The Environmental and Economic Consequences of Internalizing Border Spillovers

(Preliminary)

Shaoda Wang* Zenan Wang†

September 29, 2020

Abstract

This paper studies how centralized decision-making can help local governments internalize regional environmental spillovers, and investigates the associated economic and welfare consequences. Utilizing novel firm-level geocoded emission and production panel datasets, and exploiting more than 3000 cases of township mergers in China, we find that as township mergers eliminate the borders between neighboring townships, the negative externalities of polluting firms located on these borders are suddenly internalized by the new jurisdiction. As a result, these firms spend more effort on emission abatement, which leads to lower emissions, as well as lower output and profit levels. Further analysis suggests that household welfare improves with the internalization of border spillovers, as reflected by increased residential land price around the merging borders.

Keywords: Decentralization, Pollution, Township Merger

*University of Chicago. Email: shaoda@uchicago.edu.
†UC Berkeley. Email: zenan.wang@berkeley.edu.

1 Introduction

The issue of "centralization" versus "decentralization" for governance is long-standing in social sciences (Tiebout, 1956; Oates, 1972; Prud'homme, 1995; Fisman and Gatti, 2002; Foster and Rosenzweig, 2002). In this discussion, a key argument against decentralization is that, under decentralized decision-making, the lack of coordination among local jurisdictions would make them unable to properly internalize regional spillovers and spatial externalities, which would lead to Pareto-inefficient outcomes (Oates, 1972; Wildasin, 1991; Saavedra, 2000; Fredriksson and Millimet, 2002; Wilson, 1999).

The classic example to support this argument is the phenomenon known as "polluting your neighbor": the ambient pollutants emitted by firms located near jurisdictional boundaries affect both the host and the neighboring jurisdictions, while the associated economic benefits (e.g., tax revenue) are disproportionately enjoyed by the host jurisdiction. As a result, individual jurisdictions lack the incentives to fully internalize their environmental externalities affecting the neighboring townships, and will likely impose relatively lenient environmental regulation standards on those border polluting firms (Burgess et al., 2012; Gray and Shadbegian, 2004; Helland and Whitford, 2003; Fredriksson and Millimet, 2002; Sigman, 2002, 2005; Konisky and Woods, 2010).

Despite its importance, causally evaluating the mechanisms and consequences of internalizing border environmental spillovers remains challenging. Since polluting firms endogenously decide whether to locate on the border (or in the center) of a jurisdiction, if we simply estimate the difference in "emission and production activities" between border and central polluters, such a comparison would reflect not only the local government's differential internalization of environmental externalities, but also many other confounding differences between the border and central polluters which determined their differential locations in the first place. In addition, researchers also face a data challenge, as firm-level differences in "regulatory burdens" and "responses to regulation" can rarely be precisely measured within a narrow geographic region.

To fill in this gap in knowledge, this paper investigates how within-firm changes in "relative location to township border," created by the staggered roll-out of more than 3000 township mergers between 2002 and 2008, affect firm-level emission and production outcomes. Specifically, we exploit the fact that when two townships merge, a firm originally located on the merging border will suddenly lie in the center of the newly merged township, and such an abrupt switch from being "border" to being "central" would sharply increase the township government's internalization of that firm's pollution externalities. By comparing the same firm under "border status" and "central status" before and after a merger, we thus can causally identify the environmental and economic impacts of internalizing border spillovers.

Leveraging novel firm-level geocoded panel datasets with detailed information on production and emission activities, we find that, when a polluting firm suddenly switches from "township border" to "township center" following a merger, its compliance with environmental regulation improves substantially, as reflected by a 9.7% drop in SO_2 emissions, a 3.4% decrease in coal boiler ownership, and a 12.5% increase in abatement investment. To cope with greater enforcement of environmental standards, polluters "moving" toward township centers on average suffer from a 4.8% reduction in industrial output, a 8.4% reduction in net profit, and a 7.8% reduction in export volume. In stark contrast, when non-polluting firms "move" from township border to township center following a merger, there are no detectable changes in their production and emission activities.

To understand the overall welfare impacts of internalizing border environmental spillovers, we also analyze another geocoded dataset on the universe of land auctions in China. We find that the disappearance of a township border (due to township merger) leads to more residential projects being developed near the merging-border area, and also causes an increase in local residential land price. Both results suggest an improvement in household welfare. Further analysis suggests that these effects are predominantly driven by merging townships with at least one polluting firm located near the merging border, suggesting that the internalization of environmental spillovers, rather than other forms of spatial externalities, is driving the improvement in household welfare.

Combining the findings on manufacturing output loss, ambient emission reduction, and residential land value gain, we conduct a series of simple back-of-the-envelope calculations to shed light on public policy. We calculate that, during our sample period, the internalization of border spillovers (associated with township mergers) led to a net welfare gain of more than 140 billion RMB; and if all township governments in China were to properly internalize their border pollution externalities, there will be an overall increase in welfare of more than 460 billion RMB. Our estimates also suggest that a 10% reduction in SO_2 emissions would cost the polluting industries a 7.1% loss in output, implying that China's ambitious national emission abatement plan during our sample period could have caused an output loss of more than 200 billion RMB per year.

This paper relates to four strands of literature. First and foremost, it adds to the work on the strategic interactions among local governments, which falls more broadly into the literature on decentralization and federalism. Using "tax competition" as an example, Oates (1972) first showed that the result of inter-governmental competition is that each decentralized jurisdiction fails to internalize regional spillovers, which leads to distortions in local policies (from a central planner's perspective). This idea was later formalized (Zodrow and Mieszkowski, 1986), and extended to various other contexts, including income redistribution (Wildasin, 1991), government expenditure (Wilson, 1999), and welfare transfers (Saavedra, 2000). In this literature, the "polluting your neighbor" phenomenon has been documented extensively by previous research (Burgess et al., 2012; Gray and Shadbegian, 2004; Helland and Whitford, 2003; Fredriksson and Millimet, 2002; Sigman, 2002, 2005; Konisky and Woods, 2010). A particularly relevant paper is Lipscomb and Mobarak (2017), which shows that, when the split of two counties reduces the distance between a water monitoring station and the county boundary, the water quality reading of that station worsens significantly. Our paper supports the findings of Lipscomb and Mobarak (2017) in the context of China, and also complements it in two important ways: (1) our empirical setting allows us to pin down how the internalization of spillovers affects firm-level regulatory burdens and abatement efforts; (2) our data enables us to quantify the economic costs and residential welfare gains associated with the internalization of pollution spillovers.[1]

Second, our paper adds to the growing literature on the political economy of environmental policies (List and Sturm, 2006; Burgess et al., 2012; Kahn et al., 2015; Lipscomb and Mobarak, 2017; Jia, 2017; He et al., 2020). Consistent with He

[1] Two previous studies that investigated China's water quality issue are also relevant to this study: Kahn et al. (2015) investigates water pollution abatement across provincial boundaries and finds that tighter environmental regulation by the central government incentivized local officials to reduce border pollution according to specific criteria; Cai et al. (2016) finds that provincial governments responded to the pollution reduction mandates by shifting their enforcement efforts away from the downstream counties. Our paper complements these studies by exploiting only with-firm variation in "border" vs. "central," and quantifying the environmental, economic, and welfare consequences of internalizing border spillovers in the same context.

et al. (2020), we find that under the same nominal environmental regulation standard, the actual level of regulation enforcement could vary tremendously across spatially adjacent firms even within narrowly-defined industries. Specifically, we find that a polluting firm will face much tighter environmental regulatory enforcement when its negative externalities are fully internalized by the local government. Our findings suggest that the state-business relationship affects the effectiveness of environmental policies in important ways, and needs to be taken into account when designing optimal regulation programs.

Third, this study contributes to the discussion about the economic consequences of environmental regulation. While there exists a large empirical literature on environmental regulation in the United States (Becker and Henderson, 2000; List et al., 2003; Ryan, 2012; Greenstone, 2002; Reed Walker, 2011), much less is known about the environment-economy tradeoff in the developing world.[2] In this paper, we add to the existing literature by estimating the average pollution abatement cost for the entire Chinese manufacturing sector: exploiting within-firm variation in regulatory stringency, we find that a 10% reduction in SO2 would cause a 7.1% drop in industrial output for the polluting industries, which amounts to more than 200 billion RMB per year.

Fourth, our paper also adds to a growing literature on the socio-economic consequences of jurisdictional boundary changes, such as the impacts on market access and development (Redding and Sturm, 2008), regional population (Davis and Weinstein, 2002), local conflicts (Bazzi and Gudgeon, 2016), and water pollution (Lipscomb and Mobarak, 2017). To the best of our knowledge, this paper is the first to link jurisdictional boundary changes to precisely geocoded firm-level panel datasets, which creates a unique empirical setting to investigate the dynamics of state-business relations through the lens of environmental regulatory enforcement. Specifically, we document that the disappearance of a township border changes certain "border firms" into "central firms" for the local government, which leads to a significant increase in the regulatory burdens imposed on these firms.

The remainder of this paper is organized as follows. In Section 2, we briefly discuss the role of township governments in the Chinese governance system, introduce the township consolidation program, and explain the intuition of our

[2] One exception is He et al. (2020), which estimates that China's water regulation costs a 2.7% reduction in manufacturing TFP for a 10% abatement of Chemical Oxygen Demand emissions.

research design. In Section 3, we lay out a simple conceptual framework to guide the subsequent empirical analysis. In Section 4, we explain the datasets used in this project. In Section 5, we test our theoretical predictions regarding firm behaviors. In Section 6, we test our theoretical predictions regarding residential land sales. Section 7 discusses the policy and welfare implications of our findings. Section 8 concludes.

2 Background and Research Design

2.1 The "Township Consolidation" Program in China

China's local governance system has four official tiers: province, prefecture, county, and township.[3] Townships, which are the lowest-level formal administrative units, are comparable in size to small cities in the US: an average township contains approximately 20 villages, has about 30,000 residents, and spans nearly 80 km^2.

Due to mass rural-urban migration in the early 1990s, rural townships in China lost a large fraction of their population (to the urban areas), making the existence of some small townships economically inefficient. Moreover, many township and village enterprises (TVEs) also started to face severe deficits after their initial success in the 1980s, which severely weakened the fiscal condition of the township governments. Therefore, to mitigate the fiscal crisis and realize economies of scale in rural governance, in 1995, the Chinese central government launched a large-scale "township consolidation program (TCP)," which aimed to reduce the number of rural townships from 47,136 to roughly 30,000.

To enforce the TCP in a coordinated way, each year, the central government would assign a specific target to every provincial government, specifying the number of township mergers to be achieved. This target was then divided up and assigned to each county government, which would follow a 4-step process to implement the consolidation: (1) the county government would first propose a plan for consolidation, and consult with the involved townships; (2) upon agreement, the involved townships would then submit a formal application for consolidation to the county government; (3) upon approval, the county government would report the decision to the prefectural government and ask for its approval; (4) the

[3] Villages are "self-governed entities" rather than formal administrative units.

prefectural government would then report its decision to the provincial government, which was the final decision-maker. Upon provincial approval of the merger case, the provincial government would issue a formal notification to the county government outlining the detailed schedule for implementation.

Rich anecdotal evidence suggests that the merger suggestion made by county governments was the most critical step in this decision process: once a county government recommends that two townships merge with each other, such a decision is denied in the subsequent steps. Since the primary goal of the TCP was to reduce the economic inefficiencies caused by maintaining too many underpopulated townships, county governments mostly suggest mergers for neighboring small townships, such that, once they have merged with each other, the newly consolidated township would have a population size that is close to the national average.

The rollout of the TCP is illustrated in Figure 1. As we can see, China had 47,136 townships in 1995, and after 18 years of township consolidation, this number was reduced to 32,929 in 2013. The broad scope of this program, combined with its gradual rollout over a nearly two-decade time span, generates rich variation that can be exploited in our empirical analysis.

2.2 Township Governments and Environmental Regulation

In the late 1990s and early 2000s, after two decades of rapid economic growth, many urban areas in China started to raise their environmental standards substantially. This drove pollution-intensive industrial activities to relocate from urban areas to less populated townships, where thousands of township and village enterprises accounted for more than 45% of national industrial value-added as of 2000. As documented by Tilt (2007), in the early 2000s, township plants were responsible for the lion's share of industrial emissions in China, and the concentration of pollution was especially severe around township boundaries.

As alleviating industrial pollution in townships became an increasingly pressuring issue, the township governments also started to play more important roles in the local enforcement of environmental regulation programs. Despite being the most grassroots units in the bureaucratic hierarchy, township governments have always been pivotal in delivering public services and enforcing national programs.[4] Importantly, when the national policy goal and local vested interest are not perfectly aligned, it has been well-documented that the township government officials can wield substantial *de facto* discretion in local governance by selectively implementing national programs (O'Brien and Li, 1999). This pattern is also salient in the enforcement of environmental regulation programs.

Specifically, during our sample period, China's air monitoring system only had coverage at the prefectural city level, and could not reflect pollution sources from more grassroots areas. Therefore, in order to identify the specific polluting plants that were violating the governmental emission standards, the prefectural- and county-level environmental bureaus had to rely heavily on the decentralized information provided by the township governments. Moreover, to effectively enforce regulatory decisions such as temporary suspensions or forced adoptions of cleaner technologies, the prefectural and county environmental bureaus also needed the endorsement and assistance of the local township governments.

However, despite the township officials' key roles in the decentralized enforcement of environmental regulation, "environmental quality" was not included as part of their "promotion criteria" in this era, which consisted mainly of economic indicators. This creates an obvious principal-agent problem in environmental regulation: the lack of political rewards gives the township officials little incentive to truthfully report and stringently regulate those border polluters, whose output is fully counted as local GDP, while the associated pollution only partly affects the local land values. We formalize this problem in Section 3: when township government officials only care about economic indicators such as local tax revenue and housing price, they will enforce tighter regulations on the "central firms" while letting the "border firms" pollute their neighboring townships.

Such selective implementation of environmental regulation is also consistent with evidence documented in the existing literature. For example, it has been shown that when enforcing environmental regulation, township governments have both the incentive and the leverage to favor certain polluting firms over others by exhibiting double standards (Wang et al., 2008). As also documented by He et al. (2020), while the nominal environmental regulation standard is set at the level of

[4]A famous metaphor often cited in Chinese politics is that "various policies and programs are like thousands of threads, while grassroots governments are the single needle that weaves," which highlights the township governments' critical role in policy implementation.

the central government, local officials are able to target certain "key polluters," exploiting various policy instruments, including temporary shutdowns and compulsory abatement investments. Such differential enforcement of environmental regulations creates significant variation in *de facto* regulatory stringency, even among spatially adjacent firms within the same township-industry cluster.

2.3 Research Design

The intuition for our research design is illustrated in Figure 2. In this very simple case, there are two neighboring townships, A and B, which will eventually be merging with each other. For polluting firms located in a township, we use a circle to illustrate its range of negative externalities, where the yellow shading represents the portion of externalities absorbed by the host township, while the white shade represents the portion absorbed by the neighboring township.

As shown in the upper panel, our research design focuses on two types of border firms: (1) "non-merging border firms": firms that are located close to a border that is not shared by townships A and B, so that part of their externalities are borne by another adjacent township even after the merger between townships A and B; and (2) "merging border firms": firms that are located close to the shared border of townships A and B, so that part of their externalities are borne by the neighboring township only prior to the township merger.

As illustrated in the lower panel, once townships A and B merge with each other, their shared border disappears. This event has differential impacts on the two types of firms: for those non-merging border firms, they remain "border firms," because their externalities are still partially borne by another adjacent township; while for those merging border firms, the disappearance of the merging border would suddenly "move" them to the center of the new township (A+B), which makes them "central firms" whose externalities are fully borne by the new township government.

Our research design will exploit such abrupt within-firm switches between "border" and "central" status induced by township mergers, and will causally identify the impacts of internalizing border spillovers on firms' regulatory burdens, emission and production conditions, as well as adjacent land transaction prices. Specifically, within the same township-industry cluster, we will compare the "break-in-trends" for "merging border firms" to the "break-in-trends" for "non-merging border firms", before and after the merger occurs. When analyzing changes in consumer welfare, we will use "residential land parcel" instead of "manufacturing firm" as the unit of analysis.

3 Conceptual Framework

In this section, we present a highly stylized political economy model to conceptualize the impacts of township merger on environmental regulation enforcement, firm emission and productivity, and local housing price. The predictions of this model will be used to guide the subsequent empirical analysis.

3.1 Setup

There are three types of agents: firms, residents, and township governments. Without loss of generality, we assume there are only two neighboring townships, A and B, and two periods, $t = 1$ and $t = 2$. In $t = 1$, the townships are two separate entities, each making independent decisions to maximize its own objective function. But in $t = 2$, townships A and B are consolidated into one. The border between A and B is erased and the new government makes decisions to maximize the joint objective function for the merged township as a whole. The borders shared with other neighboring townships are assumed to remain unchanged.

Firms

There are firms located in both A and B. Firms produce outputs but generate air/water emissions as by-products, thus facing environmental regulation (r) imposed by the local government. We assume that firms produce homogeneous goods, with a Hicks-neutral continuously differentiable production function $f(K, L)$,[5] where K is capital, and L is labor. The amount of emission produced is $E(f(K, L), K_E)$, a differentiable function of total output $f(K, L)$, and emission abatement capital K_E. Emission grows as total output increases ($E_1 > 0$), and decreases with higher abatement investment ($E_2 < 0$). We assume that there is a diminishing return to abatement equipment ($E_{22} > 0$), while marginal impacts

[5]The production function has diminishing returns to scale: $f_1, f_2 > 0; f_{11}, f_{22} < 0$.

of total output on emissions is increasing ($E_{11} > 0$).[6]

Firms maximize their profits by choosing the optimal level of labor, productive capital, and abatement equipment while taking the local environmental regulation r as given:

$$\begin{aligned} &\max_{K,L,K_E} \pi(K, L, K_E) \\ =& \max_{K,L,K_E} (1-t) \cdot f(K, L) - p_k \cdot (K + K_E) - p_l \cdot L - r \cdot E(f(K, L), K_E), \end{aligned} \tag{3.1}$$

where p_k and p_l are market prices for capital and labor respectively, and t is the general production tax rate decided by the central government. The local environmental regulation r can be thought of as a fine for each unit of emission produced.[7]

Residents

A resident chooses a community to live in based on the local housing price and amenities. Without loss of generality, we define a community i to be the adjacent area surrounding a firm i, and assume firms' emissions affect only the local community.[8] Because jurisdictional boundaries may cut through border communities, we use $\alpha_{ic} \in [0, 1]$ to capture the proportion of community i's land that falls in township c.

Following Greenstone and Gallagher (2008), we use a simple hedonic function to characterize the relationship between environmental amenities and housing price. Since communities located on the shared border of merging townships constitute only a small proportion of the Chinese housing market, we assume that the general equilibrium price for amenities does not change in response to the increased supply of less polluted (merging border) communities.

[6]One way to conceptualize this assumption is that, given the fixed amount of emission abatement equipment, the more output is produced, the less effective is the emission processing.

[7]It is worth noting that, nominally, a township government is not empowered to directly tax the firms in its jurisdiction, but it benefits indirectly from firms' growth through eventual tax-sharing and potential corruption. Therefore, the tax parameter "t" is not literally a tax, but rather a conceptualization of the implicit benefits that township governments enjoy from local firm growth.

[8]According to (Currie et al., 2015), the range of a firm's environmental externalities is about 1 mile, which is consistent with our empirical findings. For simplicity, we assume people living in areas not impacted by firms' emissions will not influence local government's regulation decisions.

Consider a border community that saw an increase in local environmental quality after a township merger. The supply curve of the residential housing in the community is upward sloping in the relatively longer term and demand is downward sloping. This is depicted in Appendix Figure A1 with S and D_1 and equilibrium outcome (P_1, Q_1). When there is an exogenous increase in environmental quality after the merger, the demand curve for residential housing near the improved community shifts out, because the improved amenity will attract more individuals with higher valuation of environmental quality to move in. This is depicted as D_2. This change causes prices to increase to P_2 and quantities to increase to Q_2. The shaded area is the welfare gain from the improved amenities. When the supply is highly inelastic, the change in welfare is roughly proportional to the change in price.

Government

The township government could benefit from both increased firm tax revenues and land surpluses. Specifically, while more production means higher fiscal revenue for the township government, it also leads to deteriorated environmental amenities and thus lower land values. Therefore, when setting the environmental regulation policy (r_i) for each firm i in its jurisdiction, the township government faces a trade-off between the "land market surplus of its constituency" and the "tax revenue from local firms."

Let $\mathcal{I}_c$ be the set of firms in township c, and r_i be the regulation policy for firms i [9]. The government's objective is

$$\max_{\{r_i | i \in \mathcal{I}_c\}} \sum_i \left(\alpha_{ic} Surplus(r_i) + Tax(r_i) \right),$$

where $\alpha_{ic} Surplus(r_i)$ represents the portion of the land market surplus in community i that can benefit the local government and $Tax(r_i)$ represents tax revenue from all firms located within its boundary. Assuming that firm-specific regulation stringency r_i affects only firm i and its surrounding community i but not other firms and communities, the government's problem can be simplified as maximizing the total welfare ($\alpha Surplus + Tax$) of each firm-community unit separately,

[9] We assume that there are sufficient distances between firms so that communities around the firms do not overlap. Alternatively, we could treat firms in close proximity as one unit to satisfy this assumption.

i.e:

$$\sum_{i \in \mathcal{I}_c} \max_{r_i} \alpha_{ic} Surplus(r_i) + Tax(r_i) \tag{3.2}$$

3.2 Testable Hypotheses

In this simple framework, when the two townships merge with each other, "merging border firms" move closer to the township center, with the result that they face tighter environmental regulation. In response to the increased regulatory stringency, these firms reduce their output, increase their abatement investment, and reduce their pollutant emissions. Local housing price also increases as a result. We briefly summarize these propositions below, and prove them under general conditions in the Appendix.

Proposition 1. *When a firm becomes farther away from the township border (higher α_{ic}), it will face tighter environmental regulatory standards (higher r_i^*).*

Proposition 2. *When facing tighter environmental regulation (higher r), firms have lower output and profit levels (lower f and lower π).*

Proposition 3. *When facing tighter environmental regulation (higher r), firms increase emission abatement capital (K_E), provided that they do not suspend production.*

Proposition 4. *When facing tighter environmental regulation (r), firms reduce total emissions (E).*

Proposition 5. *When a community becomes farther away from the township border (higher α_{ic}), local housing price will increase (higher $Surplus$), and so will total welfare ($Surplus + Tax$).*

4 Data

We match detailed GIS maps of township boundaries to the administrative records of townships mergers, and combine them with novel firm-level geocoded panel datasets on pollutant emission and manufacturing production, which allows us to exploit rich and precise spatial variation to implement our research design. In addition, we utilize a transaction-level geocoded land price dataset to understand residential welfare. In the following subsections, we describe each dataset in more detail, and explain how we combine them to assemble the final sample for analysis.

4.1 GIS Maps and Administrative Records on Township Mergers

We obtain township boundary GIS maps of China in 2000 and 2010 from the National Bureau of Statistics, as shown in figure 3. By overlaying these two maps, we can precisely measure all the boundary changes caused by township mergers during this decade [10]. We digitize the administrative records of township boundary changes published yearly by the Ministry of Civil Affairs (MCA) between 2002 and 2008[11], and match this information with the boundary changes observed in the township GIS maps. Utilizing the merged GIS-administrative dataset, we are able to identify the exact timing of every merger case, and the associated township boundary changes.

4.2 Production Data of Industrial Firms

Our firm-level production information comes from the Annual Survey of Industrial Firms (ASIF) between 1998 and 2013. The dataset, collected and maintained by the National Bureau of Statistics (NBS), includes all the industrial enterprises with annual sales exceeding 5 million RMB. It contains a rich set of information obtained from the accounting books of these firms, such as input, output, sales, taxes, subsidies, etc. This dataset has been widely used in economic research, and more details about its construction and cleaning processes can be found in Brandt et al. (2012).[12]

[10] Township splits or repeated mergers in multiple years are rare. In our analysis, we keep only the townships that merged once, excluding township splits and repeated merger.

[11] The ministry did not keep track of township-level changes until 2002, and stopped keep records again after 2008

[12] A widely-known issue with the ASIF dataset is that it contains outliers. We follow standard procedures in the literature to clean the data. We first drop observations missing key financial indicators or having negative values for value-added, employment, and fixed capital stock; then we exclude observations that violate accounting principles: liquid assets, fixed assets, or net fixed assets larger than total assets; current depreciation larger than accumulated depreciation. Finally, we trim the data by dropping observations with the values of key variables outside the range of the 0.5 to 99.5 percentile.

Following Yang (2015), we deflate firm-level output, value-added, and profit using yearly output price indexes for every 2-digit industry, which are collected from the "Urban Price Yearbook 2011" published by the NBS. Because those price indexes are linked across different years, we can use them to deflate yearly nominal output to real output in 2000.

Importantly, the ASIF dataset provides a detailed street address for each sampled firm in every year. Utilizing this address information, we geocode the location of all the firms that ever appeared in our sample.

4.3 Emission Data of Polluting Firms

For environmental outcomes, we use firm-level emission data from China's Environmental Survey and Reporting (ESR) database. The ESR database is maintained by the Ministry of Environmental Protection (MEP) and is used to monitor the polluting activities of all major polluting sources, including heavily polluting industrial firms, hospitals, residential pollution discharging units, hazardous waste treatment plants, and urban sewage treatment plants.

The ESR documents firm-level (polluting-source-level) emissions of various pollutants, such as SO2, NOX, and COD. Another unique feature of this dataset is that it contains information on each firm's investment in emission abatement equipment and facilities.

The sampling criteria in the ESR is based on the cumulative distribution of emissions in each county. Polluting sources are ranked based on their emission levels by different "criteria" pollutants, and those jointly contributing to the top 85% of total emissions in a county are included in the database. Polluting sources are required to report their emission levels of various pollutants to county-level Environmental Protection Bureaus (EPBs). Local EPBs are responsible for checking the quality of the data and upper-level EPBs then verify the data.[13]

Like the ASIF dataset, the ESR records the detailed street address of each firm. Using this information, we find the coordinates for every firm that ever appeared in this dataset during our sample period.

[13] More details of the database are described in He et al. (2020) and Cai et al. (2016).

4.4 Land Transaction Data

To examine the merger's impacts on the local land market, we collect land auction data from 2004 to 2015 published by the Land Bureau of China. Land ownership in China is communal in principle, therefore the primary land market is hosted by local governments, in which potential land developers gain "land development rights" through open auctions. Since the establishment of this primary land market in 2004, the details of each transaction have been publicized on the Land Bureau's official website (Cai et al., 2013). For each parcel of land sold by the government to the developers, our scraped data include detailed information such as the address, area, land use category, final sales value, and sales date. All prices are adjusted for inflation using a GDP deflator with 2010 as the base year.

Utilizing the location of each plot reported by the government record, we geocode each land transaction during our sample period.

4.5 Data Matching

To match these different sources of information, we first overlay the 2000 and 2010 township GIS maps, which identifies all the eventually merged township pairs.[14] Then, we overlay all the geocoded ASIF firms on that map based on their coordinates.

Figure 4 demonstrates a zoomed-in version of this overlaid map. The red dashed lines represent the 2000 borders that became defunct after the merger, whereas the black lines represent the official borders in 2010. We categorize firms into two groups based on changes in "distance to border" in 2000 and 2010. If the change in distance for a firm is larger than some small number ϵ [15], we label it as a red dot on the map, otherwise a blue circle. The red dot firms were all adjacent to the merging border, whereas the blue circle firms were not. After the merger, the red dot firms "move" closer to the center of the township, and will likely to be subject to stricter regulation according to our theoretical predictions. The discontinuous change in "distance to border" thus constitutes an intuitive

[14]Specifically, we first take GIS information (polygons) of all the townships that existed in 2000. Then the GIS software projects their positions onto the 2010 map and determines to which township they belong in 2010.. If multiple 2000 towns belong to the same town in 2010, we will code it as a merge. There are rare cases where a town in 2000 split into several smaller towns in 2010. Those towns are dropped from our sample.

[15]We use an ϵ instead of 0 to tolerate measurement errors.

measure of the "centralization shock" received by each firm.

Therefore, for each geocoded firm, we compute its minimum distance to township border, in both 2000 and 2010. Figure 5 illustrates the distributions of firms' "minimum distance to township border" in 2000 and 2010 respectively. As can be seen, for a substantial fraction of firms, their "minimum distance to township border" increased by 1-3 kilometers due to township mergers during our sample period. To put such a shock in context, Figure 5 also shows the distribution of township radius, which is defined as the radius of the maximum inscribed circle of a township. The average township radius is about 4 km, which means that the average shock created by township mergers effectively moves a firm on the merging border halfway to the center of that township.

We repeat the same spatial exercise for the geocoded ESR firms and land plots. Ideally, we would like to merge the ASIF and ESR datasets with each other, but due to differences in sampling strategies and coding systems, less than 10% of the ESR firms can be matched with the ASIF dataset. As a result, we will analyze these two datasets separately, but provide discussions on how to incorporate estimates from these different datasets in the same back-of-the-envelope calculation.

4.6 Summary Statistics

Following the aforementioned procedure, we identified 1115 township merger cases that happened between 2002 and 2008, which involved 3282 pre-merger townships in 2000. In these townships, there were 8324 polluting ASIF firms (11276 polluting ESR firms) located on a non-merging border, and 1063 polluting ASIF firms (1903 polluting ESR firms) located on a merging border.[16]

Table 1 compares the pre-merger characteristics of firms by their relative locations to township border. As shown in Panel A, in the ASIF sample, polluting firms located on merging borders would experience substantial increases in their distances to township boundary once the mergers happen, while firms located on non-merging borders do not experience such shock. Other than this stratifying variable, merging border firms and non-merging border firms are well-balanced

[16]A firm is defined as a "border firm" if its minimum distance to township boundary is below 2 km. A border firm is defined as being located on the "merging border," if its "distance to border" changed by more than 2 km after the merger took place.

along all the main dimensions of production, including industrial composition, firm age, input and output, TFP, taxes and subsidies, etc. As shown in Panel B, the ESR sample shows similar patterns: polluting firms located on merging borders would experience substantial increases in their distances to township boundary once the mergers happen, but the emission and abatement conditions are comparable between "merging border" and "non-merging border" firms, again supporting the validity of our empirical design.

5 Empirical Analysis: Firm Activities

In Section 3, our model predicts that when a polluting firm shifts from border to center following a township merger, it will reduce its output while increasing its abatement investment, which leads to lower profit, but also lower pollutant emissions. In this section, we confront these theoretical propositions empirically.

5.1 Production Activities

5.1.1 Main Results

We first investigate the impacts of internalizing border spillovers on firm production activities. Specifically, we focus on border firms in the ASIF dataset, and estimate the following econometric model:[17]

$$Y_{ijst} = \alpha \cdot Distance_{ist} + \beta \cdot Distance_{ist} \cdot Polluting_j + \sigma_i + \lambda_{st} + \gamma_{jt} + \epsilon_{ijst} \quad (5.1)$$

where Y_{ijst} is the outcome of interest for firm i in industry j in township s in year t. $Distance_{ist}$ is the nearest geographic distance between firm i and the border of township s in year t. This variable increases for "merging border firms" after the integration of two townships, but is constant for firms located on non-merging borders. $Polluting_j$ is a dummy variable that equals 1 if industry j is one of the 16 "heavily polluting" industries defined by the Ministry of Environmental Protection (MEP), and 0 otherwise.[18] σ_i is the firm fixed effect. λ_{st} is the township-by-year fixed effect. γ_{jt} is the industry-by-year fixed effect. ϵ_{ijst} is the error term. Standard

[17] "Border" is defined as within 2 km away from township border in 2000. We conduct robustness checks using various alternative definitions in Appendix Table X.

[18] See Appendix Table A1 for more details on the classification of polluting industries.

errors are two-way clustered at the township and industry levels.

The inclusion of firm fixed effects σ_i captures any time-invariant characteristics of the firm, such as endogenous location choices. The township-by-year fixed effects λ_{st} pick up township specific shocks, including the common shock of township merger on local firms. The industry-by-year fixed effect controls for industry-specific shocks, such as changes in regulatory standards. In this highly saturated model, we are essentially comparing firms (in different locations relative to the merging border) within the same industry, the same township, and the same year, before and after the township merger. Therefore, as long as the timing of a township merger is not driven by a factor associated with a future break in trends between "merging border" and "non-merging border" firms within that township, α would identify the causal effect of a non-polluting firm being "moved" away from border by 1 kilometer, and β would causally identify the additional effect of a polluting firm being "moved" away from border by 1 kilometer.[19]

We first investigate the theoretical hypotheses regarding output and profit levels. As shown in Table 2, when a non-polluting firm "moves" from border to center due to township merger, there is no noticeable change in its production activities. In stark contrast, when a polluting firm "moves" from border to center due to township merger, for every kilometer it shifts away from township border, it experiences a 2.8% reduction in net profit, a 1.6% reduction in total output, and a 2.6% reduction in total exports.

These results are consistent with our theoretical predictions, suggesting that the local governments indeed internalize border pollution by penalizing polluting firms that move farther away from the township border. As shown in Figure 5, the average change in "distance to border" for the merging-border firms is 3 km. From this perspective, the "average treatment effect on the treated" can be inferred by multiplying the coefficients by three: after township mergers, border polluting firms on average face a 8.4% reduction in net profit, a 4.8% reduction in total output, and a 7.8% reduction in total exports.[20]

[19]While it is unclear whether non-polluting firms have positive or negative externalities, pollution is associated with additional negative externalities. Therefore, our main theoretical predictions correspond to the sign of β in the econometric model.

[20]A related thought experiment is that, if an average-sized township had the shape of a perfect circle, then its radius would be approximately 4km (see Figure 5). Therefore, our estimates suggest thatd if a polluting firm is moved from the very border to the very center of a township, it faces on average a 11.2% drop in profit, a 6.4% drop in output, and a 10.4% drop in exports.

In Table 3, we investigate the theoretical prediction regarding productivity. Similar to the previous findings, the non-polluting firms do not appear to be affected by township mergers. However, when polluting firms "move" away from township border, their value-added decreases significantly, without any noticeable changes in labor and capital inputs. These findings confirm that productivity, as measured by the amount of value-added not attributable to labor and capital inputs, drops when a polluting firm's environmental spillovers suddenly get internalized by the local government.

5.1.2 Township Merger and Upper-Level Political Incentives

As discussed in Section 2, township leaders influence firm-level regulatory burdens mainly by selectively cooperating with the upper-level environmental bureaus, who hold the actual powers to punish polluting firms by issuing fines or temporary shutdowns.

Therefore, when a township merger happens, the extent to which the newly merged township government could effectively alleviate environmental spillovers on the merging border depends not only on the township government itself, but also the upper-level government's incentives to enforce environmental regulation. During our sample period, the local environmental bureaus were led by the local prefectural governments, as a result, the political incentives of the prefectural city leaders should play important roles in the internalization of township spillovers.

To investigate this hypothesis, we match our baseline firm data to the resumes of the local prefectural party secretaries (PPSs), who are the top-ranked officials at the prefectural city level. As documented by Wang (2016) and He et al. (2020), a PPS can rarely be promoted to the provincial level once he reaches the age of 57, which substantially reduces the PPS's incentives in implementing the central government's policies. Leveraging this unique feature, we define a dummy variable "incentive," which equals one if the local PPS is younger than 57 in that year, and zero otherwise. We interact this variable with the terms in Equation 5.1 to investigate how the political incentives of the PPSs affect the internalization of pollution spillovers following township mergers.

As shown in Appendix Table A2, in prefectural cities with "incentivized" leaders, township mergers have significantly stronger impacts on the merging-border polluting firms, while there is no similar heterogeneity for their non-polluting counterparts. These findings are consistent with our interpretation of an underlying principal-agent problem between the township officials and their upper-level leaders.

5.2 Emission Activities

5.2.1 Main Results

To investigate the environmental consequences of internalizing border spillovers, we use the ESR data and estimate a modified version of Equation (5.1):

$$Y_{ijst} = \beta \cdot Distance_{ist} + \sigma_i + \lambda_{st} + \gamma_{jt} + \epsilon_{ijst} \tag{5.2}$$

where the outcomes of interest are pollutant emissions (SO_2 and NO_x) and abatement investment. Since the ESR data samples only the heavy polluters, the modified econometric model no longer distinguishes between polluting and non-polluting firms using the interaction term $Distance_{ist} \cdot Polluting_j$.

As shown in Table 4, "moving" a polluting firm away from the township border by 1 kilometer would decrease its SO_2 emissions by more than 3.2%, reduce its ownership of coal boilers by 1.1%, and increase its abatement investment by 4.2%. Considering that the average "treated" firm moves away from township border by roughly 3 km after township merger, in terms of the "average treatment effect on the treated," our estimates indicate a nearly 10% emission reduction and a more than 12% increase in abatement investment.[21]

5.2.2 Corroborating with ASIF Data

We also corroborate the emission results using the ASIF data. In the 2004 ASIF data, firms reported the "emission fees" they had to pay in that year, which can be considered as a monetary measure of pollutant emissions, and is also the only environmental indicator documented in this dataset. Since this information only exists in 2004, we can no longer exploit within-firm variation in "distance to border"

[21] Ideally, we also would like to investigate other types of pollutant emissions. Unfortunately, the other main measure of air pollution, NOX, was not documented in the ESR dataset until 2005. Given that our township merger information ends in 2008, we do not have enough statistical power to analyze the impacts of merger on NOX.

for identification. Instead, to conduct a DiD analysis using this cross-sectional data, we need a stronger identifying assumption: among all the townships that merged during the 2002-2008 time window, whether the merger happened before or after 2004 is orthogonal to the relative emissions of merging border polluting firms and non-merging border polluting firms. Under this assumption, we can compare "the difference between merging border firms and non-merging border firms in townships that had already merged in 2004" to "the difference between merging border firms and non-merging border firms in townships that had not yet merged in 2004," which provides the DiD estimate for the impacts of "shifting" from township border to township center.

Formally, we estimate the following econometric specification:

$$Y_{is} = \alpha \cdot Merged_s + \beta \cdot \log\left(\Delta Distance_{is}\right) + \gamma \cdot Merged_s \cdot \log\left(\Delta Distance_{is}\right) + \epsilon_{is} \tag{5.3}$$

where Y_{is} is the log amount of emission fees paid by firm i in township s in 2004. $Merged_s$ is a dummy variable which equals 1 if township s has already merged in 2004, and 0 otherwise. We define $log(\Delta Distance)_{it}$ to be the logarithm of the change in firm i's distance to the border after the township merger, so the larger this measure is, the closer the firm is to the merging border.

We estimate Equation 5.3 separately for polluting firms and non-polluting firms, and report the results in Table 5. As shown in Column 1, before the merger, polluters located on the merging border paid higher emission fees than other firms in the township (which includes both central firms and "non-merging border" firms). However, after the merger, this gap in emission fees is completely eliminated.

In Column 2, we restrict the sample to only border firms (within 2 km from border in 2000), which compares "merging border polluters" to "non-merging border polluters." As we can see, before the merger, the two groups paid similar amounts of emission fees, while after the merger, "merging border polluters" started to pay significantly lower fees, again suggesting that the internalization of pollution spillovers reduced firm emissions.

In Columns 3 and 4, as a placebo test, we repeat the same analysis for non-polluting firms. As we can see, non-polluting firms pay much lower emission fees in general, and the magnitude of their emission fees is not affected by the township merger.

5.3 Threats to the Baseline Analysis and Robustness Checks

In this section, we address the potential threats to the validity of our baseline research design, and conduct a series of additional checks for robustness.

5.3.1 Pre-Merger Trends

Our baseline research design relies on the assumption that, had there not been a township merger, "polluters located near merging borders" and "polluters located near non-merging borders" should have parallel trends in production and emission activities.

To infer the validity of this assumption, we test for parallel trends prior to township merger. Specifically, we keep only the pre-merger sample, and (counterfactually) code the jump in $Distance_{ist}$ as if it happened one year, two years, or three years before the actual township merger. If the "treated" and "control" firms had differential trends before the merger actually happened, then the coefficients for the 1-year and 2-year leads of $Distance_{ist}$ should be statistically different from zero. As shown in Tables A3 and A4, for all the main production and emission variables, the lead coefficients for polluting industries are precisely estimated and statistically indistinguishable from zero. This reassures us of the validity of our research design.

5.3.2 Alternative Measures of Internalization

In the baseline analysis, we use the change in a firm's minimum distance to township border to proxy for the internalization of its environmental spillovers: the closer a firm moves towards the township center, the more its externalities will be internalized. However, there can be measurement errors associated with this proxy, especially for firms located near the "corners" of a township, as illustrated in Appendix Figure A2.[22]

To address this issue, we come up with an alternative way to measure the internalization of externalities. For each firm, we draw a 2-km radius circle around

[22] For a "merging border polluter" located in township corner, the township merger barely changes its minimum distance to border, but substantially increases the internalization of its externalities.

it, and identify how much of this circle falls in the host township of the firm.[23] The larger this "overlapping area" is, the more a firm's externalities will be internalized.

As shown in Table 6, for a non-polluting firm, having its externalities internalized has no significant impact on production outcomes. On the other hand, for a polluting firm, when the entirety of its externalities are fully internalized, its profit will drop by 15.6%, its output will fall by 13.2%, and its exports will decrease by 12.6%. Suppose a polluter lies exactly on the merging border; then, prior to the merger, half of its externalities go to the neighboring township. Our estimates suggest that having its cross-border spillovers fully internalized would lead to a 7.8% drop in profit, a 6.6% drop in output, and a 6.3% drop in exports, which are comparable in magnitudes to the baseline impacts according to Table 2.

5.3.3 Alternative Definitions of "Border Firm"

In our baseline specification, we kept only "border firms," defined as firms within 2 km of township border in 2000. This bandwidth choice is motivated by the range of negative externalities caused by industrial pollution documented in the existing literature (Currie et al., 2015).

However, our main findings are highly robust to this bandwidth choice: as shown in Appendix Table A5, if we keep firms within 1.5 km or 3 of the 2000 township border, all the baseline findings remain quantitatively similar.

5.3.4 Replacing Log Values with Inverse Hyperbolic Sine Functions

In the baseline analysis, we use "$Log(1+x)$" to construct the outcome variables. One potential concern is that some variables in the production dataset, such as profit and value-added, contain a small amount of negative values, which were dropped when taking log.

To investigate the robustness of our findings with respect to alternative ways of variable construction, we also use the inverse hyperbolic sine function (arcsinh) to transform all the key variables, which does not require dropping negative values.

[23] We use 2 km because this is consistent with the range of pollution externalities documented in the existing literature (Currie et al., 2015), but our findings are robust to any range between 1.5 km and 3 km.

As shown in Appendix Tables A6 and A7, when we re-run the baseline analyses using these alternative variables, all the main findings still go through.

5.3.5 Inter-Firm Spillover Effects

Because we are comparing firms within the same township, one potential concern is that the non-merging border firms are also affected by the increased emission regulation of merging border firms after the integration, which means the Stable Unit Treatment Value Assumption (SUTVA) is violated. For instance, there might exist a "monitoring capacity constraint" for township governments. When a township government decides to spend more effort monitoring the merging border firms, it mechanically has to spend less effort monitoring other firms if bounded by the "monitoring capacity constraint." If this is the case, our baseline specification would over-estimate the effects of internalizing border emissions.

To test this, we replace the township-year fixed effects in the baseline specification with the province-year fixed effects. If the SUTVA violation is indeed driving the main result, then the baseline result should disappear after relaxing the control group from "firms within the same township" to "firms from the same province," because the treatment and control groups would no longer be subject to the same monitoring constraint.

The results are shown in the Table A8. We find that the estimated coefficients with province-year fixed effects are barely different from the baseline coefficients, indicating that the violation of SUTVA should not be a major concern.

5.3.6 Population Distribution

Another possibility is that when the merger happens, a large fraction of population from both townships starts to concentrate close to the center of the new township (the merging border). Since higher population density would make pollution in that area more detrimental, even if the merger did not lead to any internalization of border spillovers, the township government still has an incentive to impose more stringent regulation on the firms close to the merging border.

Given that we are comparing firms immediately before and after the merger happened, to the extent that migration might need more time to respond to a township merger, we think this is unlikely to be confounding our results. Another argument against this migration channel is that when population concentrates towards the merging border, other regions in this new township should have a smaller population. In that case, we would expect the township government to relax the regulation on other polluting firms after the merger, which means that there should be spillover effects between merging-border firms and other firms. But as discussed in Section 5.3.5, spillover effects do not play important roles in our data, confirming that migration is not the mechanism behind the main findings.

Nevertheless, we directly investigate the robustness of our findings to population distribution by leveraging geocoded population census data, which allows us to directly control for local population density in our main specifications. As shown in Appendix Table A9, none of our main findings changes in any substantial way following the inclusion of this population control, again confirming that "changes in population distribution" cannot confound the main results of this paper.

6 Empirical Analysis: Residential Land Transactions

Our model predicts that when border spillovers are internalized, both the amount of local land transactions (extensive margin) and local land price (intensive margin) would go up. In this section, we investigate both margins using the universe of land auction data scraped from the official website of the Ministry of Land and Resources. Because each parcel of land appears only once in the government auction data,[24] we can no longer exploit within-plot variation as we did in the previous firm-level analyses. Instead, we use two modified econometric models to investigate the extensive and intensive margins respectively.

6.1 Extensive Margin: Amount of Land Transactions

To understand the extensive margin effect – the effect of internalizing border spillovers on the quantity of residential land transactions – we divide the area of each township into two groups, "*Merging Border Area*" and "*Other Area*," as

[24]Subsequent sales on the secondary markets are not done via these government auctions.

illustrated in Figure 6.[25].

By counting the number of land parcels auctioned in each group in each year, we construct group-year level panel data of land transactions. Combining the data with information on township mergers, we estimate the following model:

$$\begin{aligned} Y_{ist} =& \alpha \cdot PostMerge_{st} + \beta \cdot MergingBorder_i \\ &+ \gamma \cdot PostMerge_{st} \cdot MergingBorder_i + \delta_s + \eta_t + \epsilon_{ist} \end{aligned} \tag{6.1}$$

where Y_{ist} is the outcome of interest for land area i in township s in year t. $PostMerge_{st}$ is an indicator variable for whether township s in year t completed a merger. $MergingBorder_i$ is also an indicator that equals 1 if land area i is a *Merging Border* area, and 0 otherwise. δ_s and η_t, respectively, are township and year fixed effects. α identifies the overall impact of the township merger, and β identifies the difference between *Merging Border Area* and *Other Area* before the merger. γ is the coefficient of interest. It estimates the average effects of changing from being a border area to a non-border area on the likelihood of new residential development.

In Table 7, we present the results of estimating Equation 6.1. We divide each township into two areas and define a dummy variable *Residential Construction* that equals 1 if any residential development occurs in the area. Column 1 shows that before the merger, the merging border areas are 14% less likely to have residential projects, while after the merger, this gap is almost completely eliminated (12.5% increase). Column 2 controls for the area fixed effects and reports similar results.

6.2 Intensive Margin: Land Price

To estimate the intensive margin effects – the impacts of internalizing border spillovers on land prices – we compare the prices of residential land parcels close to and far away from the merging border, before and after the merger. Specifically, we use $log(\Delta Distance)_{it}$ to measure how much a land parcel's distance to the border would change after the township merger happens, which is a proxy for its

[25] *Merging Border* is defined as the areas whose distance to border increases more than 2 km after the merger. The rest of areas in the pre-merger township is defined as *other* area. The results are similar using other thresholds between 1.5 km and 3 km.

closeness to the merging border. We estimate the following econometric model:

$$\begin{aligned}Y_{ist} =&\alpha \cdot PostMerger_{st} + \beta \cdot \log(\Delta Distance)_{it} \\ &+ \gamma \cdot PostMerger_{st} \cdot \log(\Delta Distance)_{it} + \delta_s + \eta_t + \epsilon_{ist}\end{aligned} \tag{6.2}$$

where $PostMerger_{st}$ is a dummy variable measuring whether the merger has already happened for township s in year t.

Table 8 presents the estimates of the impacts on the residential land price. We use the logarithm of per unit residential land price as the outcome variable. Column 1 reports the results of estimating Equation 6.2. As we can see, before the merger, land parcels on the merging borders were substantially cheaper, while after the merger, this price gap is completely eliminated. Column 2 controls for more characteristics of the land and yields a similar estimate. These findings are again highly consistent with the predictions of our model.

To further investigate the underlying mechanisms of our findings, Columns 3 and 4 employ a triple difference design. We define a dummy variable *Polluting Border* to be 1 if at least one polluting firm is located on the merging border (whose change in distance to the border is larger than 1.5km). By including the interaction terms with *Polluting Border*, we add another layer of variation to the DiD framework: comparing the pairs of merging townships with polluting firms on the shared border to those merging pairs without. Such comparison is illustrated in Figure 7.

As shown in Columns 3 and 4, all the results (initial price gap and subsequent price equalization) are almost entirely driven by cases where there are polluting firms located on the merging border, suggesting that the internalization of pollutant emissions, rather than other forms of externalities, is the main drivingforce behind the increase in residential land prices.

7 Welfare and Policy Implications

The findings in the previous sections suggest that, when the local governments make coordinated decisions following the merger of townships, they would internalize border emission spillovers and impose tighter environmental regulation on polluting firms located at the merging border. Such a change in regulation enforcement then leads to significant reductions in both pollutant emissions and industrial output for polluting firms located near the merging borders.

In this section, we discuss the welfare and policy implications of our empirical findings. Specifically, we leverage our baseline estimates to address two questions: (1) What is the net welfare effect of internalizing border spillovers in our sample, taking into account both the loss in output and the gain in residential rent? (2) What is the average abatement cost of Chinese manufacturing firms, and what does that tell us about the overall economic cost of various environmental regulation policies in China?

7.1 Net Welfare Calculation

As documented in our reduced-form analysis, the internalization of border spillovers came with both a loss in industrial output and a gain in residential welfare. To understand the net welfare consequences, we compare the pecuniary value of these two effects for the "treated" townships in our sample.

Our baseline estimates suggest that, for an average polluting firm, moving away from township border by 1 km reduces its total output by 1.6%. Linking this parameter to each polluting firm's annual total output and change in distance to border after merger, a simple calculation within our sample indicates that the annual output loss due to internalizing merging border spillovers exceeded 14.5 billion Chinese yuan.[26]

To measure the total increase in residential land rent, we need to take into account that not only are those newly-sold residential land plots more expensive, but all land plots close to merging borders gained potential market value.[27] Therefore, for every township merger during our sample period, we identify the "merging border area" whose externalities are internalized after township merger.[28] Linking the size of this "merging border area" to the average price change identified in our reduced-form analysis, we calculate that the overall increase in residential welfare exceeds 155 billion Chinese yuan.

Combining the two estimates, we estimate the net welfare gain of internalizing border spillovers to be around 140 billion Chinese yuan.

[26] $Total\ Output\ Loss = \sum_i (\frac{Output_i}{1-\Delta Distance_i \cdot 0.0159} - Output_i)$

[27] For plots sold before the merger, land owners are residual claimants of this appreciation; for plots that are yet to be sold, the governments will eventually be the residual claimants.

[28] We identify this area by drawing 2-km inward buffers for both 2000 and 2010 township boundaries, and compare the difference between the two.

Looking beyond townships that merged during our sample period, we can also calculate another, more "ideal" policy counterfactual: what would happen had all inter-township pollution spillovers been properly internalized? To answer this question, we calculate the total output loss if all the border polluters in ASIF are treated as if they are central firms (more than 2 km away from border), and the total residential land price gain if border land plots experienced the same appreciation as the merging-border plots. Our calculation suggests an economic cost of 46 billion RMB, and a residential welfare gain of 510 billion RMB.

7.2 Average Abatement Cost

Our baseline results suggest that when an average ASIF firm moves away from township border by 1 km, its total output decreases by 1.6%, and when an average ESR firm moves away from township border by 1 km, its SO_2 emission drops by 3.2%. Since the ASIF and ESR datasets have different sampling strategies, these two estimates are not directly comparable. Utilizing a simple re-weighting technique developed by He et al. (2020), we are able to link these two estimates and compute the average abatement cost (output loss per unit of emission reduction): our calculation indicates that, for the average polluting industrial firm in China, a 10% reduction in SO_2 would lead to a 7.1% loss in industrial output.

Leveraging this estimated average abatement cost, we can evaluate the overall economic costs of various environmental regulation policies in China. For instance, according to the 11th Five-Year Plan published by the MEP, industrial SO_2 emissions should decrease by 10% between 2001 and 2005. Based on the ASIF dataset, in 2000, the total industrial output of the 16 polluting industries was approximately 2.76 trillion RMB. Therefore, a simple calculation indicates that in order to fully comply with the centrally-designated abatement target (10% reduction in SO_2 emissions), the Chinese manufacturing sector had to bear an annual output loss of more than 200 billion RMB.

8 Conclusion

This paper aims to contribute to the debate about centralization versus decentralization by empirically estimating the economic and environmental impacts of internalizing border spillovers. The "township merger program" in China, where thousands of pairs of neighboring townships were merged over the last two decades, provides a good opportunity to study the border pollutions problem because the border changes introduce exogenous variations in firms' distance to the border. Exploiting the panel variations in firms' distance to the border and using firms at non-merging borders as counterfactual proxies enables us to adopt a stronger identification strategy than that previously used in the literature.

To get a comprehensive understanding of the impacts, we compile a novel panel dataset that consists of multiple sources of information. We start with two township GIS maps in 2000 and 2010. Then we geocode all the firms and parcels in the firm production data, firm emission data, and land transaction records so that those economic statistics can be linked to the changes in the border. Administrative records of township boundary changes from the Ministry of Civil Affairs are also digitized and incorporated into township GIS maps to help identify the year of border change.

Our empirical results show that when local governments make coordinated decisions, they can indeed better internalize border environmental spillovers. When facing more stringent regulatory enforcement, firms originally located on merging borders would reduce their emissions and invest more in abatement equipment, which leads to lower output, profit, and export levels. We also found evidence that both land prices and new developments of residential buildings as a result of the internalization of border pollution, indicating that household welfare increases.

Taken together, our results suggest that when local governments in China are acting independently, they are not able to perfectly coordinate or negotiate to fully internalize the border spillovers. The distortion from such failure in coordination between local governments is economically significant: our back-of-the envelope calculations suggest that if all township governments in China were to properly internalize their border environmental spillovers, there will be an overall welfare increase of more than 460 billion RMB.

Our findings have important policy implications. First, as the Township Consolidation Program is still ongoing, our estimates can help policymakers make projections when deciding which pairs of townships to merge with each other. Second, our results are relevant in other contexts for designing optimal environmental regulation regulation. A key lesson from our estimates is that spillovers can be internalized through coordinated decision-making. Cooperation among local governments can be achieved in many ways without forcing an annexation.

For example, a regional emission regulation commission may be established to take control of environmental regulatory decisions similar to the transportation commissions in many metropolitan areas.

It is important to note that our findings do not yield a clear-cut prediction for overall welfare impacts from the consolidation of township governments. Besides making joint decisions about environmental regulation, there are many other important advantages and disadvantages of merging local governments. For example, the merged governments may utilize economies of scale to provide public goods more efficiently, which improves the welfare of citizens. But larger administrative areas are also typically associated with less varieties in public goods to satisfy heterogeneous preferences, therefore decreasing the welfare of citizens. Our empirical approach is agnostic about the impacts from those other channels because any impacts affecting the entire township are captured by the township-year fixed effects in our econometric model. As a result, our estimated effects are only partial effects of the township consolidation through the environmental regulation channel, and we are unable to draw conclusions about overall welfare effects of the township consolidation. More research is needed to answer this important question.

References

Bazzi, Samuel and Matthew Gudgeon (2016) "Local Government Proliferation, Diversity, and Conflict," *Boston University*.

Becker, Randy and Vernon Henderson (2000) "Effects of air quality regulations on polluting industries," *Journal of Political Economy*, DOI: `http://dx.doi.org/10.1086/262123`.

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*, Vol. 97, pp. 339–351.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber (2012) "The political economy of deforestation in the tropics," *Quarterly Journal of Economics*, DOI: `http://dx.doi.org/10.1093/qje/qjs034`.

Cai, Hongbin, Yuyu Chen, and Qing Gong (2016) "Polluting thy neighbor: Unintended consequences of China's pollution reduction mandates," *Journal of Environmental Economics and Management*, DOI: `http://dx.doi.org/10.1016/j.jeem.2015.01.002`.

Cai, Hongbin, J Vernon Henderson, and Qinghua Zhang (2013) "China's land market auctions: evidence of corruption?" *The Rand journal of economics*, Vol. 44, pp. 488–521.

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*, Vol. 105, pp. 678–709.

Davis, Donald R. and David E. Weinstein (2002) "Bones, bombs, and break points: The geography of economic activity," *American Economic Review*, DOI: `http://dx.doi.org/10.1257/000282802762024502`.

Fisman, Raymond and Roberta Gatti (2002) "Decentralization and corruption: Evidence across countries," *Journal of Public Economics*, DOI: `http://dx.doi.org/10.1016/S0047-2727(00)00158-4`.

Foster, Andrew D. and Mark R. Rosenzweig (2002) "Household division and rural economic growth," *Review of Economic Studies*, DOI: `http://dx.doi.org/10.1111/1467-937X.00228`.

Fredriksson, Per G. and Daniel L. Millimet (2002) "Strategic interaction and the determination of environmental policy across U.S. States," *Journal of Urban Economics*, DOI: `http://dx.doi.org/10.1006/juec.2001.2239`.

Gray, Wayne B. and Ronald J. Shadbegian (2004) "'Optimal' pollution abatement - Whose benefits matter, and how much?," in *Journal of Environmental Economics and Management*, DOI: http://dx.doi.org/10.1016/j.jeem.2003.01.001.

Greenstone, Michael (2002) "The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures," *Journal of Political Economy*, DOI: http://dx.doi.org/10.1086/342808.

Greenstone, Michael and Justin Gallagher (2008) "Does hazardous waste matter? Evidence from the housing market and the superfund program," DOI: http://dx.doi.org/10.1162/qjec.2008.123.3.951.

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

Helland, Eric and Andrew B. Whitford (2003) "Pollution incidence and political jurisdiction: Evidence from the TRI," *Journal of Environmental Economics and Management*, DOI: http://dx.doi.org/10.1016/S0095-0696(03)00033-0.

Jia, Ruixue (2017) "Pollution for promotion," *21st Century China Center Research Paper*.

Kahn, Matthew E., Pei Li, and Daxuan Zhao (2015) "Water pollution progress at borders: The role of changes in China's political promotion incentives," *American Economic Journal: Economic Policy*, DOI: http://dx.doi.org/10.1257/pol.20130367.

Konisky, David M. and Neal D. Woods (2010) "Exporting air pollution? regulatory enforcement and environmental free riding in the United States," *Political Research Quarterly*, DOI: http://dx.doi.org/10.1177/1065912909334429.

Lipscomb, Molly and Ahmed Mushfiq Mobarak (2017) "Decentralization and pollution spillovers: Evidence from the re-drawing of county borders in Brazil," *Review of Economic Studies*, DOI: http://dx.doi.org/10.1093/restud/rdw023.

List, John A., Daniel L. Millimet, Per G. Fredriksson, and W. Warren McHone (2003) "Effects of environmental regulations on manufacturing plant births: Evidence from a propensity score matching estimator," DOI: http://dx.doi.org/10.1162/003465303772815844.

List, John A and Daniel M Sturm (2006) "How elections matter: Theory and evidence from environmental policy," *The Quarterly Journal of Economics*, Vol. 121, pp. 1249–1281.

Oates, Wallace E (1972) "Fiscal federalism," *Books*.

O'Brien, Kevin J. and Lianjiang Li (1999) "Selective policy implementation in rural China," *Comparative Politics*, DOI: http://dx.doi.org/10.2307/422143.

Prud'homme, Rémy (1995) "The dangers of decentralization," *World Bank Research Observer*, DOI: http://dx.doi.org/10.1093/wbro/10.2.201.

Redding, Stephen J. and Daniel M. Sturm (2008) "The costs of remoteness: Evidence from German division and reunification," *American Economic Review*, DOI: http://dx.doi.org/10.1257/aer.98.5.1766.

Reed Walker, W. (2011) "Environmental regulation and labor reallocation: Evidence from the clean air act," in *American Economic Review*, DOI: http://dx.doi.org/10.1257/aer.101.3.442.

Ryan, Stephen P (2012) "The costs of environmental regulation in a concentrated industry," *Econometrica*, Vol. 80, pp. 1019–1061.

Saavedra, Luz Amparo (2000) "A Model of Welfare Competition with Evidence from AFDC," *Journal of Urban Economics*, DOI: http://dx.doi.org/10.1006/juec.1999.2141.

Sigman, Hilary (2002) "International spillovers and water quality in rivers: Do countries free ride?," *American Economic Review*, DOI: http://dx.doi.org/10.1257/00028280260344687.

——— (2005) "Transboundary spillovers and decentralization of environmental policies," *Journal of Environmental Economics and Management*, DOI: http://dx.doi.org/10.1016/j.jeem.2004.10.001.

Tiebout, Charles M. (1956) "A Pure Theory of Local Expenditures," *Journal of Political Economy*, DOI: http://dx.doi.org/10.1086/257839.

Tilt, Bryan (2007) "The political ecology of pollution enforcement in China: a case from Sichuan's rural industrial sector," *The China Quarterly*, Vol. 192, pp. 915–932.

Wang, Mark, Michael Webber, Brian Finlayson, and Jon Barnett (2008) "Rural industries and water pollution in China," *Journal of Environmental Management*, DOI: http://dx.doi.org/10.1016/j.jenvman.2006.12.019.

Wang, Shaoda (2016) "Fiscal Competition and Coordination: Evidence from China."

Wildasin, David E (1991) "Income redistribution in a common labor market," *The American Economic Review*, pp. 757–774.

Wilson, John Douglas (1999) "Theories of Tax Competition," *National Tax Journal*.

Yang, Rudai (2015) "Study on the total factor productivity of Chinese manufacturing enterprises," *Economic Research Journal*, Vol. 2, pp. 61–74.

Zodrow, George R. and Peter Mieszkowski (1986) "Pigou, Tiebout, property taxation, and the underprovision of local public goods," *Journal of Urban Economics*, DOI: `http://dx.doi.org/10.1016/0094-1190(86)90048-3`.

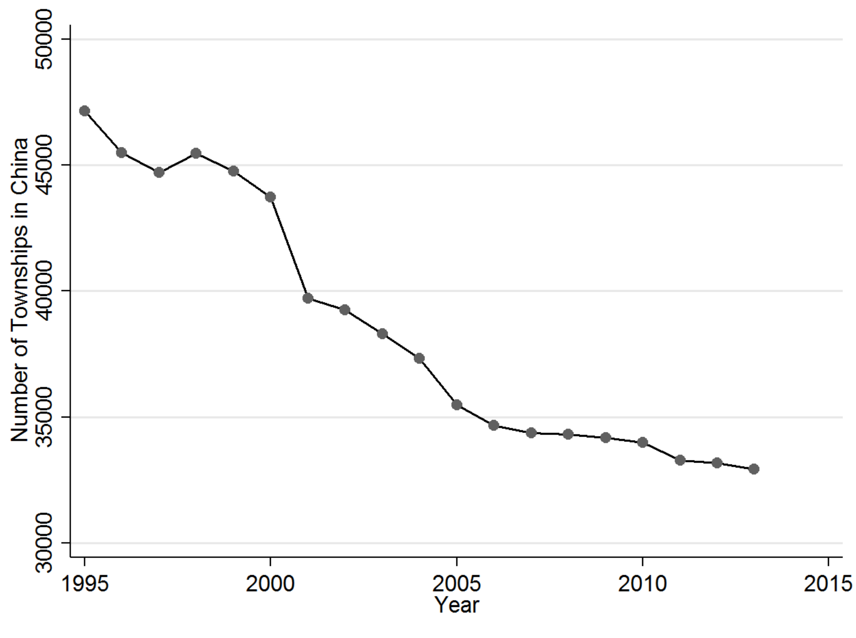

Figure 1: Number of Townships Over Time

Note: This figure documents the total number of townships in China between 1995 and 2013.

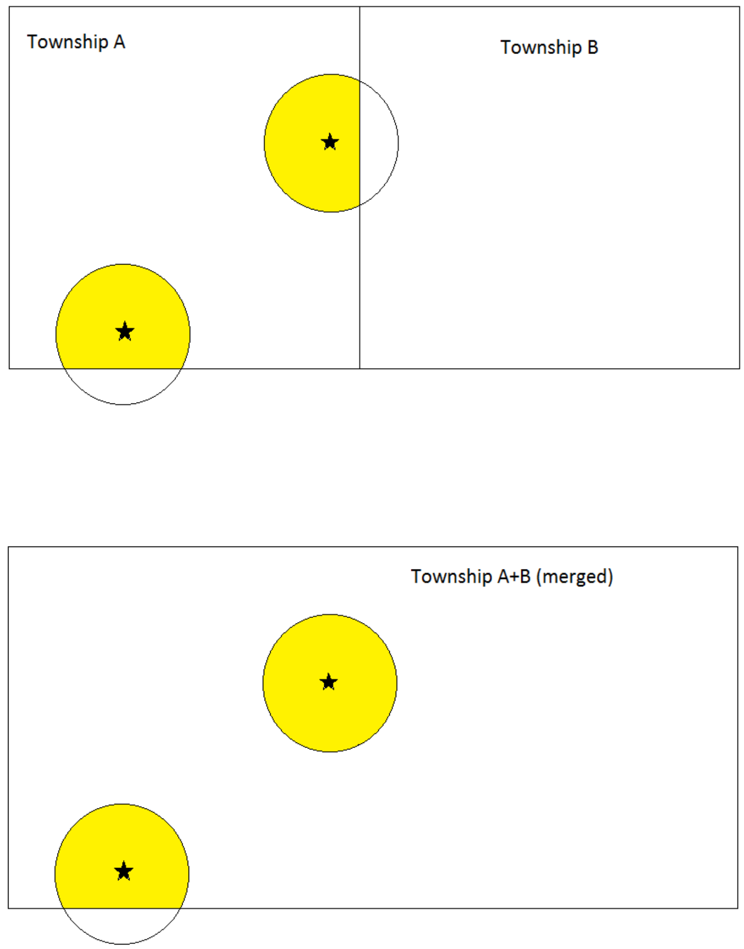

Figure 2: Research Design

Note: This figure illustrates the research design, which compares "merging border polluters" and "non-merging border polluters," before and after the township merger.

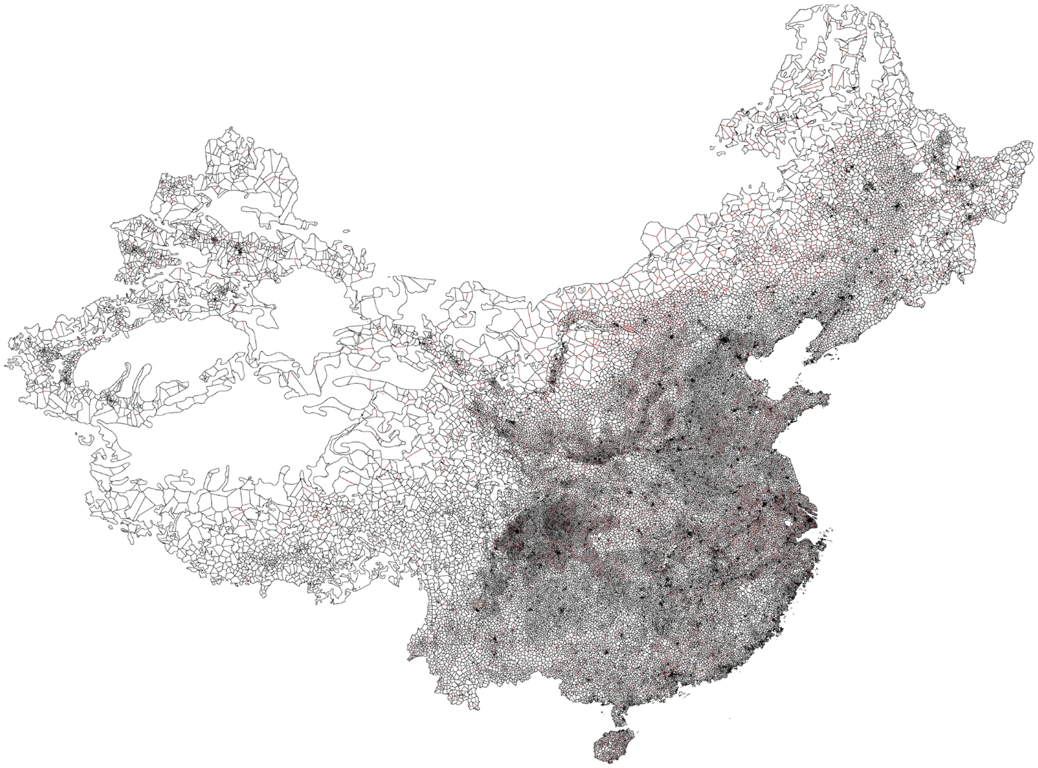

Figure 3: Township Boundary Changes

Note: In this figure, we overlay the 2010 township boundary map (in black) and the 2000 township boundary map (in red). Therefore, the remaining red borders displayed in this overlaid map represent those "merging borders" that existed in 2000 but disappeared before 2010.

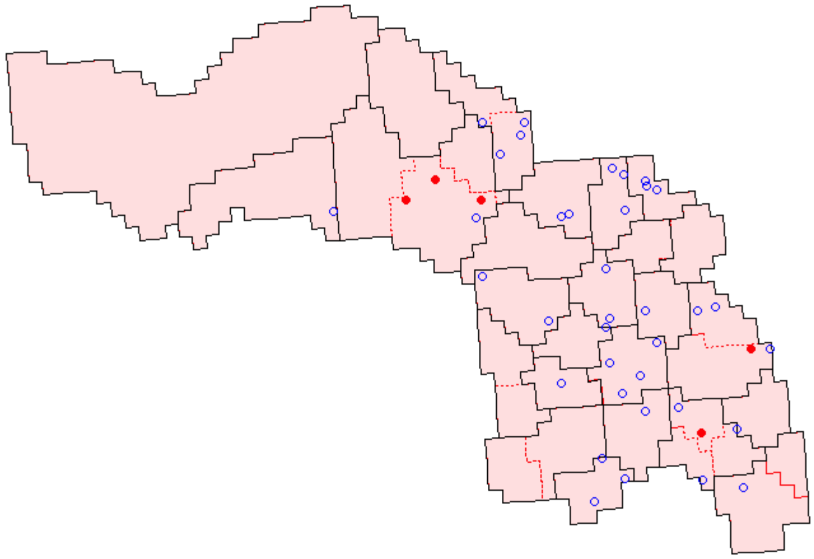

Figure 4: Illustration of Border Change in one County

Note: In this figure, we zoom in on one specific county to illustrate the construction of our key variables: the red borders are “merging borders,” while the black borders are “non-merging borders”; the red dots represent “merging border firms,” while the blue dots represent “non-merging border firms.”

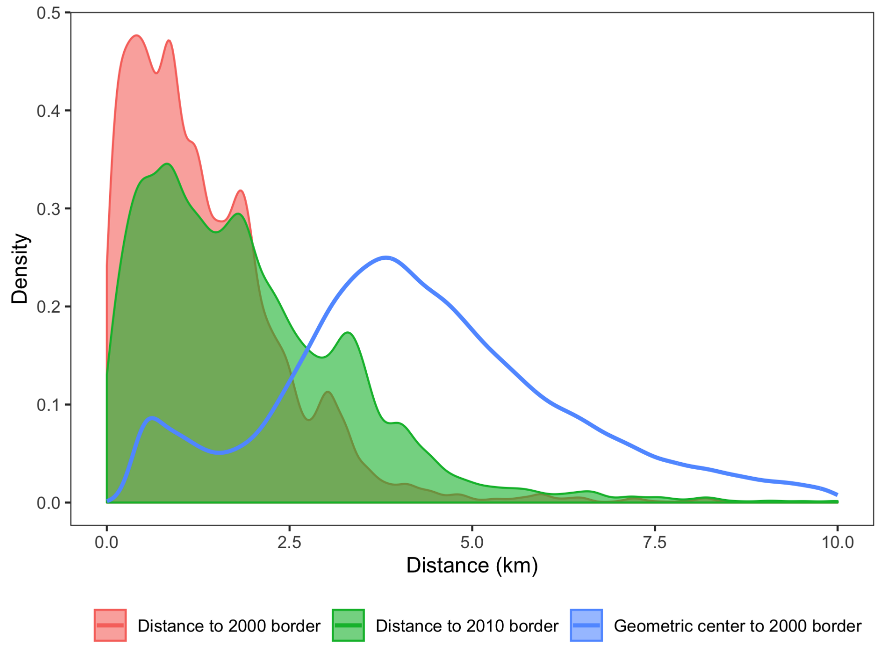

Figure 5: Distance to Border

Note: This figure illustrates the distributions of firm's minimum distance to township border in 2000 and 2010, as well as the distribution of the inscribed circle radius for 2000 townships.

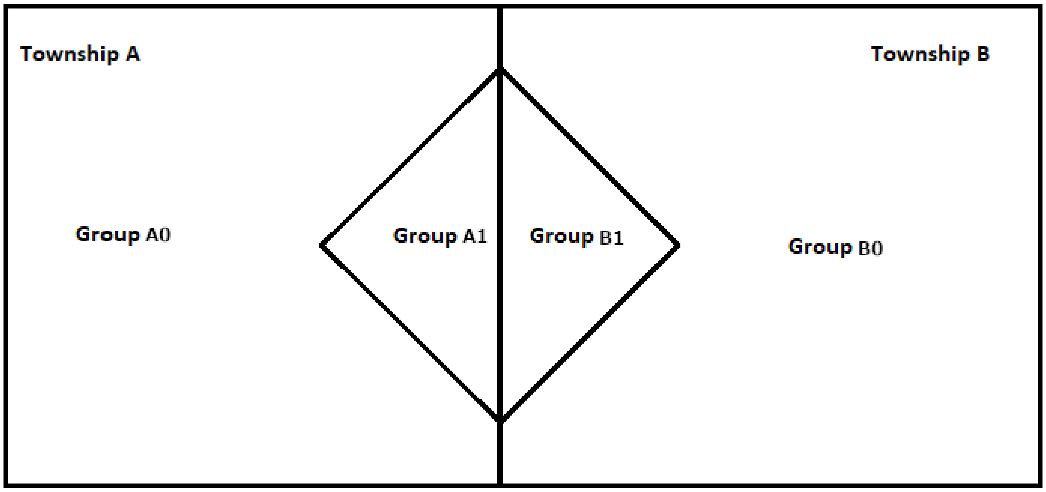

Figure 6: Residential Land: Extensive Margin

Note: This figure illustrates our research design for analyzing the "extensive margin" effect on land sales: we compare the amount of land sales in "merging border regions" and "other regions," before and after the township merger.

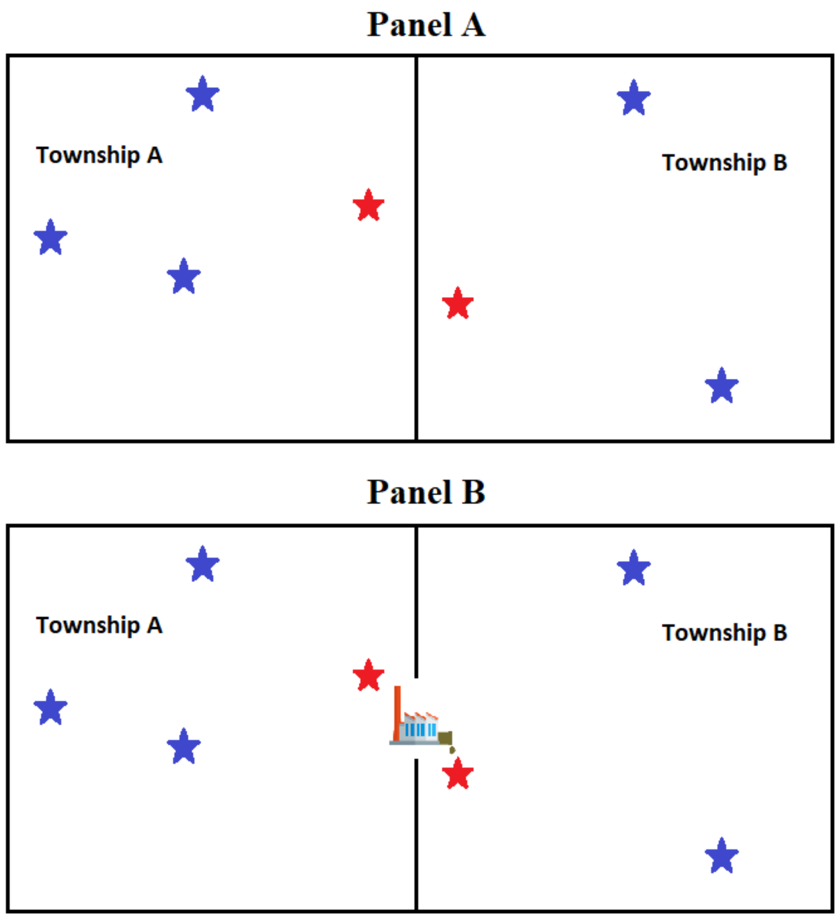

Figure 7: Residential Land: Intensive Margin

Note: These figures illustrate our research design for analyzing the "intensive margin" effect on land price: we compare the price of "merging border plots" to that of "other plots," before and after the township merger. We also compare cases with and without polluting firms located near the merging border.

Table 1: Pre-Merger Balance Test

	(1) Non-Merging Border	(2) Merging Border	(3) Difference
Panel A: Production Data			
Dist10 - Dist00 (km)	0.212	3.650	3.438***
	(0.474)	(2.816)	(0.213)
Log(Profit)	1.567	1.529	-0.038
	(1.299)	(1.326)	(0.081)
Log(Output)	5.175	5.144	-0.031
	(0.903)	(0.920)	(0.047)
Log(Value Added)	3.880	3.868	-0.013
	(1.131)	(1.120)	(0.054)
Log(Export)	0.963	0.843	-0.120
	(1.885)	(1.790)	(0.089)
Log(Labor)	4.758	4.806	0.049
	(1.076)	(1.097)	(0.068)
Log(Capital)	8.510	8.688	0.178*
	(1.744)	(1.864)	(0.108)
Log(TFP)	2.702	2.653	-0.050
	(1.009)	(1.025)	(0.050)
Log(Tax)	0.292	0.309	0.017
	(0.465)	(0.460)	(0.023)
Log(Subsidy)	0.186	0.226	0.040
	(0.584)	(0.639)	(0.029)
State Owned	0.086	0.070	-0.016
	(0.272)	(0.250)	(0.012)
Exit Market	0.129	0.141	0.012
	(0.269)	(0.285)	(0.013)
Observations	8,324	1,063	9,387
Panel B: Emission Data			
Dist10 - Dist00 (km)	0.240	3.819	3.579***
	(0.524)	(2.616)	(0.180)
Log(SO2)	7.403	7.446	0.043
	(4.252)	(4.194)	(0.189)
Log(Boilers)	0.514	0.491	-0.023
	(0.534)	(0.525)	(0.025)
Log(Cleaning Fee)	1.326	1.241	-0.086
	(1.453)	(1.433)	(0.066)
Observations	11,276	1,903	13,179

Note: Standard errors are reported in parentheses. Only border firms (within 2 km away from 2000 township border) are kept in the sample. "Merging border firms" are defined as firms whose minimum distance to township border would have increased by at least 2 km following the township merger; "non-merging border firms" include all other border firms.

Table 2: Township Merger and Firm Production

	(1) Log(Profit)	(2) Log(Output)	(3) Log(Export)
Dist to Border	0.0095	-0.0022	0.0070
	(0.0114)	(0.0076)	(0.0123)
Dist to Border*Polluting	-0.0284***	-0.0159**	-0.0258**
	(0.0089)	(0.0077)	(0.0108)
Constant	1.8195***	4.6564***	0.9890***
	(0.0144)	(0.0082)	(0.0152)
Mean of Dep Variable	1.817	4.646	0.985
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	237032	261262	216473
R squared	0.697	0.954	0.794

Note: Standard errors are two-way clustered at the township and industry levels. "Polluting" is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. "Dist to Border" is the minimum distance between a firm and the corresponding township border in a given year. Firm FE, Industry-by-Year FE, and Township-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table 3: Township Merger and Firm Productivity

	(1) Log(VA)	(2) Log(Labor)	(3) Log(Capital)
Dist to Border	0.0030	-0.0055	0.0067
	(0.0095)	(0.0059)	(0.0101)
Dist to Border*Polluting	-0.0223*	0.0008	-0.0095
	(0.0114)	(0.0055)	(0.0092)
Constant	4.0213***	4.8039***	8.4799***
	(0.0097)	(0.0064)	(0.0098)
Mean of Dep Variable	4.015	4.798	8.483
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	179141	250935	172509
R squared	0.787	0.837	0.892

Note: Standard errors are two-way clustered at the township and industry levels. “Polluting” is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. “Dist to Border” is the minimum distance between a firm and the corresponding township border in a given year. Firm FE, Industry-by-Year FE, and Township-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table 4: Township Merger and Emission Abatement

	(1) Log(SO2)	(2) Log(Boilers)	(3) Log(Abatement Cost)
Dist to Border	-0.0324**	-0.0113***	0.0416**
	(0.0139)	(0.0034)	(0.0175)
Constant	7.3979***	0.6189***	1.0636***
	(0.0176)	(0.0036)	(0.0223)
Mean of Dep Variable	7.357	0.607	1.117
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	70202	40529	70202
R squared	0.879	0.905	0.830

Note: Standard errors are two-way clustered at the township and industry levels. "Dist to Border" is the minimum distance between a firm and the corresponding township border in a given year. Firm FE, Industry-by-Year FE, and Township-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table 5: Township Merger and Emission Fees

	(1) Log(Emission Fee)	(2) Log(Emission Fee)	(3) Log(Emission Fee)	(4) Log(Emission Fee)
Merged	-0.1107***	-0.1381***	0.0175	0.0227
	(0.0410)	(0.0464)	(0.0230)	(0.0256)
Log(Dis10-Dis00)	0.1178***	0.0602	-0.0139	-0.0225
	(0.0413)	(0.0507)	(0.0253)	(0.0279)
Merged*Log(Dis10-Dis00)	-0.1924***	-0.1401*	0.0015	0.0003
	(0.0675)	(0.0771)	(0.0431)	(0.0472)
Constant	1.1968***	1.2127***	0.5581***	0.5592***
	(0.0251)	(0.0284)	(0.0136)	(0.0152)
Mean of Dep Variable	1.165	1.158	0.561	0.561
Border Firms	No	Yes	No	Yes
Polluting Industry	Yes	Yes	No	No
Number of Observations	10590	8241	16151	12987

Note: Standard errors are two-way clustered at the township and industry levels. Columns 1 and 2 focus on polluting industries, while Columns 3 and 4 focus on non-polluting industries. Columns 2 and 4 restrict the sample to only border firms (within 2 km of the 2000 township border). "Log(Dis10-Dis00)" is the log level of a firm's change in minimum distance to border after township merger, which measures a firm's closeness to the merging border. * significant at 10% ** significant at 5% *** significant at 1%.

Table 6: Township Merger and Production: 2km Circles

	(1) Log(Profit)	(2) Log(Output)	(3) Log(Export)
Internalized Ratio	0.0641	-0.0261	-0.0017
	(0.0831)	(0.0594)	(0.0497)
Internalized Ratio*Polluting	-0.1557***	-0.1322***	-0.1260***
	(0.0487)	(0.0321)	(0.0382)
Constant	1.6741***	4.4843***	0.8091***
	(0.0584)	(0.0445)	(0.0319)
Mean of Dep Variable	1.681	4.428	0.773
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	208663	222879	190595
R squared	0.682	0.959	0.782

Note: Standard errors are two-way clustered at the township and industry levels. "Polluting" is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. "Dist to Border" is the minimum distance between a firm and the corresponding township border in a given year. Firm FE, Industry-by-Year FE, and Township-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table 7: Township Merger and Residential Land Sales

	(1) Residential Land Development	(2) Residential Land Development
Close	-0.1417***	
	(0.0481)	
Merged	0.0125	-0.0010
	(0.0284)	(0.0324)
Close*Merged	0.1252***	0.1190*
	(0.0477)	(0.0623)
Constant	0.4512***	0.4574***
	(0.0278)	(0.0304)
Mean of Dep Variable	0.461	0.472
Township FE	Yes	No
Group FE	No	Yes
Year FE	Yes	Yes
Number of Observations	29119	28152
R squared	0.347	0.461

Note: Standard errors are clustered at the township level. "Close" is a dummy variable indicating whether an area is close to the merging border (minimum distance to border would change by more than 2 km following the township merger); "Merged" is a dummy variable indicating whether the township merger had already happened that year. * significant at 10% ** significant at 5% *** significant at 1%.

Table 8: Township Merger and Residential Land Price

	(1) Land Price	(2) Land Price	(3) Land Price	(4) Land Price
Merged	0.1184	0.1254	0.1171	0.1314
	(0.1505)	(0.1478)	(0.1364)	(0.1381)
Log(Dis10-Dis00)	-0.3863**	-0.3760**	-0.1142	-0.1151
	(0.1572)	(0.1549)	(0.2418)	(0.2449)
Merged*Log(Dis10-Dis00)	0.4089***	0.3980***	0.0282	0.0288
	(0.1517)	(0.1492)	(0.2756)	(0.2782)
Log Land Area		-0.1007***		-0.0941**
		(0.0213)		(0.0395)
Land Quality		0.1043***		0.1050*
		(0.0349)		(0.0453)
Merged*Polluting Border			0.0879	0.0609
			(0.3254)	(0.3281)
Log(Dis10-Dis00)*Polluting Border			-0.1719***	-0.1727***
			(0.0468)	(0.0385)
Merged*Log(Dis10-Dis00)*Polluting Border			0.3268***	0.3271***
			(0.0430)	(0.0400)
Constant	14.8658***	14.8257***	14.8713***	14.8331***
	(0.1512)	(0.1507)	(0.1085)	(0.1204)
Mean of Dep Variable	14.991	14.991	15.045	15.045
Township FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Transaction Type FE	Yes	Yes	Yes	Yes
Number of Observations	79115	79115	73155	73155
R squared	0.682	0.683	0.678	0.679

Note: Standard errors are clustered at the township level. "Merged" is a dummy variable indicating whether the township merger already happened in that year; "Log(Dis10-Dis00)" is the log level of a firm's change in minimum distance to border after township merger, which measures a firm's closeness to the merging border; "Polluting Border" indicates whether there is at least one polluting firm located near the merging border area. Township FE, Year FE, and Auction Type FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Appendix

Proofs

Proof of Proposition 1

Proof. From Implicit Function Theorem, we have:

$$\frac{\partial r^*}{\partial \alpha} = -\frac{\frac{\partial Surplus(r^*)}{\partial r}}{\alpha \frac{\partial^2 Surplus(r^*)}{\partial r^2} + \frac{\partial^2 Tax(r^*)}{\partial r^2}}$$

Since $\frac{\partial^2 Tax(r^*)}{\partial r^2} > 0$. Since township-level governments only receive small and indirect benefits from firm development (relative to their benefits from real estate projects), the *de facto* tax rate t is small. This means when trading off between "land revenue" and "output" loss caused by internalizing border spillovers, the marginal benefit of regulation on Surplus dominates the marginal cost of regulation on Tax. Therefore we have: $\alpha \frac{\partial^2 Surplus(r)}{\partial r^2} + \frac{\partial^2 Tax(r)}{\partial r^2} < 0$.[29]

As a result, $\frac{\partial r^*}{\partial \alpha} > 0$, which shows that the optimal level of regulation r^* is increasing in α if the condition is met. ■

Proof of Proposition 2

Proof. Suppose the regulation policy (r) is still reasonable after the increase such that the optimization problem always has an interior solution. Then the first order condition for K is:

$$\begin{aligned} \left.\frac{\partial \pi(K, L, K_E)}{\partial K}\right|_{K^*, L^*, K_E^*} &= (1-t) f_K(K^*, L^*) - p_r - r E_1(f^*, K_E^*) f_K(K^*, L^*) = 0 \\ \Rightarrow f_K(K^*, L^*) &= \frac{p_r}{1 - t - r E_1(f^*, K_E^*)} > 0 \\ \Rightarrow 1 - t - r E_1(f^*, K_E^*) &> 0 \end{aligned} \tag{8.1}$$

From the Implicit Function Theorem, we get:

$$\begin{aligned} \frac{\partial K^*}{\partial r} &= -\frac{\left.\frac{\partial^2 \pi(K, L, K_E)}{\partial K \partial r}\right|_{K^*, L^*, K_E^*}}{\left.\frac{\partial^2 \pi(K, L, K_E)}{\partial K^2}\right|_{K^*, L^*, K_E^*}} \\ &= \frac{E_1(f^*, K_E^*) f_K(K^*, L^*)}{(1 - t - r E_1(f^*, K_E^*)) f_{KK} - r E_{11}(f^*, K_E^*) f_K(f^*, K_E^*)^2} \end{aligned} \tag{8.2}$$

[29] In an extreme case, if $t = 0$ and the government maximizes only land rent, all our theoretical propositions still hold.

It's straightforward to see that the numerator is positive, as $E_1 > 0$ and $f_K > 0$. Using the result in equation 8.1, we know the first term in the denominator is negative because $f_{KK} < 0$. The second term in the denominator is positive because $E_{11} > 0$ and $f_K^2 > 0$, which means that the denominator as a whole is negative. Taken together, $\frac{\partial K^*}{\partial r} < 0$, meaning K^* decreases in r. Following the same logic, it's easy to show that L^* decreases in r. As a result, total output decreases as r increases.

We can derive r's impacts on profit by directly applying the envelope theorem:

$$\frac{\partial \pi(K^*, L^*, K_E^*)}{\partial r} = -E(f(K^*, L^*), K_E^*) < 0 \tag{8.3}$$

So the firm's profit also decreases in r .

If the magnitude of regulation (r) is so large that marginal revenue of capital becomes negative, there will not be an interior solution and the firm will shut down production. The proposition still holds in this scenario. The negative impacts of regulation policy (r) on output (f) can be shown similarly. ■

Proof of Proposition 3

Proof. The first order condition for K_E is

$$\left.\frac{\partial \pi(K, L, K_E)}{\partial K_E}\right|_{K^*, L^*, K_E^*} = -p_r - rE_2(f^*, K_E^*) = 0$$

From the Implicit Function Theorem, we get:

$$\begin{aligned}
\frac{\partial K_E^*}{\partial r} &= -\frac{\left.\frac{\partial^2 \pi(K,L,K_E)}{\partial K_E \partial r}\right|_{K^*,L^*,K_E^*}}{\left.\frac{\partial^2 \pi(K,L,K_E)}{\partial K_E^2}\right|_{K^*,L^*,K_E^*}} \\
&= -\frac{E_2(f^*, K_E^*)}{rE_{22}(f^*, K_E^*)} \\
&= \frac{p_r}{r^2 E_{22}(f^*, K_E^*)} < 0.
\end{aligned}$$

The last inequality follows because $E_{22} > 0$. This proves that firms increase emission abatement capital as r increases. ■

Proof of Proposition 4

Proof.

$$\frac{\partial E(f^*, K_E^*)}{\partial r} = E_1(f^*, K_E^*)\frac{\partial f^*}{\partial r} + E_2(f^*, K_E^*)\frac{\partial K_E^*}{\partial r} < 0.$$

The inequality holds because $\frac{\partial f^*}{\partial r} < 0$ and $\frac{\partial K_E^*}{\partial r} > 0$, as shown in Propositions 2 and 3, and $E_1 > 0$ and $E_2 < 0$ by assumptions. ∎

Proof of Proposition 5

Proof. From Proposition 1, we know that optimal enforcement increases in α, therefore larger α also means higher surplus.

The total welfare for the community is $W(r) = Surplus(r) + Tax(r)$. Let r^* be the maximizer of total welfare. Since $\frac{\partial^2 Surplus(r)}{\partial r^2} + \frac{\partial^2 Tax(r)}{\partial r^2} < 0$, this means that

$$\left.\frac{\partial W}{\partial r}\right|_{r<r^*} = \left.\frac{\partial Surplus(r)}{\partial r}\right|_{r<r^*} + \left.\frac{\partial Tax(r)}{\partial r}\right|_{r<r^*} > 0$$

Hence, the total welfare for the community increases as enforcement increases. ∎

Appendix Figures

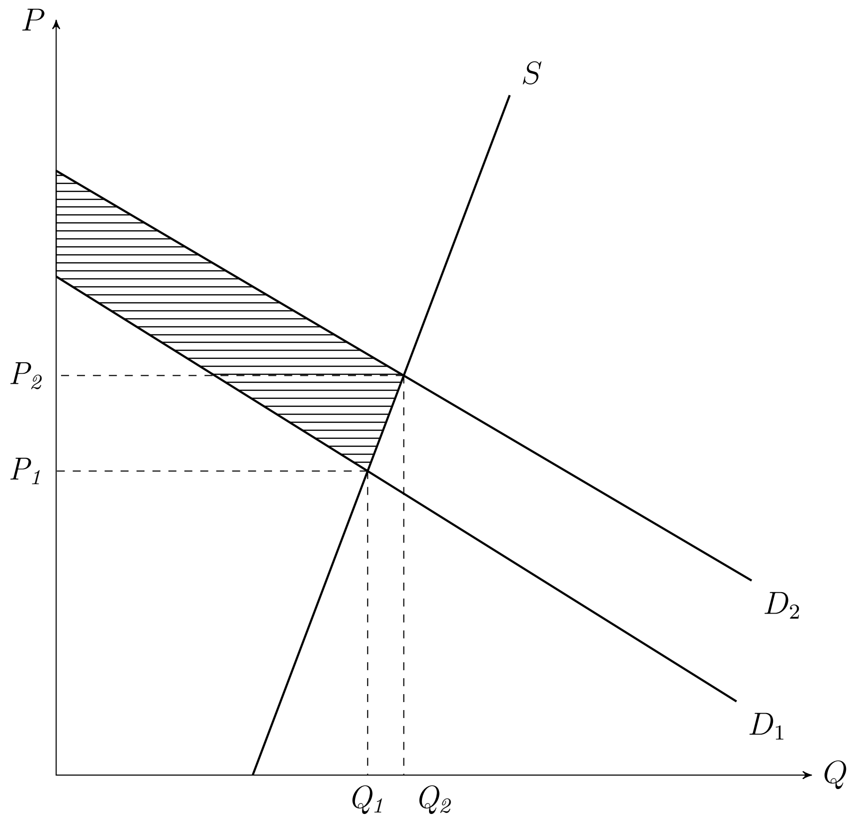

Figure A1: Welfare gains due to amenity improvements

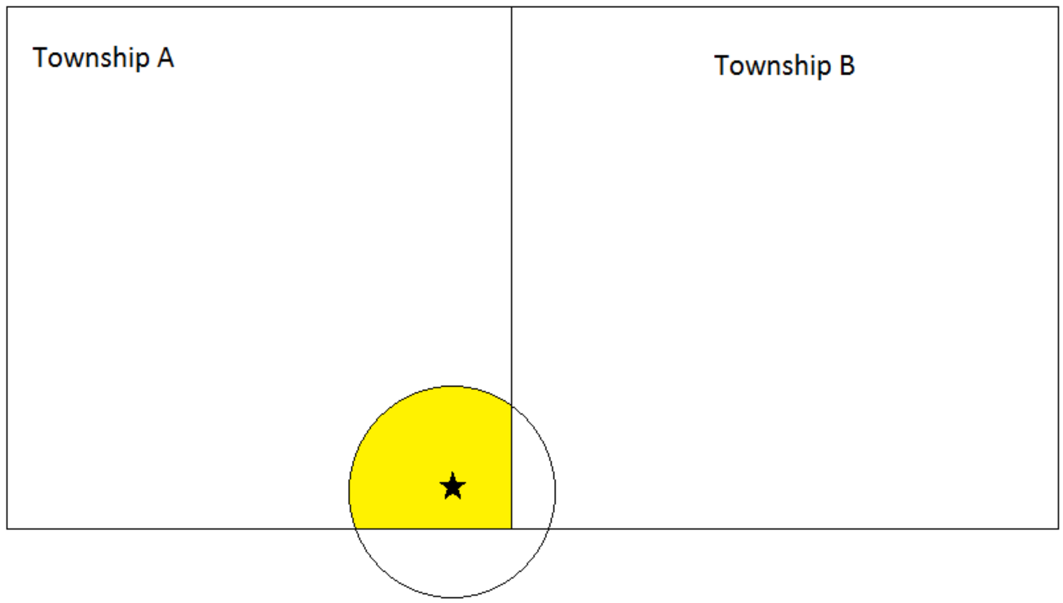

Figure A2: Measuring the Border Spillover of "Corner Firms"

Note: This figure illustrates how using "distance to border" to proxy for border spillovers would lead to measurement error for firms located near the "corner" of a township.

Appendix Tables

Table A1: Polluting vs Non-Polluting Industries

Polluting Industries		Non-Polluting Industries	
Industry	Code	Industry	Code
Mining and Washing of Coal	6	Forestry	2
Mining and Processing of Ferrous Metal Ores	8	Extraction of Petroleum and Natural Gas	7
Mining and Processing of Non-metallic Mineral	10	Mining and Processing of Non-ferrous Metal Ores	9
Fermentation	14 (6)	Agricultural and Sideline Food Processing	13
Beverage Manufacturing	15	Food Manufacturing	14
Textiles Mills	17	Tobacco Manufacturing	16
Leather, Fur and Related Products Manufacturing	19	Wearing Apparel and Clothing Accessories Manufacturing	18
Pulp and Paper Manufacturing	22 (1, 2)	Wood and Bamboo Products Manufacturing	20
Petrochemicals Manufacturing	25	Furniture Manufacturing	21
Chemical Products Manufacturing	26	Paper Products Manufacturing	22
Medicine Manufacturing	27 (1, 2, 4)	Printing and Reproduction of Recorded Media	23
Chemical Fibers Manufacturing	28	Education and Entertainment Articles Manufacturing	24
Non-Metallic Mineral Products Manufacturing	31	Medical Goods Manufacturing	27
Iron and Steel Smelting	32 (1, 2)	Rubber Products Manufacturing	29
Non-Ferrous Metal Smelting	33 (1)	Plastic Products Manufacturing	30
Fossil-Fuel Power Station	44 (1)	Basic Metal Processing	32
		Non-Ferrous Metal Processing	33
		Fabricated Metal Products Manufacturing	34
		General Purpose Machinery Manufacturing	35
		Special purpose Machinery Manufacturing	36
		Transport Equipment Manufacturing	37
		Electrical Equipment Manufacturing	39
		Computers and Electronic Products Manufacturing	40
		General Instruments and Other Equipment Manufacturing	41
		Craftworks Manufacturing	42
		Renewable Materials Recovery	43
		Electricity and Heat Supply	44
		Gas Production and Supply	45
		Water Production and Supply	46

Table A2: Township Merger and Political Incentives of Prefectural Leaders

	(1) Log(Profit)	(2) Log(Output)	(3) Log(Export)
Dist to Border	0.0080	-0.0023	0.0021
	(0.0130)	(0.0070)	(0.0132)
Dist to Border*Polluting	-0.0195**	-0.0083	-0.0194*
	(0.0094)	(0.0079)	(0.0101)
Dist*Incentive	0.0071	0.0044	0.0180
	(0.0073)	(0.0062)	(0.0113)
Polluting*Incentive	0.0727**	-0.0148	0.1176***
	(0.0296)	(0.0195)	(0.0255)
Dist*Polluting*Incentive	-0.0211**	-0.0085	-0.0244**
	(0.0099)	(0.0064)	(0.0111)
Constant	1.7492***	5.4925***	0.9267***
	(0.0161)	(0.0093)	(0.0175)
Mean of Dep Variable	1.764	5.484	0.950
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	219250	216191	206485
R squared	0.701	0.823	0.802

Note: Standard errors are two-way clustered at the township and industry levels. "Polluting" is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. "Dist to Border" is the minimum distance between a firm and the corresponding township border in a given year. "Incentive" is a dummy variable that equals 1 if the local prefectural party secretary is younger than 57 years old in that year, and 0 otherwise. Firm FE, Industry-by-Year FE, and Township-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table A3: Pre-Merger Trend: Production

	(1) Log Tot Profit	(2) Log Ouput Value	(3) Log Tot Export
Panel A: One Year Before			
F1.Distance	-0.0443	0.0236	0.0464
	(0.0592)	(0.0384)	(0.0613)
F1.Distance*Polluting	-0.0098	-0.0082	-0.0279
	(0.0426)	(0.0362)	(0.0599)
Panel B: Two Years Before			
F2.Distance	-0.0101	0.0132*	-0.0086
	(0.0137)	(0.0068)	(0.0143)
F2.Distance*Polluting	0.0018	-0.0087	0.0050
	(0.0155)	(0.0066)	(0.0133)
Mean of Dep Variable	1.427	5.341	0.615
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	61341	64596	56383
R squared	0.741	0.877	0.845

Note: Standard errors are two-way clustered at the township and industry levels. "Polluting" is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. The main explanatory variables are one-year and two-year leads of a firm's minimum distance to township border. Only the pre-merger sample is used. Firm FE, Industry-by-Year FE, and Township-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table A4: Pre-Merger Trend: Emission

	(1) Log(SO2)	(2) Log(Boilers)	(3) Log(Abatement Cost)
Panel A: One Year Before			
F1.Distance	0.0664	0.0153	0.0464
	(0.1631)	(0.0141)	(0.0381)
Panel B: Two Years Before			
F2.Distance	0.0548	0.0069	-0.0099
	(0.0571)	(0.0070)	(0.0199)
Mean of Dep Variable	8.653	0.648	2.008
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	16209	15056	16209
R squared	0.899	0.924	0.883

Note: Standard errors are two-way clustered at the township and industry levels. The main explanatory variables are one-year and two-year leads of a firm's minimum distance to township border. Only the pre-merger sample is used. Firm FE, Industry-by-Year FE, and Township-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table A5: Alternative Bandwidths

	(1) Log(Profit)	(2) Log(Output)	(3) Log(Export)
Panel A: 1.5-km Bandwidth			
Dist to Border	0.0079	-0.0009	0.0118
	(0.0135)	(0.0085)	(0.0164)
Dist to Border*Polluting	-0.0277**	-0.0173**	-0.0228*
	(0.0121)	(0.0085)	(0.0114)
Number of Observations	191177	210645	174790
Panel B: 3-km Bandwidth			
Dist to Border	0.0184*	-0.0002	0.0160
	(0.0105)	(0.0077)	(0.0118)
Dist to Border*Polluting	-0.0310***	-0.0167**	-0.0346***
	(0.0089)	(0.0081)	(0.0126)
Number of Observations	278524	305936	254969
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes

Note: Standard errors are two-way clustered at the township and industry levels. "Polluting" is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. The main explanatory variables are one-year and two-year leads of a firm's minimum distance to township border. In Panel A, the sample consists of firms located within 1.5 km of the 2000 township border; in Panel B, the sample consists of firms located within 3 km of the 2000 township border. Firm FE, Industry-by-Year FE, and Township-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table A6: Arcsinh Results: Firm Production

	(1) asinh(Profit)	(2) asinh(Output)	(3) asinh(Export)
Dist to Border	0.0014	-0.0022	0.0056
	(0.0298)	(0.0039)	(0.0171)
Dist to Border*Polluting	-0.0736**	-0.0098**	-0.0588***
	(0.0321)	(0.0042)	(0.0188)
Constant	5.1021***	11.0018***	2.6818***
	(0.0344)	(0.0045)	(0.0195)
Mean of Dep Variable	5.068	10.994	2.661
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	265611	246047	248429
R squared	0.547	0.831	0.792

Note: Standard errors are two-way clustered at the township and industry levels. "Polluting" is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. Firm FE, Industry-by-Year FE, and Prefecture-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table A7: Arcsinh Results: Firm Productivity

	(1) asinh(VA)	(2) asinh(Labor)	(3) asinh(Capital)
Dist to Border	-0.0046	-0.0056	0.0077
	(0.0149)	(0.0035)	(0.0050)
Dist to Border*Polluting	-0.0308*	0.0009	-0.0082
	(0.0175)	(0.0038)	(0.0057)
Constant	9.1690***	5.4826***	9.2220***
	(0.0159)	(0.0040)	(0.0053)
Mean of Dep Variable	9.150	5.476	9.227
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	179199	250935	170797
R squared	0.544	0.835	0.908

Note: Standard errors are two-way clustered at the township and industry levels. "Polluting" is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. Firm FE, Industry-by-Year FE, and Prefecture-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table A8: Inter-Firm Spillovers

	(1) Log(Profit)	(2) Log(Output)	(3) Log(Export)
Dist to Border	0.0069	-0.0012	0.0077
	(0.0083)	(0.0059)	(0.0080)
Dist to Border*Polluting	-0.0321***	-0.0157*	-0.0249**
	(0.0110)	(0.0090)	(0.0108)
Constant	1.8210***	4.6494***	0.9815***
	(0.0117)	(0.0058)	(0.0087)
Mean of Dep Variable	1.814	4.640	0.979
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Prefecture-Year FE	Yes	Yes	Yes
Number of Observations	239305	263668	218464
R squared	0.674	0.951	0.786

Note: Standard errors are two-way clustered at the township and industry levels. "Polluting" is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. Firm FE, Industry-by-Year FE, and Prefecture-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.

Table A9: Controlling for Population Density

	(1) Log Tot Profit	(2) Log Ouput Value	(3) Log Tot Export
Distance	0.0096	-0.0022	0.0071
	(0.0115)	(0.0076)	(0.0125)
Distance*Polluting	-0.0283***	-0.0159**	-0.0257**
	(0.0089)	(0.0077)	(0.0108)
Population (2km Circle)	0.0005	-0.0001	0.0007
	(0.0005)	(0.0003)	(0.0005)
Constant	1.8050***	4.6583***	0.9682***
	(0.0237)	(0.0115)	(0.0176)
Mean of Dep Variable	1.817	4.646	0.985
Firm FE	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes
Township-Year FE	Yes	Yes	Yes
Number of Observations	237032	261262	216473
R squared	0.697	0.954	0.794

Note: Standard errors are two-way clustered at the township and industry levels. "Polluting" is a dummy variable indicating whether a firm belongs to one of the 16 heavily polluting industries as defined by the MEP. Firm FE, Industry-by-Year FE, and Prefecture-by-Year FE are controlled for. * significant at 10% ** significant at 5% *** significant at 1%.