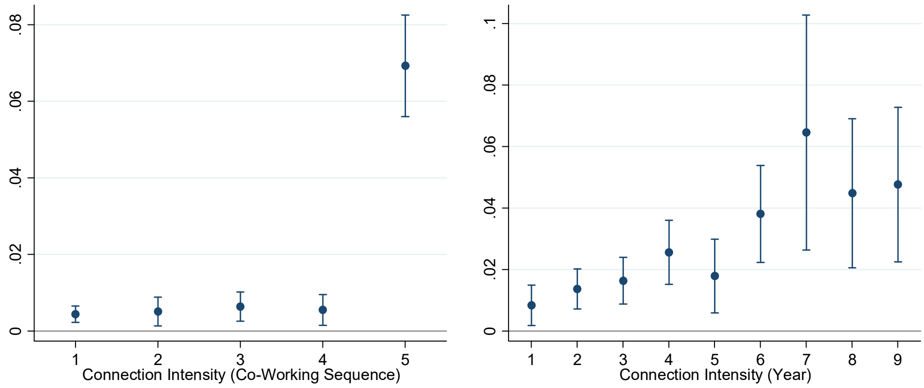
Notes: This figure shows the estimated effect of the intensity of connection on land parcel price a firm pays without parametric restrictions. Each subfigure is created by estimating a variant of the regression in Column (1) of Table 4, with Connection replaced by a set of separate dummies for different values of Connection Intensity (0 is omitted), and plotting the coefficients and associated 95% standard errors on the dummies. The dependent variable is the unit price of a land parcel over average unit price of other similar land parcels (same type and sold in the same year) within a 2 km radius of that land parcel. Connection Intensity (Co-Working Sequence) takes values from 0 to 5. Five indicates the closest case where a firm's headquarters is located in a local leader's most recent work prefecture, immediately before the leader's move. One indicates the farthest case where a firm's headquarters is located in a local leader's initial work prefecture. Connection Intensity (Year) denotes how many years a local leader previously worked in the firm's headquarters prefecture.

Figure A6: Intensity Effects on Land Usage Efficiency without Parametric Restrictions



Notes: This figure shows the estimated effect of the intensity of connection on land usage efficiency without parametric restrictions. Each subfigure is created by estimating a variant of the regression in Column (1) of Table 5, with Connection replaced by a set of separate dummies for different values of Connection Intensity (0 is omitted), and plotting the coefficients and associated 95% standard errors on the interaction between Post Deal and these dummies. The dependent variable is the light intensity of a parcel sold to a firm in a prefecture-year divided by the average light intensity of all parcels sold to all firms in that prefecture-year. Connection Intensity (Co-Working Sequence) takes values from 0 to 5. Five indicates the closest case where a firm's headquarters is located in a local leader's most recent work prefecture, immediately before the leader's move. One indicates the farthest case where a firm's headquarters is located in a local leader's initial work prefecture. Connection Intensity (Year) denotes how many years a local leader previously worked in the firm's headquarters prefecture.

Figure A7: Intensity Effects on Firm Entry without Parametric Restrictions



Notes: This figure shows the estimated effect of the intensity of connection on firm entry without parametric restrictions. Each subfigure is created by estimating a variant of the regression in Column (1) of Table A10, with Connection replaced by a set of separate dummies for different values of Connection Intensity (0 is omitted), and plotting the coefficients and associated 95% standard errors on these dummies. The dependent variable denotes firm entry in a prefecture, a dummy variable which takes the value of 1 after a firm purchases a land parcel until politician turnover in that prefecture, and 0 otherwise. Connection Intensity (Co-Working Sequence) takes values from 0 to 5. Five indicates the closest case where a firm's headquarters is located in a local leader's most recent work prefecture, immediately before the leader's move. One indicates the farthest case where a firm's headquarters is located in a local leader's initial work prefecture. Connection Intensity (Year) denotes how many years a local leader previously worked in the firm's headquarters prefecture.

Table A1: Summary Statistics of Main Variables

Variable	Observations	Mean	Standard Deviation
Normalized Number	9,547,267	1.00	36.21
Normalized Area	9,547,267	1.00	41.42
Normalized Value	9,451,113	1.00	42.60
Connection	9,547,267	0.03	0.18
Connection Intensity (Co-Working Sequence)	9,547,267	0.11	0.65
Connection Intensity (Year)	9,547,267	0.15	0.86
Unit Price / Average Unit Price within a 2 KM Radius	13,227	1.00	1.92
Normalized Light Intensity	262,119	1.00	0.75

Notes: Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in a local leader's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Connection Intensity (Co-Working Sequence) takes values from 0 to 5. Five indicates the closest case where a firm's headquarters is located in a local leader's most recent work prefecture, immediately before the leader's move. One indicates the farthest case where a firm's headquarters is located in a local leader's initial work prefecture. Connection Intensity (Year) denotes how many years a local leader previously worked in the firm's headquarters prefecture. Unit Price / Average Unit Price within a 2 KM Radius is the unit price of a land parcel over the average unit price of other similar land parcels (same type and sold in the same year) within a 2 km radius of that land parcel. Normalized Light Intensity is the light intensity of a parcel sold to a firm in a prefecture-year divided by the average light intensity of all parcels sold to all firms in that prefecture-year.

Table A2: Summary Statistics of Land Data Obtained by Publicly Listed Firms

Category	Number
Number of land parcels obtained by public firms	43,773
-Industrial land	19,252
-Residential land	7,196
-Commercial land	11,202
-Other land	6,123
Number of public firms that have obtained at least one land parcel	2,188
Average number of land parcels per land buyer	20
Number of land parcels obtained by non-local public firms	38,259

Sources: Land China (2016), WIND (2016), and authors' own calculation.

Table A3: Turnover Frequency by Province from 2006 to 2016

Province Name	GDP Per Capita in 2016 (*yuan*)	Turnover Number
Anhui	39,561	50
Beijing	118,198	4
Chongqing	58,502	4
Fujian	74,707	34
Gansu	27,643	47
Guangdong	74,016	69
Guangxi	38,027	22
Guizhou	33,246	33
Hainan	44,347	14
Hebei	43,062	52
Heilongjiang	40,432	28
Henan	42,575	42
Hubei	55,665	48
Hunan	46,382	38
Inner Mongolia	72,064	46
Jiangsu	96,887	28
Jiangxi	40,400	61
Jilin	53,868	36
Liaoning	50,791	47
Ningxia	47,194	14
Qinghai	43,531	23
Shaanxi	51,015	23
Shandong	68,733	48
Shanghai	116,562	4
Shanxi	35,532	30
Sichuan	40,003	57
Tianjin	115,053	5
Tibet	35,184	16
Xinjiang	40,564	36
Yunnan	31,093	54
Zhejiang	84,916	36

Sources: CSMAR (2017), National Bureau of Statistics of China (2017), and authors' own calculation.

Table A4: Group Difference between Connected and Unconnected Firms

	(1) Connection=0	(2) Connection=1	(3) Difference	(4) Conditional Difference
TFP	0.0023	0.0077	0.0054	0.0057
	(0.0001)	(0.0006)	(0.0006)	(0.0014)
ROA	0.0370	0.0389	0.0020	0.0020
	(0.00002)	(0.0001)	(0.0001)	(0.0004)
ROE	0.0653	0.0721	0.0068	0.0076
	(0.0001)	(0.0003)	(0.0003)	(0.0007)
Firm age	15.3458	15.2328	-0.1130	-0.3265
	(0.0019)	(0.0103)	(0.0105)	(0.0794)
SOE	0.4744	0.5161	0.0416	0.0517
	(0.0002)	(0.0011)	(0.0011)	(0.0065)
Debt / Assets	0.2117	0.1986	-0.0131	-0.0106
	(0.0001)	(0.0004)	(0.0004)	(0.0013)
ln(Assets)	21.8377	22.0275	0.1898	0.1697
	(0.0005)	(0.0030)	(0.0027)	(0.0155)

Notes: This table reports the difference and conditional difference between connected and unconnected firms. These examined characteristics are key ones in existing literature. Columns (1) and (2) report the means and associated standard errors (in parentheses) for connected and unconnected firms respectively. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Column (3) reports the mean differences and associated standard errors (in parentheses) for these two groups. Column (4) reports the conditional mean differences and associated standard errors (in parentheses) for these two groups conditional on prefecture and year fixed effects.

Table A5: Capital Prefecture vs Non-Capital Prefectures for Provincial Leaders

	(1)	(2)	(3)	(4)	(5)
Dependent variable	Normalized Value				
Connection	3.00				
	(0.31)				
Connection (Prefecture)		3.64			3.61
		(0.36)			(0.36)
Connection (Province, Capital Prefecture)			1.51		1.18
			(0.47)		(0.48)
Connection (Province, Non-Capital Prefectures)				-0.01	0.05
				(0.08)	(0.07)
Prefecture-year FEs	Yes	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes	Yes	Yes
p(Capital Prefecture=Non-Capital Prefectures)					0.02
Observations	9,451,113	9,451,113	9,451,113	9,451,113	9,451,113
R-squared	0.06	0.06	0.06	0.06	0.06

Notes: This table compares the effects of connection in capital prefectures to the effects of connection in non-capital prefectures for provincial leaders. Normalized Value denotes the value of land parcels bought by a firm in a prefecture-year divided by the average value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Connection (Prefecture) denotes whether a firm's headquarters is in a prefectural politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Connection (Province, Capital Prefecture) denotes whether a firm's headquarters is in a provincial politician's previous work prefectures and is equal to 1 in that politician's governing province's capital prefecture if yes, and 0 otherwise. Connection (Province, Non-Capital Prefecture) denotes whether a firm's headquarters is in a provincial politician's previous work prefectures and is equal to 1 in that politician's governing province's non-capital prefectures if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A6: Alternative Clustering Strategies

Dependent variable	(1) Normalized Number	(2) Normalized Area	(3) Normalized Value	(4) Normalized Number	(5) Normalized Area	(6) Normalized Value
	Two-way Clustering: Prefecture and Year			Provincial-level Clustering		
Connection	2.37	3.36	3.00	2.37	3.36	3.00
	(0.32)	(0.43)	(0.42)	(0.46)	(0.54)	(0.49)
Prefecture-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes	Yes	Yes	Yes
Wild bootstrap *p*	0.00	0.00	0.00	0.00	0.00	0.00
Observations	9,547,267	9,547,267	9,451,113	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06	0.19	0.06	0.06
Number of clusters	326 and 11	326 and 11	326 and 11	27	27	27

Notes: This table reports the estimated effect of whether a firm and a politician in a prefecture-year have a connection on the firm's land purchases in that prefecture-year, using alternative clustering strategies. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at various levels are reported in parentheses. Columns (1)-(3) use two-way clustering at the prefecture and year level. Columns (4)-(6) use provincial-level clustering. To alleviate the concern of small number of clusters, especially at the province level, we also try the wild bootstrap approach suggested by Cameron, Gelbach, and Miller (2008). We use the procedure developed by Roodman et al. (2019) to implement the wild bootstrap and report the *p*-values.

Table A7: Economic Geography - Adjacent Prefectures

	(1)	(2)	(3)
Dependent variable		Normalized Value	
Connection	3.00		3.00
	(0.31)		(0.31)
Connection (Adjacent Prefectures)		-0.20	-0.35
		(0.68)	(0.67)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
p(Baseline= Adjacent)			0.00
Observations	9,451,113	9,451,113	9,451,113
R-squared	0.06	0.06	0.06

Notes: This table compares the land purchases of a connected firm to the land purchases of an unconnected firm headquartered in a prefecture adjacent to the connected firm's headquarters prefecture. Normalized Value denotes the value of land parcels bought by a firm in a prefecture-year divided by the average value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Connection (Adjacent Prefectures) denotes whether an unconnected firm's headquarters prefecture is adjacent to a connected firm's headquarters prefecture and is equal to 1 if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A8: Connection and Land Usage Efficiency - Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Normalized Light Intensity					
	Size > 0.02 Square KM (Median Size)	Size > 0.5 Square KM	Size > 1 Square KM	Distance > 26 KM (Median Distance)	Distance > 50 KM	Grid Cells with One Parcel
Post Deal	-0.05	-0.03	-0.03	-0.02	-0.02	-0.05
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Post Deal × Connection	-0.05	-0.05	-0.05	-0.05	-0.06	-0.06
	(0.01)	(0.02)	(0.03)	(0.01)	(0.02)	(0.01)
Land parcel FEs	Yes	Yes	Yes	Yes	Yes	Yes
Prefecture-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	140,492	25,509	23,221	147,543	99,154	118,580
R-squared	0.93	0.94	0.94	0.92	0.92	0.93

Notes: This table reports the robustness of the estimated effect of Connection on land usage efficiency. The unit of analysis is land parcel. Normalized Light Intensity is the light intensity of a parcel sold to a firm in a prefecture-year divided by the average light intensity of all parcels sold to all firms in that prefecture-year. Post Deal is a dummy variable, taking the value of 1 after a land parcel's sale, and 0 otherwise. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Column (1) keeps only land parcels larger than the sample median which is 0.02 square km. Column (2) keeps only land parcels larger than 0.5 square km. Column (3) keeps only land parcels larger than 1 square km. Column (4) keeps only land parcels more than 26 km away from the prefecture center, which is the median distance. Column (5) keeps only land parcels more than 50 km away from the prefecture center. Column (6) keeps only grid cells containing one land parcel. All regressions control for prefecture-by-year, firm-by-year, and land parcel fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A9: The High Management Cost Channel for Lower Light Growth

	(1)
Dependent variable	Normalized Light Intensity
Post Deal	-0.06
	(0.01)
Post Deal × Remote	0.00
	(0.02)
Land parcel FEs	Yes
Prefecture-year FEs	Yes
Firm-year FEs	Yes
Observations	262,119
R-squared	0.92

Notes: This table tests whether higher management cost for a firm leads to lower land usage efficiency. The unit of analysis is land parcel. Normalized Light Intensity is the light intensity of a parcel sold to a firm in a prefecture-year divided by the average light intensity of all parcels sold to all firms in that prefecture-year. Post Deal is a dummy variable, taking the value of 1 after a land parcel's sale, and 0 otherwise. Remote is a dummy variable and is equal to 1 for land parcels out of the province where a firm is headquartered, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and land parcel fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A10: Firm Entry

	(1)	(2)	(3)
Dependent variable		Entry	
Connection	0.026	0.003	0.015
	(0.002)	(0.003)	(0.003)
Connection Intensity (Co-Working Sequence)		0.010	
		(0.001)	
Connection Intensity (Year)			0.004
			(0.001)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	9,547,428	9,547,428	9,547,428
R-squared	0.118	0.124	0.119
Mean of dependent variable	0.003	0.003	0.003

Notes: This table reports the estimated effect of Connection or Connection Intensity on firm entry in terms of land purchase in a prefecture. Entry denotes firm entry in a prefecture, a dummy variable which takes the value of 1 after a firm purchases a land parcel until politician turnover in that prefecture, and 0 otherwise. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, 0 otherwise. Connection Intensity (Co-Working Sequence) takes values from 0 to 5. Five indicates the closest case where a firm's headquarters is located in a local leader's most recent work prefecture, immediately before the leader's move. One indicates the farthest case where a firm's headquarters is located in a local leader's initial work prefecture. Connection Intensity (Year) denotes how many years a local leader previously worked in the firm's headquarters prefecture. All regressions control for prefecture-by-year, firm-by-year, and origin-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A11: Alternative Fixed Effects

	(1)	(2)	(3)
Dependent variable	Normalized Number	Normalized Area	Normalized Value
Panel A			
Connection	2.30	3.24	2.90
	(0.24)	(0.32)	(0.30)
Prefecture FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.17	0.05	0.05
Panel B			
Connection	2.30	3.26	2.91
	(0.25)	(0.32)	(0.30)
Prefecture FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06
Panel C			
Connection	2.36	3.34	2.98
	(0.25)	(0.33)	(0.31)
Prefecture-year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.17	0.05	0.05
Panel D			
Connection	2.43	3.46	3.08
	(0.25)	(0.33)	(0.31)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province-year FEs	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06

Notes: This table reports the estimated effect of whether a firm and a politician in a prefecture-year have a connection on the firm's land purchases in that prefecture-year, using baseline regressions with alternative sets of fixed effects. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A12: Prefectural Leaders vs Provincial Leaders

	(1)	(2)	(3)	(4)
Dependent variable		Normalized Value		
Connection	3.00			
	(0.31)			
Connection (Prefecture)		3.64		3.61
		(0.36)		(0.36)
Connection (Province)			1.51	1.17
			(0.47)	(0.48)
Prefecture-year FEs	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes	Yes
p(Prefecture=Province)				0.00
Observations	9,451,113	9,451,113	9,451,113	9,451,113
R-squared	0.06	0.06	0.06	0.06

Notes: This table compares the effects of connection with prefectural politicians to the effects of connection with provincial politicians. Normalized Value denotes the value of land parcels bought by a firm in a prefecture-year divided by the average value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. The politician could be a prefectural leader or a provincial leader. Connection (Prefecture) denotes whether a firm's headquarters is in a prefectural politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Connection (Province) denotes whether a firm's headquarters is in a provincial politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A13: Baseline Results without Politicians Who Have Worked in Beijing

Dependent variable	(1) Normalized Number	(2) Normalized Area	(3) Normalized Value
Connection	2.52	3.52	3.13
	(0.26)	(0.35)	(0.33)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06

Notes: This table reports the estimated effect of whether a firm and a politician in a prefecture-year have a connection on the firm's land purchases in that prefecture-year, for the subsample of politicians without experiencing working in Beijing where the central government is located. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A14: Baseline Results without Firms Headquartered in Beijing or Shanghai

Dependent variable	(1) Normalized Number	(2) Normalized Area	(3) Normalized Value
Connection	3.81	5.31	4.69
	(0.33)	(0.46)	(0.44)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	7,914,186	7,914,186	7,834,484
R-squared	0.02	0.01	0.02

Notes: This table reports the estimated effect of whether a firm and a politician in a prefecture-year have a connection on the firm's land purchases in that prefecture-year, for the subsample of firms without their headquarters in Beijing or Shanghai. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A15: Baseline Regression by Land Types

Dependent variable	(1) Normalized Number	(2) Normalized Area	(3) Normalized Value	(4) Normalized Number	(5) Normalized Area	(6) Normalized Value
	Industrial Land			Commercial/Residential Land		
Connection	1.57	2.22	1.89	3.71	4.12	4.38
	(0.24)	(0.29)	(0.27)	(0.34)	(0.39)	(0.40)
Prefecture-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,934,272	8,934,272	8,754,318	7,526,887	7,526,887	7,412,008
R-squared	0.25	0.12	0.11	0.03	0.02	0.02

Notes: This table reports the estimated effect of whether a firm and a politician in a prefecture-year have a connection on the firm's land purchases in that prefecture-year, for a subsample with only industrial land in Columns (1)-(3) and a subsample with only commercial and residential land in Columns (4)-(6). Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A16: Heterogeneity by Pre-Connection Firm Performance

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Normalized Number	Normalized Area	Normalized Value	Normalized Number	Normalized Area	Normalized Value
Connection	2.37	3.56	3.24	2.30	3.67	3.27
	(0.42)	(0.53)	(0.52)	(0.33)	(0.48)	(0.44)
Connection	-0.07	-0.45	-0.49			
× High Sales Growth	(0.41)	(0.54)	(0.51)			
Connection				0.03	-0.64	-0.56
× High ROA				(0.39)	(0.56)	(0.52)
Prefecture-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,405,735	9,405,735	9,311,005	9,405,735	9,405,735	9,311,005
R-squared	0.19	0.06	0.06	0.19	0.06	0.06

Notes: This table tests the heterogeneity effects of Connection by pre-connection firm performance. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. High Sales Growth is a dummy variable denoting whether the firm's sales growth was higher than the median when the politician was working at the firm's headquarters prefecture and is equal to 1 if yes, and 0 otherwise. High ROA is a dummy variable denoting whether the firm's ROA was higher than the median when the politician was working at the firm's headquarters prefecture and is equal to 1 if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A17: Political Incentives

Dependent variable	(1) Normalized Number	(2) Normalized Area	(3) Normalized Value	(4) Normalized Number	(5) Normalized Area	(6) Normalized Value
Connection	2.35	3.34	3.00	2.36	3.36	3.00
	(0.25)	(0.33)	(0.31)	(0.25)	(0.33)	(0.31)
Connection × Age	0.07	0.05	0.02			
	(0.05)	(0.06)	(0.06)			
Connection × Term Length				0.10	0.00	0.02
				(0.16)	(0.18)	(0.16)
Prefecture-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,547,162	9,547,162	9,451,008	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06	0.19	0.06	0.06

Notes: This table tests the role of political incentives, by comparing politicians with different ages and term lengths. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Connection denotes whether a firm's headquarters is in a politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Age and Term Length denote the age and the term length of the connected politician, respectively. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A18: Market Reactions to Corruption Investigation

	(1)	(2)	(3)	(4)
Dependent variable	CAR[-1,1]	CAR[-10,10]	CAR[-1,1]	CAR[-10,10]
Connection	-0.006	-0.024	-0.005	-0.023
	(0.002)	(0.007)	(0.003)	(0.009)
Firm-investigation year FEs	No	No	Yes	Yes
Observations	6,237	6,236	5,012	5,010
R-squared	0.003	0.005	0.383	0.462

Notes: This table reports the estimated effects of politically connected firms' cumulative abnormal returns (CARs) around the corruption investigations of the connected politicians. The CARs are calculated using the Fama-French three factor model with an estimation window of [-140, -20]. Connection denotes whether a firm's headquarters is in that politician's previous work prefectures and is equal to 1 if yes, and 0 otherwise. Standard errors clustered at the firm level are reported in parentheses.

Table A19: Locally Headquartered Firms

	(1)	(2)	(3)	(4)
Dependent variable	Normalized Number	Normalized Area	Normalized Value	Unit Price / Average Unit Price within a 2 KM Radius
Local	2.45	2.67	1.59	-0.04
	(0.28)	(0.20)	(0.17)	(0.16)
Prefecture-year FEs	Yes	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes	Yes
Observations	9,533,071	9,537,719	9,441,662	13,212
R-squared	0.02	0.06	0.06	0.46

Notes: This table reports the estimated effect of whether a firm is locally headquartered on the land purchases of that firm, and land parcel price that firm pays. In Columns (1)-(3), the dependent variables Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. In Column (4), the dependent variable is the unit price of a land parcel over average unit price of other similar land parcels (same type and sold in the same year) within a 2 km radius of that land parcel. Local denotes whether a firm's headquarters is in that prefecture-year and is equal to 1 if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A20: Hometown as Another Potential Source of Favoritism

	(1)	(2)	(3)
Dependent variable	Normalized Number	Normalized Area	Normalized Value
Hometown	0.00	-0.22	0.01
	(0.44)	(0.55)	(0.52)
Prefecture-year FEs	Yes	Yes	Yes
Firm-year FEs	Yes	Yes	Yes
Origin-destination province FEs	Yes	Yes	Yes
Observations	9,547,267	9,547,267	9,451,113
R-squared	0.19	0.06	0.06

Notes: This table reports the estimated effect of whether a firm's headquarters is in the hometown prefecture of a politician in a prefecture-year on the firm's land purchases in that prefecture-year. Normalized Number, Normalized Area, and Normalized Value denote the number, area, and value of land parcels bought by a firm in a prefecture-year divided by the average number, area, and value of land parcels bought by all firms in that prefecture-year, respectively. Hometown denotes whether a firm's headquarters is in that politician's hometown prefecture and is equal to 1 if yes, and 0 otherwise. All regressions control for prefecture-by-year, firm-by-year, and origin province-by-destination province fixed effects. Robust standard errors clustered at the prefecture level are reported in parentheses.

Table A21: The Effect of "Go with the Politician" on Politician's Promotion and Prefectural GDP

	(1)	(2)	(3)	(4)
Dependent variable	Promotion$_{t+1}$		ln(GDP)$_{t+1}$	
Fraction of Land Value to Connected Firms	-0.04	-0.03	-0.00	-0.00
	(0.02)	(0.03)	(0.01)	(0.01)
GDP Growth Rate		0.25		0.84
		(0.16)		(0.05)
ln(Population)		-0.10		0.43
		(0.20)		(0.07)
Prefecture FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Observations	3,370	2,864	2,875	2,863
R-squared	0.52	0.52	0.98	0.99

Notes: This table reports the estimated effect of whether a higher fraction of land sales to connected firms in a prefecture-year might lead to that prefecture leader's promotion or lead to higher GDP next year. The dependent variable is Promotion$_{t+1}$ in Column (1) and (2), a dummy variable taking the value of 1 if that prefecture leader is promoted next year, and 0 otherwise. In Column (3) and (4), the dependent variable is ln(GDP), denoting the logarithm of GDP in that prefecture next year. GDP Growth Rate is defined as the change in the logarithm of GDP from year t-1 to year t. Robust standard errors clustered at the prefecture level are reported in parentheses.

Appendix B: TFP Calculation

TFP is total factor productivity, calculated as the difference between actual and predicted output of a firm. Specifically, we estimate the following log-linear Cobb-Douglas production function, following the literature (Bertrand and Mullainathan, 2003; Giroud, 2013):

$$y_{it} = \alpha + \beta_1 k_{it} + \beta_2 l_{it} + \beta_3 m_{it} + \varepsilon_{it}$$

where i and t denote firm and year, respectively. y_{it} is the logarithm of sales, k_{it} is the logarithm of total assets, l_{it} is the logarithm of the number of employees, and m_{it} is the logarithm of expenditure for material inputs. TFP is captured by the residual. The estimation is conducted for each industry-year separately to account for varying factor intensities.

Appendix References:

Bertrand, Marianne, and Sendhil Mullainathan. 2003. "Enjoying the Quiet Life? Corporate Governance and Managerial Preferences." *Journal of Political Economy* 111 (5): 1043-1075.

Giroud, Xavier. 2013. "Proximity and Investment: Evidence from Plant-level Data." *The Quarterly Journal of Economics* 128 (2): 861-915.