Leveraging Political Incentives for Environmental Regulation: Evidence from Chinese Manufacturing Firms*

Guojun He Shaoda Wang Bing Zhang

Abstract

This paper estimates the effect of environmental regulation on firm productivity using a spatial regression discontinuity design implicit in China's water quality monitoring system. Because water quality readings are important for political evaluations, and the monitoring stations only capture emissions from their upstream regions, local government officials are incentivized to enforce tighter environmental standards on firms immediately upstream of a monitoring station, rather than those immediately downstream. Exploiting this discontinuity in regulation stringency with novel firm-level geocoded emission and production datasets, we find that upstream polluting firms face a 27% reduction in Total Factor Productivity (TFP), and a 48% reduction in emission intensity, as compared to their downstream counterparts. We find that the discontinuity in TFP does not exist in non-polluting industries, only emerged after the government explicitly linked political promotion to water quality readings, and was entirely driven by prefecture cities with career-driven leaders. Linking the TFP estimate with the emission estimate, a back of the envelope calculation indicates that China's current water-pollution abatement target leads to an annual economic loss of more than 30 billion dollars.

Keywords: political incentives; total factor productivity; water quality monitoring; water pollution; environmental policy; COD

JEL: Q56, Q58, O13, O44, D24

* He: Hong Kong University of Science and Technology; Email: gjhe@me.com. Wang: University of California, Berkeley; Email: shaoda@berkeley.edu. Zhang: Nanjing University; Email: zhangb@nju.edu.cn. We thank Ariel BenYishay, Peter Berck, Fiona Burlig, Steve Cicala, Alain de Janvry, Andrew Foster, Eyal Frank, Marco Gonzalez-Navarro, Michael Greenstone, Koichiro Ito, Ryan Kellogg, Arik Levinson, Jeremy Magruder, Mushfiq Mobarak, Albert Park, Jay Shimshack, Elisabeth Sadoulet, Joseph Shapiro, Heiwai Tang, Reed Walker, Chris Woodruff, Ye Wang, Rudai Yang, Noam Yuchtman, and Jane Zhang for their valuable suggestions. We also appreciate comments from seminar participants at UChicago, LSE, UC Berkeley, MIT, HKUST, Duke Kunshan, SHUFE, GRIPS, and conference participants at NBER Chinese Economy Working Group Meeting, ABCDE, WCERE, Occasional Workshop, PacDev, MIEDC, Econometric Society Asian Meeting, CES, Oxford PEDL Workshop, Columbia IPWSD, and Camp Resources. Chun Wai Cheung, Ziteng Lei, and Chenxinyu Zhao offered excellent research assistance. Financial supports from University Grants Committee of HK (General Research Fund, Project No. 16500617) and PEDL Exploratory Grant are gratefully acknowledged.

I. Introduction

The question of whether environmental regulation hinders factor productivity has long been important and controversial. On the one hand, neoclassical models suggest that environmental regulations will increase production costs, lead to a reallocation of labor and capital, and reduce the competitiveness of firms. On the other hand, proponents of environmental protection argue that stringent regulations provide incentives for polluters to invest in cleaner and less costly technologies to reduce pollution, which can in turn enhance productivity.[1] This debate entails particularly significant policy ramifications for rapidly-developing countries such as China and India, where the economies still rely heavily on manufacturing output in dirty industries, causing billions of people to suffer from extreme levels of water and air pollutions (Greenstone and Hanna, 2014; Ebenstein et al., 2017).

However, while there exists a large empirical literature investigating the relationship between environmental regulation and productivity (e,g, Jaffe et al. 1995; Berman and Bui, 2001; Greenstone, 2002; Greenstone et al., 2012), these works mostly focus on developed countries, and much less is known about the developing world, despite tremendous potential policy relevance. Developing countries may face a very different regulation-productivity tradeoff partly because they fall behind in industrial structures and technologies. But more importantly, developing countries tend to have weaker formal institutions, where local bureaucrats have drastically different incentives and constraints compared to those in the developed world.[2] Understanding the political economy of regulation is thus critical for the design of effective environmental programs (Acemoglu and Robinson, 2013; Greenstone and Jack, 2015).

In this study, we focus on China, the largest manufacturer and emitter in the world, where high-powered political incentives were leveraged to help alleviate pollution. Specifically, we exploit the political incentives implicit in China's surface water quality monitoring system, and estimate how tighter water emission controls affect the production and emission activities of Chinese manufacturing firms. Because water quality monitors can only pick up pollution information from upstream regions, and because the readings from these monitors are

[1] Notably, Porter (1991) argues that, if one country adopts more stringent environmental standards than a competitor, firms in this country will invest more in clean innovations, which in turn will enhance the country's growth. Evidence in favor of Porter's hypothesis is summarized in a recent review by Ambec et al. (2013).

[2] It has been shown that local bureaucrats in the developing world have considerable discretions in public service delivery (Martinez-Bravo 2014; Munshi and Rosenzweig 2015; He and Wang, 2017).

important for political evaluations, local government officials have particularly strong incentives to abate the emissions from upstream firms. As a result, within a small neighborhood around a water quality monitoring station, upstream firms typically face tighter environmental standards than downstream firms. By focusing on a narrow geographic band that only stretches a few kilometers upstream and downstream of each surface water monitoring station, we are able to causally identify the impacts of water quality controls on manufacturing productivity.

Exploiting this spatial discontinuity in regulation stringency embedded in China's water quality monitoring system, we find that polluting firms in the near upstream of monitoring stations have a 27% lower TFP, and a 48% lower emission intensity, as compared to their near downstream counterparts. To put the magnitude in context, the TFP gap is on average equivalent to upstream polluting firms losing 2 years of productivity growth.[3] Importantly, we show that the discontinuity is not driven by the endogenous location choices of the monitoring stations, nor by the endogenous sorting of polluting firms around monitoring stations. Instead, a rich set of evidence suggests that the baseline findings are indeed driven by upstream firms receiving tighter regulation: (1) the upstream-downstream TFP gap only exists in polluting industries, and does not apply to non-polluting industries; (2) the discontinuity only emerged after the central government started to link water quality readings to political promotions; and (3) only firms within a few kilometers upstream are regulated, as emissions from the further upstream would dissipate quickly and have small influence on water quality readings.

In addition to the baseline economic impacts, we conduct a series of additional analyses to understand the political economy of environmental regulation in China, which has largely been a black box to both academia and the public. First, we find that upstream firms pay more emission fees and taxes than do downstream firms, even though they actually have lower levels of outputs and emissions. This implies that local governments use double standards in environmental regulation. Second, the baseline effect is completely driven by prefecture cities with politically motivated leaders, and there is no discontinuity at all when the leaders do not have promotion incentives. Third, we find that the discontinuity is more salient around "automatic" monitoring stations, whose data are less susceptible to local political influence, suggesting potential manipulation of environmental data among those traditional "manual" stations.

[3] With the same dataset and same method for TFP estimation, Brandt et al. (2012) finds that the average TFP growth among Chinese manufacturing firms in 2005 was 14%.

Our findings contribute to the ongoing debate on the economic costs of environmental regulation in several important ways. First, as mentioned, most studies to date have focused on developed countries (e.g., Jaffe et al., 1995; Henderson 1996; Becker and Henderson, 2000; Berman and Bui, 2001; Greenstone, 2002; Walker, 2011; Greenstone, List, and Syverson, 2012; Ryan, 2012; Kahn and Mansur, 2013; Walker, 2013). In this paper, we investigate China, the largest developing country in the world, and highlight the significant economic costs of environmental regulation in a rapidly-growing manufacturing economy. Moreover, most of the existing literature on environmental regulation focuses on air pollution, and the few exceptions that look at water regulation mostly only evaluate the environmental consequences of water regulation (Greenstone and Hanna, 2014; Keiser and Shapiro, 2018), leaving a gap in knowledge regarding the economic costs of water regulation, which our paper fills in.

Second, our paper adds to the growing literature on the political economy of pollution (List and Sturm, 2006; Burgess et al., 2012; Kahn et al., 2015; Lipscomb and Mobarak, 2017; Jia, 2017). Specifically, our analyses on the political economy of water quality monitoring shed light on how environmental regulations are implemented at the local level, which could guide future policy design in important ways. We find huge variations in the effectiveness of environmental regulation and the spatial allocation of regulatory burdens, which can largely be explained by regional differences in the incentives and constraints of local government officials. We also find that the regulation effect is stronger when improved technology makes it harder to manipulate water readings, suggesting that local politicians have incentives to misreport environmental data in the absence of a precise monitoring practice.

Third, the detailed firm-level data also allow us to explore the channels through which firms are affected, and helps us understand how different types of firms respond to regulations. For instance, we show that upstream firms have to make substantially more investments in machineries (abatement capital) to cope with tighter regulation. Heterogeneous analyses reveal that the TFP loss is almost exclusively experienced by private Chinese firms, so tightening environmental regulations in the future is likely to damage the competitiveness of private Chinese firms rather than state-owned or foreign firms. Tests on sorting suggest that to avoid the large impact of regulation on productivity, upstream firms tend to relocate in the long run. These results imply a redistribution of production, income, environmental quality and social welfare between upstream and downstream regions. This study therefore also speaks to several lines of literature on the impacts of environmental regulation on production (Becker and Henderson, 2000), employment (Greenstone, 2002; Walker, 2011), plant location choice (List

et al., 2003), income and total welfare (Ryan, 2012), and foreign direct investment (FDI) (Fredriksson et al., 2003; Hanna, 2010; Cai et al., 2016).

Finally, understanding firms' abatement costs and responses to regulation is critical for optimal policy design. Combining the TFP estimates with the emission estimates, we can calculate the economic costs of tightening water pollution regulations. We estimate that a 10% reduction in chemical oxygen demand (COD) emissions leads to a 2.49% decrease in TFP, and China's target of reducing total COD emissions by 10% between 2016 and 2020 would cause a total loss in industrial output value of 990 billion Chinese yuan (159 billion US dollars for five years) under current policy design and enforcement practices.

The rest of this paper is structured as follows. Section II describes the institutional background and research design. Section III discusses the data and presents descriptive statistics. Section IV presents the baseline results. Section V examines the channels, explores the political economy of environmental regulation, and tests whether emission measures also differ across the monitoring stations. Section VI interprets the results and benchmarks their economic significance. Section VII concludes the paper.

II. Research Design and Empirical Setup

A. Water Quality Monitoring and Water Pollution Controls in China

As the world's largest developing country, China faces a variety of pressing environmental challenges, including prevalent water and air pollution. According to the World Bank (2007), roughly 70 percent of China's rivers were polluted and contained water deemed unsafe for human consumption (at the time of that report). Poor surface water quality has driven policymakers to propose regulations to protect water bodies and reverse the process of degradation.

To gather surface water quality information, a national water quality monitoring system was established in the 1990s to monitor surface water quality in major river segments, lakes and reservoirs, known as the "National Environmental Quality Monitoring Network-Surface Water Monitoring System" (NEQMN-SWMS). At the beginning, the monitoring system was intended mainly for scientific rather than regulatory purposes, and most of the station-level monitoring data were kept confidential by the government. No strict emission abatement targets were set by the Chinese government between 1990s to early 2000s because economic

growth was considered the country's priority. As a result, along with China's rapid economic growth, the country witnessed severe degradation of its ecological systems.

In 2002, Hu Jintao became the new political leader of China, taking power from Jiang Zemin, and held office until 2012. Given the country's mounting environmental challenges, the new president started to emphasize the importance of seeking a balance between economic growth and environmental sustainability. Notably, in 2003, President Hu proposed the "Scientific Outlook of Development" (SOD),[4] which sought integrated sets of solutions to economic, environmental and social problems, opening an era of environmental regulation.

Responding to the SOD slogan, the Ministry of Environmental Protection (MEP) increased its efforts to resolve the water pollution issue. In 2003, the MEP issued an updated version of NEQMN-SWMS and spread the "Technical Specification Requirements for Monitoring of Surface Water and Wastewater" to local governments, explicitly highlighting the importance of surface water quality monitoring. The new policy documents imposed emission reduction targets for all the 419 state-controlled stations at the time, and also started a trend of automation of these stations. To introduce public monitoring of environmental quality, the water quality readings from all state-controlled stations also became available for the public and were published in various environmental yearbooks starting from 2003.

During President Hu's political regime, the importance of clean surface water was emphasized and the central government adopted a target-based abatement system to control environmental pollutants. In particular, during the 11th Five-Year Plan (2006–2010), the emission abatement targets included (but are not limited to): (1) reducing COD emissions by 10% (from 141.4 million tons in 2005 to 127.3 million tons in 2010), (2) reducing the percentage of monitored water sections failing to meet Grade V National Surface Water Quality Standards from 26.1% in 2005 to 22% by 2010, and (3) increasing the ratio of monitored water sections (of the seven main bodies of water in China) meeting Grade III National Surface Water Quality Standards from under 41% in 2005 to 43% by 2010.[5] With these targets, the central government then allocated binding abatement requirements to each province, and provincial governors were required to sign individual responsibility contracts with the central government, documenting their emission abatement plans in detail. Provincial governors further assigned strict abatement mandates to prefectures and counties and used local

[4] SOD can also be translated as the "Scientific Development Concept" or the "Scientific Development Perspective."

[5] Source: http://www.mep.gov.cn/gzfw_13107/zcfg/fg/gwyfbdgfxwj/201605/t20160522_343144.shtml

environmental performance along with other criteria to assess and promote local government officials. Water pollution control thus became an important political task for local government officials.

Because rivers flow from higher to lower elevation, water quality monitoring stations can only capture emissions from their upstream areas, but not from downstream areas. Under the new political regime, local officials would have strong incentives to enforce tighter environmental regulations in upstream regions than in downstream regions. We exploit this spatial discontinuity and estimate the causal impact of tighter water pollution regulation on productivity. Because the Chinese government did not enforce stringent industrial pollution controls until 2003, we expect that, if water quality monitoring indeed influences firm productivity, this effect should be weaker or non-existent before 2003, and become stronger afterward.

B. Location Choice of Water Quality Monitoring Stations

The purpose of establishing a water quality monitoring network is to achieve a comprehensive understanding of the country's surface water quality. The monitoring system covers the country's major rivers, lakes, and reservoirs. A monitoring station should be set in a way that can is spatially representative to its neighborhood water bodies and can properly reflect changes in water pollutants over time. Consequently, the locations of the monitoring stations were chosen based mainly on hydrological considerations.

According to the MEP, the monitoring stations must be placed in rivers with steady flows, wide water surfaces, and stable river beds, and must avoid stagnant water areas, backwater areas, sewage outfalls, rapids and shallow water. The MEP also requires that monitoring stations be established to serve "long-term" purposes, ensuring that short-term needs (such as avoiding or targeting pollution from a specific region or a specific firm) cannot be accommodated.

An important feature of station placement is that the MEP requires monitoring stations to be built close to hydrological stations, which enables the government to combine hydrological parameters with water quality information. Most hydrological stations were built in the 1950s-1970s and are used to collect meteorological and hydrological data.

In this paper, we focus on the state-controlled surface water quality monitoring stations. State-controlled stations are established and supervised by the MEP and the State Council of China. The water quality readings from the state-controlled stations are reported directly to the

MEP to ensure data quality. Yearly average water quality readings from the stations are reported in the environmental yearbooks and the central government used these data to assess the environmental performance of local governments.

Aside from state-controlled stations, there are also local water quality monitoring stations and special stations designed to monitor the emissions of major polluters. The special monitoring stations are placed immediately downstream from the polluter to monitor its environmental performance. We do not have data for these types of stations.

C. Research Design and Econometric Model

We exploit the spatial discontinuity in regulation stringency around water monitoring stations to estimate the causal effect of regulation on TFP. The distance between a firm and a monitoring station serves as the running variable. We examine whether firms located immediately upstream from the monitoring station have lower productivity than adjacent downstream firms. This empirical strategy is similar in spirit to recent works that also exploits the flow of pollution along rivers for identification (Kaiser and Shapiro, 2017; Lipscomb and Mobarak, 2017), but is novel in that it utilizes a unique spatial discontinuity setting around the monitoring stations and explores the political incentives behind the regulation.

Specifically, the identifying assumption of our research design is that, due to spatial adjacency, firms located immediately upstream and downstream of monitoring stations should be balanced *ex ante* along various dimensions but will differ from each other only because upstream firms become more tightly regulated. The discontinuity can be estimated by both parametric and non-parametric approaches. Gelman and Imbens (2017) show that the parametric RD approach, which uses a polynomial function of the running variable as a control in the regression, tends to generate RD estimates that are sensitive to the order of the polynomial and have some other undesirable statistical properties. As a result, estimators based on local linear regression or other smooth functions are often preferred, because they can assign larger weights to observations that are closer to the threshold and therefore can produce more accurate estimates. We thus focus on a local linear approach, which can be estimated by the following equation:

$$(1)\quad TFP_{ijk} = \alpha_1 Down_{ijk} + \alpha_2 Dist_{ijk} + \alpha_3 Down_{ijk} Dist_{ijk} + u_j + v_k + \varepsilon_{ijk}$$

$$s.t.\ \ -h \leq Dist_{ijk} \leq h$$

where TFP_{ijk} is the total factor productivity of firm i in industry j around monitoring station k. $Down_{ijk}$ is an indicator variable that equals 1 if firm i (in industry j) is downstream from monitoring station k, and 0 otherwise. $Dist_{ijk}$ measures the distance between firm i and monitoring station k, and h is the bandwidth length (i.e., the acceptable distance from the discontinuity for sample inclusion).

To account for industry and location-specific TFP determinants in the non-parametric estimations, we control for industry and monitoring station fixed effects u_j and v_k. The model essentially compares upstream and downstream firms in the same industry around the same monitoring station. The estimation of this non-parametric RD model with fixed effects is implemented using the two-step approach suggested by Lee and Lemieux (2010), where industry and station fixed effects are absorbed by running an OLS regression of TFP on a set of industry and station-specific dummies, and then apply the non-parametric estimations on the residual TFP obtained from OLS estimation.[6]

The choice of the optimal bandwidth h involves balancing the conflicting goals of focusing comparisons near the monitoring stations, where the identification assumption is most likely to be satisfied, and providing a large enough sample for reliable estimation. In this study, we rely on an MSE-optimal bandwidth h proposed by Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, Farrell (forthcoming), and experiment with various kernel weighting functions to ensure robustness.

The standard error is clustered at the monitoring station level to deal with the potential spatial correlation of the error term, as suggested by Cameron and Miller (2015). We also try two-way clustering at both the industry and the monitoring station levels and get quantitatively similar standard errors.

As a way to check the robustness, we also estimate a parametric RD model:

$$(2) \qquad TFP_{ijk} = \alpha_1 Down_{ijk} + f(Dist_{ijk}) + Down_{ijk} f(Dist_{ijk}) + u_j + v_k + \varepsilon_{ijk}$$

where TFP_{ijk} is the total factor productivity of firm i in industry j around monitoring station k. $f(Dist_{ijk})$ is a polynomial in distance between firm i in industry j and monitoring station k. The polynomial function is interacted with the treatment dummy to allow flexible functional form on both sides of the cutoff, and u_j and v_k are industry and station fixed effects.

[6] Lee and Lemieux (2010) argue that, if there is no violation of the RD assumption that unobservables are similar on both sides of the cutoff, using a residualized outcome variable is desirable because it improves the precision of estimates without violating the identification assumption.

III. Data and Summary Statistics

A. Data

Our analyses are based on several datasets that together provide comprehensive information on the socio-economic conditions of townships, the production and performance of industrial firms, and emissions from heavy polluters centered around the monitoring stations.

Water Quality Monitoring Stations

We collect data from water quality monitoring stations from surface water quality reports in various environmental yearbooks from 2003-2010, which include the China Environmental Yearbooks, China Environmental Statistical Yearbooks, and China Environmental Quality Statistical Yearbooks. Data available in more than two different sources are cross-validated. The number of state-controlled monitoring stations varied slightly between years in these reports, ranging from 400 to 500 stations. We geocoded all the water quality monitoring stations.

Annual Survey of Industrial Firms Database

Our firm-level TFP is calculated using data from the Annual Survey of Industrial Firms (ASIF) from 2000 to 2007. The ASIF data include all the private industrial enterprises with annual sales exceeding 5 million Chinese yuan and all the state-owned industrial enterprises (SOEs). The data are collected and maintained by the National Bureau of Statistics (NBS) and contain a rich set of information obtained from the accounting books of these firms, such as inputs, outputs, sales, taxes, and profits.

The detailed production information allows us to construct TFP measures for each firm in each year. There are several approaches to estimating firm-level TFP and each requires different assumptions (Van Biesebroeck, 2007). In this paper, we use the consistent semi-parametric estimator suggested by Olley and Pakes (1996). The Olley-Pakes method addresses the simultaneity and selection biases in estimating TFP and has been the most widely-used method for the investigation of Chinese firms' productivity in the literature (see for example, Brandt et al., 2012; Yang, 2015). Using Olley-Pakes TFP estimator ensures that our estimates can be compared with the previous ones. The details of estimating TFP using the Olley-Pakes method are discussed in Appendix A. For robustness checks, we also construct alternative TFP

measures based on other estimators, such as the Ackerberg et al. (2015) approach and the naïve labor/capital productivity measures.[7]

The ASIF data have been used in many studies investigating Chinese firms. A well-known issue is that the data contain outliers. We follow standard procedures documented in the literature to clean the data. We first drop observations with missing key financial indicators or with negative values for value added, employment, and fixed capital stock. We then drop observations that apparently violate accounting principles: liquid assets, fixed assets, or net fixed assets larger than total assets; or current depreciation larger than cumulative depreciation. Finally, we trim the data by dropping observations with values of key variables outside the range of the 0.5^{th} to 99.5^{th} percentile.[8]

The ASIF data have detailed address information for sampled firms in each year. We geocode the location of the 952,376 firms that appeared in the sample and then compute precise distance measures between each firm and its closest water quality monitoring station.

Because our research design is fundamentally cross-sectional, in the baseline analysis, we collapse the multi-year data into a cross-section and apply the RD estimators to it. The interpretation of the coefficients is therefore an average effect that persists for years. To fully utilize the dynamic structure, however, we also apply non-parametric RD estimators to different years and examine how the discontinuity changes over time.

Environmental Survey and Reporting Database

To investigate whether water quality monitoring indeed reduces water-related emissions, we collect firm-level emission data from China's Environmental Survey and Reporting (ESR) database, which is managed by the MEP.

The ESR database is the most comprehensive environmental dataset in China that provides firm-level (polluting-source level) emissions for various pollutants. The ESR database monitors 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. In this study, we keep only the ESR firms that are in the same polluting industries as the ASIF firms.

---

[7] Note that the TFP measure we construct only accounts for private factors of production. Environmental goods are not considered as inputs, which is sometimes done in the environmental economics literature.

[8] More details about the construction and cleaning processes of the ASIF data can be found in Hsieh and Klenow (2009), Songet et al. (2011), Yu (2015), and Huang et al. (forthcoming).

The sampling criteria in the ESR database is based on the cumulative distribution of emissions in each county. Polluting sources are ranked based on their emission levels of different "criteria pollutants," and those jointly contributing to the top 85% of total emissions in a county are included in the database. In this study, we use ESR data between 2000 and 2007, the same period as the ASIF database.

During our sample period, the "criteria pollutants" changed over time. In 2000, only chemical oxygen demand (COD) emissions and sulfur dioxide (SO_2) were "criteria pollutants." Polluting sources included in the database were therefore chosen based on their contributions to COD emissions or SO_2 emissions. In 2007, ammonia nitrogen (NH_3) and NOx also became "criteria pollutants."

Among all the pollutants, COD is most relevant to this study. COD is a widely-used water quality indicator that measures the amount of oxygen required to oxidize soluble and particulate organic matter in water.[9] It assesses the effect of discharged wastewater on the water environment. Higher COD levels mean a greater amount of oxidizable organic material in the sample, which reduces dissolved oxygen levels. A reduction in dissolved oxygen can lead to anaerobic conditions, which are deleterious to higher aquatic life forms.

We focus on COD emissions because COD is the first water-related "criteria" pollutant used by the MEP, and the government explicitly set a 10% abatement target for COD emissions in the 11th Five-Year Plan. We also corroborate the findings on COD emissions by looking at the amount of wastewater discharge.

Like the ASIF, this dataset also includes detailed address information. We therefore geocode all the ESR firms and compute their distances to the nearest monitoring sites. The dataset allows us to construct total emission levels and emission intensity measures (emission levels divided by total output value) for large polluters in each county.

Unlike the ASIF, however, we are less confident about the quality of ESR data. While later we will show that the results for emissions are consistent with our story, we must acknowledge that the quality of the ESR data is relatively poor. For instance, we find that there exist a large number of missing values and zeros in the ESR data, and location, industry, firm, and year fixed effects combined can only explain a small portion of firms' emissions. These patterns seem suspicious to us. In fact, this data quality issue can be the reason that the government

[9] For example, COD abatement is used by the central government of China as a key performance indicator to assess local government efforts in environmental protection. In the 10th and 11th Five-Year Plans (2001-2005 and 2006-2010), COD was used as a primary criterion (along with ammonia-nitrogen) to set national abatement targets and conduct performance appraisals.

relies on the readings of monitoring stations, instead of the ESR data collected from polluting firms, to measure the effectiveness of pollution control, although the latter would be more powerful and better-targeted in a world without data manipulation. Given such, in subsequent analyses, we primarily focus on outcomes from the ASIF dataset and use the emission results as supportive evidence.

Township-level Socio-economic Data

The National Bureau of Statistics (NBS) conducts the "Township Conditions Survey (TCS)" on an annual basis. It is a longitudinal survey that collects township-level socio-economic data for all the townships in China. We have access to the TCS data for 20 provinces in 2002 and use the township-level data to assess similarities between upstream and downstream townships.

Geo-data

We obtained township-level GIS boundary data in 2010 from the Michigan China Data Center. We use GIS data of China's water basin system from the Ministry of Water Resources. We use GIS elevation data to identify upstream and downstream relationships. These GIS datasets are then matched to our geocoded township and firm datasets.

B. Data Matching

The data we have compiled are, to our knowledge, the most comprehensive and disaggregated collection ever assembled on water pollution and firm-level economic and environmental performance in China. The matching process involves several steps and is illustrated in Figure 1.

First, we only keep river monitoring stations, because lake and reservoir stations would not allow us to identify the upstream-downstream relationship. Then, we put a layer of the water basin system on the township GIS map, and only keep townships that have at least one river passing through it. Then, using each monitoring station as a center, we draw a circle with 10 km radius. We overlay all the geocoded firms (from the ASIF dataset and the ESR dataset) on the map, and only keep those falling in a 10-km circle, and lying in a township with a river passing through it. This procedure allows us to identify all the firms that are relevant for our research design.

After that, we calculate each firm's distance to its nearest monitoring station. In some cases (mostly in the eastern coastal areas), the distribution of monitoring stations can be very dense and multiple tributaries or branch rivers merge into the trunk streams. As a result, some 10-km circles overlap with each other, making it difficult to identify upstream and downstream

relationships (i.e., an adjacent upstream firm for one monitoring station can be in the adjacent downstream of another monitoring station). We therefore exclude these water monitoring stations from our dataset. In some less-developed regions (mainly in the Western areas), the distribution of large industrial firms is so sparse that the 10-km circles around monitoring stations contain no firms from the ASIF or ESR datasets. We also drop these monitoring stations from our sample. After these exclusions, we are able to use 161 water quality monitoring stations. The distribution of our sampled monitoring stations is represented in Figure 2.

For each firm kept in the sample, we project its location onto the nearest river basin, and extract the elevation of that projected point. Then, we compare this elevation to the elevation of the adjacent monitoring station, so that we could decide for each firm whether it is in the upstream or downstream of its adjacent monitoring station. In the end, our sample includes 19,150 unique ASIF firms and 9,888 ESR firms from 544 townships, located around 161 water quality monitoring stations.

We attempted to match the firms across the ASIF and ESR samples. However, because these two datasets use different sampling criteria and are managed by different government agencies (using different coding systems), we were only able to match 10% of the ASIF firms with the ESR firms. The matched sample is too small for us to draw any credible statistical inferences. In addition, as mentioned, we have concerns about ESR data quality. Therefore, in subsequent sections, we analyze these two datasets separately and mainly emphasize the results from the ASIF data.

C. Balance Checks

The underlying assumption for our RD design is that, in the absence of environmental regulation, upstream and downstream firms should be *ex ante* identical. Since environmental regulation may affect many production decisions, it is in general difficult to test this assumption using firm-level data, which primarily contains time-varying variables. The only two (arguably) time-invariant covariates in the ASIF data, i.e. “firm establishment date” and “firm ownership type”, are indeed smoothly distributed around monitoring stations, as shown in Table 1 Panel A.

As discussed in Section II Part A, surface water regulation was not strictly enforced until President Hu Jintao came into power in 2003, thus water monitoring stations should not affect

upstream firms in the early years of our data.[10] We thus compare upstream firms with downstream firms using pre-2003 data and test the differences in key variables such as value added, profit, employment, capital, intermediate input and tax. As can be seen from Panel B of Table 1, the vast majority of these covariates are balanced between the two groups. These results are consistent with our identifying assumption.

In addition to the balance tests using firm-level data, we also conduct balance tests using township level data, which include a rich set of variables that are important for firm production, to provide additional evidence that upstream and downstream regions are comparable in aspects other than water regulation. The results are summarized in Appendix Table S2. We see that basic township characteristics are balanced, including township area, arable area, distance to county center, whether the township is an old-region town, whether it is an ethnic minority town, the number of residents, and the number of administrative villages.[11] In Panel B, we test whether basic infrastructure measures are similar between upstream and downstream townships before 2003. Again, the length of roads, number of villages with road access, number of villages with electricity access, and number of villages with tap water access are similar between upstream and downstream areas before water regulation became a binding constraint. Finally, production requires labor. We examine whether human capital differs significantly between upstream and downstream townships prior to 2003. In the township data, we have two relevant variables: the number of primary schools and the number of students enrolled in primary schools. Again, we find no evidence that upstream townships differ from downstream townships in this regard.

The results in Table 1 and Appendix Table S2 are encouraging, as they indicate that upstream and downstream firms are well-balanced for both time-invariant characteristics and pre-2003 covariates, and these firms are located in townships that are highly comparable. While it is, of course, impossible to completely rule out the presence of unobserved factors discontinuously change across the monitoring stations, these balance checks lend additional credibility to our research design.

[10] The dynamic analysis of the RD results will be discussed in Section V.

[11] An old region refers to a Communist Party's revolutionary base region. An administrative village is organized by one village committee and may include several natural villages.

IV. Results

A. Effects of Water Quality Monitoring on TFP

We begin the analysis by visualizing our main findings. Figure 3 plots log TFP (adsorbing station FE and industry FE) against distance to the corresponding monitoring station. Each dot represents the average log TFP for firms within a bin of distance; their 95% confidence intervals are also presented. A fitted function is then overlaid on the graph to illustrate the discontinuity at the monitoring stations.

We divide the firms in ASIF into polluting industries and non-polluting industries based on the definition of polluting industries used by the MEP (in terms of COD emissions).[12] In Panel A, we show the RD plot for residual log TFP in the polluting industries. We see a sharp change in TFP at precisely the locations where the water quality controls take effect. The TFP of upstream firms is significantly lower than that of downstream firms in polluting industries. In addition, it can be seen from Panel A that the treatment effects only apply to firms in the immediate upstream (<5km), rather than firms in the further upstream. This corresponds to the fact that surface water pollutions tend to dissipate over space, so emissions from the further upstream have limited impact on water quality readings. In contrast, in Panel B, we do not observe any comparable discontinuity in TFP in non-polluting industries.

Table 2 quantifies the graphical findings in Figure 3. Panel A presents the RD estimates without any controls, for both polluting and non-polluting industries. We see that polluting firms located immediately downstream from monitoring stations have higher TFP, but there is no TFP difference for the non-polluting firms. The estimates are not statistically significant because of large standard errors.

Our sample covers 161 water quality monitoring stations in 34 manufacturing industries. A simple RD regression, as reported in Panel A, would compare upstream and downstream firms from different clusters (monitoring stations) and industries, creating noise in the statistical inference. To address this issue, we control for station fixed effects in Panel B, and control for both station and industry fixed effects in Panel C. By doing so, we effectively compare the TFP differences station by station and industry by industry, and then average the differences across stations and industries. Comparing the RD estimates in Panels B and C to Panel A, we see that the magnitudes of the estimated impacts are largely unchanged throughout different

[12] Details of the polluting and the non-polluting industries are summarized in Appendix Table S1.

specifications. This is important because it suggests that station- and industry-specific characteristics, while being important determinants of firms' TFP, are uncorrelated with the treatment status. As we control for the fixed effects, the RD coefficients get more precisely estimated, and thus become more statistically significant.

The estimates for the non-polluting industries are close to zero and none of them are statistically significant. This suggests that our results are indeed driven by environmental regulation, rather than other confounding differences between upstream and downstream areas. For both sets of results, the RD estimates are robust to different choices of kernel functions.

In our preferred specifications, which account for both station and industry fixed effects, the estimated gap in log TFP between upstream and downstream firms ranges from 0.31 to 0.35 for the polluting industries. These estimates imply that the water quality monitoring has reduced upstream firms' TFP levels by 26.7% ($e^{-0.31}$-1) to 29.5% ($e^{-0.35}$-1). While a 27% change in TFP is certainly substantial, the magnitude is better interpreted in its specific context: a period during which China experienced unparalleled overall industrial TFP growth. According to Brandt et al. (2012), who use the same data and the same method for TFP estimation, the average TFP growth among the ASIF firms was 14% in 2005. Having that benchmark in mind, our RD estimates indicate that more stringent environmental regulation in the immediate upstream of monitoring stations effectively slowed down firm productivity growth by two years. As will be discussed in the mechanisms section, this is likely due to the fact that upstream polluting firms need to invest capital in abatement facilities, instead of in production.

B. TFP Effects by Firm Ownership, Size, and Firm Age

In Table 3, we explore whether the effect of water quality monitoring on TFP varies by ownership, firm size and firm age. In light of the findings reported in Table 2, we focus on residual TFP with station and industry fixed effects absorbed.

In Panel A, we estimate the RD by firm ownership type and find that the baseline TFP loss is driven mainly by private Chinese firms. Water quality monitoring has no significant impact on the TFP of state-owned enterprises (SOEs) and foreign firms. This result may reflect the fact that environmental regulations are not binding for SOEs or foreign firms as a practical matter; they generally have greater bargaining power over local governments and thus face less stringent enforcement. Another possible explanation is that SOEs and foreign firms generally have superior *ex-ante* environmental performance compared to private Chinese firms and therefore are not affected by tighter regulations. However, given the relatively small number

of observations for SOEs and foreign firms in our sample, these sub-sample null results should be interpreted with caution.

In Panel B, we investigate the impact heterogeneity by firm size. In China, the government adopts a policy strategy called "Grasping the Large and Letting Go of the Small" ("*Zhua Da Fang Xiao*"). "Grasping the large" means that policymakers mainly target large enterprises, while "letting go of the small" means that the government exerts less control over smaller enterprises. The phenomenon has been widely documented in the context of economic reforms and policy implementation (see, for example, Hsieh and Song, 2015). In environmental regulation, many policies are also designed in such a way that larger firms need to meet larger abatement targets.[13] We investigate if this phenomenon is true in our setting. We define small firms as having less than 50 employees and the rest are categorized as large firms. The results in Panel B show that the TFP impacts are statistically significant only for larger firms. In other words, the "Grasping the Large and Letting Go of the Small" strategy seems to be applied to the context of water quality regulations, too.

In Panel C, we compare the TFP loss by firm age. We are interested in whether old firms and young firms respond differently to water quality monitoring. We define new firms as firms born in or after 2003, when China's environmental regulations became stringent. We then estimate the discontinuities separately for old and young firms using post-2003 data. We find that the TFP loss caused by water quality monitoring is statistically significant for both old and young firms. This finding is inconsistent with the "grandfathering" phenomenon, in which new environmental policies are often designed or implemented in such a way that older firms can be exempted from tighter regulations, because the cost of retrofitting existing facilities is higher than that of building new sources with cleaner technology. In the context of China's water quality monitoring, both young and old upstream firms have been under tighter regulation than downstream firms since 2003. In addition, the magnitude of treatment effects for young firms established after the regulation became stringent in 2003 is comparable to that of old firms, suggesting that the selective locational choices of young firms are not driving the baseline findings.

[13] See, for example, "The Top 10,000 Energy-Consuming Enterprise Program," which requires only large firms to abate carbon emissions: http://www.ndrc.gov.cn/zcfb/zcfbtz/201112/t20111229_453569.html

C. Results by Year

The stringency of water quality regulations has changed substantially during our sample period. Specifically, in 2003, President Hu Jintao proposed the "Scientific Outlook of Development" initiative to address the pressing environmental challenges in China. In the same year, the MEP upgraded the surface water quality monitoring system. In addition, starting in 2006, COD abatement became a key indicator in evaluating local environmental performance, which were explicitly linked to political promotions.

We thus hypothesize that the TFP effect of water monitoring should be intensified after 2003 and 2006, respectively. In Figure 4, we provide RD estimates separately for each year. We find that the TFP differences between upstream and downstream firms exactly match the policy changes we have discussed. Specifically, the estimate is close to zero from 2000 to 2002, and becomes larger in 2003, the year President Hu took office. The effect becomes statistically significant starting in 2006, the first year of the 11th Five Year Plan. The corresponding results are summarized in Table 4.

The finding that the monitoring effect was close to zero and statistically insignificant prior to 2003 is consistent with the balance tests presented in Table 1, and further justifies our identifying assumption: in the absence of tighter water quality regulations, upstream and downstream firms around the same water quality monitoring station have similar levels of productivity.

The dynamic pattern of the RD coefficients is also re-assuring in terms of ruling out alternative explanations: to the extent that one thinks the baseline results are driven by confounding factors, such factors have to be specific not only to upstream vs. downstream firms, polluting vs. non-polluting industries, but also specific to the timing of two independent political events happened in China in 2003 and 2006, respectively.

D. Endogenous Location of Monitoring Stations

Our qualitative discussions on the rules of setting up monitoring stations, the balance tests of firm-level and township level variables, the finding that the discontinuity is only evident for polluting industries, and the result that the discontinuity only emerges after 2003, all suggest that the identifying assumption in our RD design is likely to hold.

Nevertheless, one may still be concerned about the endogenous location of monitoring stations. For instance, a politically connected polluting firm has strong incentives to lobby the

local government, so that the monitoring station would be established upstream of that firm. If these connected firms also receive other forms of benefits from the government that could affect their productivity, such as subsidies or loans, our RD estimates would be biased.

In this section, we use an instrumental variable (IV) approach to directly address this concern. We exploit the fact that, when monitoring stations were set up, local governments typically attempted to locate them closer to existing hydrological stations, so that data, equipment and technicians could be shared in order to achieve economies of scale in water monitoring.

A hydrological station collects hydrological data such as water levels, flow velocity, flow direction, waves, sediment concentration, water temperature, and ice conditions, as well as data on meteorological conditions such as precipitation, evaporation, air temperature, humidity, air pressure and wind. Because hydrological stations were set up between the 1950s and 1970s (a period when China barely had any industrial pollutions at all), and because their locations were chosen based purely on hydrological considerations, these locations should be orthogonal to the future socio-economic conditions of their neighborhoods. All the hydrological stations were built and supervised by the Ministry of Water Resources (MWR), instead of the Ministry of Environmental Protection (MEP), and play no roles in collecting any measures of water pollution.

Therefore, one would expect that, except for inducing the establishment of monitoring stations, the existence of a hydrological station alone should have minimal impact on the production and emission behaviors of adjacent firms. Utilizing this "exclusion restriction," we adopt "whether a firm is in the near upstream area of a hydrological station" as an instrumental variable (IV) for "whether a firm is in the near upstream area of a monitoring station" and estimate a 2SLS to quantify the impacts of water quality monitoring on TFP.

Empirically, we estimate the following first-stage regression:

(3) $$UpMoni_{ijk} = \alpha \cdot UpHydro_{ijk} + \lambda_j + \sigma_k + \epsilon_{ijk}$$

where $UpMoni_{ij}$ is a dummy variable indicating whether firm i in industry j is in the near upstream area (10 km) of monitoring station k; $UpHydro_{ijk}$ is a dummy variable indicating whether firm i in industry j is in the near upstream area (10 km) of a hydro-station k; λ_j and σ_k represent industry and monitoring site fixed effects; and ϵ_{ijk} is the error term. We then estimate the second stage regression:

(4) $$TFP_{ijk} = \alpha \cdot \widehat{UpMoni}_{ijk} + \lambda_j + \sigma_k + \epsilon_{ijk}$$

where TFP_{ijk} is the TFP of firm i in industry j near the neighborhood of monitoring station k; $\widehat{UpMon}_{ijk}$ is the predicted value from the first stage regression; λ_j and σ_k are industry and monitoring site fixed effects; and ϵ_{ijk} is the error term.

The regression results are presented in Table 5. We estimate the effects separately for firms in the polluting industries and for firms in the non-polluting industries. First, we find that the locations of hydrological stations can strongly predict the locations of water quality monitoring stations (Columns 1 and 3): if a firm is near the upstream of a hydrological station, it is also more likely to near the upstream of a monitoring station. The 2SLS estimates show that being in the near upstream of a water monitoring station decreases the TFP of a polluting firm by 0.35 logarithmic units (Column 2), but it does not affect the productivity of non-polluting firms (Column 4).

Note that the regression results in Table 5 are not readily comparable to those in Table 2, as these two approaches use very different sources of variation in the data and estimate different treatment effects with different identifying assumptions. The RD design estimates the average treatment effect at the cutoff, whereas the IV estimates the local average treatment effect for firms near a hydrological station. Nevertheless, the closeness of the magnitudes of the estimates between the two approaches (0.31 to 0.35 versus 0.35), and the consistent findings in both sets of results, provide additional support to the causal relationship between water quality monitoring and firm TFP.

E. Sorting of Polluting Firms

Environmental policies can affect firms' production plans and their location choices. In particular, the pollution haven hypothesis (PHH) posits that polluting capital would flow from places with more stringent environmental regulations to places with less stringent regulations. This issue is important because it will affect the interpretation of the RD estimates: if certain types of firms are systematically more likely to sort away from the upstream, then our baseline findings could be partly driven by sample selection, instead of intensive margin treatment effect.

In this section, we conduct a series of analysis to formally investigate whether there indeed exists significant sorting, and whether it affects the interpretation of our baseline findings. First, we follow Cattaneo et al. (2017a, 2007b) and conduct a data-driven manipulation test. As shown in Panel A of Figure 5, using the baseline sample (collapsed cross-section for 2000-2007), we find that there is no discontinuity in the distribution of polluting firms around the

location of monitoring stations. The formal statistical tests are summarized in Appendix Table S3. This test suggests that there does not significant sorting in our baseline sample, so sorting could not explain the baseline RD results.

The intuition for the lack of sorting is the following: the firms enter into our ASIF dataset are generally large ones, for whom it is very costly and difficult to relocate. Even if they want to relocate, they need to buy a new piece of land from another local government, and build a new factory, before they can move the labor and capital to the new site. For most large manufacturing firms, this whole process of relocation could take years to happen. Recall that in Figure 4, we showed that environmental regulation only became a binding constraint after 2003, and the regulation effect only became significant after 2005, while the sample ends in 2007. So even if polluting firms decided to relocate immediately after regulation started to hurt them, the time window in our sample period is simply too short for that to happen, which explains why we could not detect any significant sorting in the data-driven manipulation test.

Fortunately, while we no longer have TFP information after 2007, the ASIF dataset still allows us to identify the location of polluting firms up until 2013, which is a decade after water quality monitoring started to matter for polluting firms. To verify the intuition on the lack of sorting in the short run, we take the ASIF data in 2013, and re-run the same set of density tests.[14] As shown in Panel B of Figure 5, in 2013, the density of the polluting firms indeed becomes slightly discontinuous at the cut-off: fewer polluting firms are located in the immediate upstream of monitoring stations. This finding suggests that more polluting firms leaving (or less firms emerging in) the upstream regions to avoid tighter regulation in the long run. The corresponding density tests results are reported in Panel B of Appendix Table S3.

Taken together, results from these two data-driven manipulation tests suggest that sorting might exist, but only in the long run, not in the short run. Therefore, sorting could not explain our baseline RD results in the short run.

Since the data-driven manipulation tests are essentially comparing the density of polluting firms around the cut-off, a potential concern is that while the overall density remains unchanged, different types of polluting firms might differentially enter\exit the upstream and downstream areas, creating a confounding difference in the composition of firms. To address this concern, we keep only a balanced panel of polluting firms during 2000-2007, and re-run the baseline RD for this sub-sample. Since these firms were already in the sample before

[14] The "value-added" variable is only available in the ASIF dataset until 2007; therefore, data in the later years could not be used for TFP analysis.

regulation became relevant in 2003, and remained in the sample until 2007, the RD coefficient estimated using this subsample should solely reflect the "intensive margin treatment effect," rather than any form of "sample selection effect." As shown in Appendix Table S4, while the point estimates are noisier due to reduced sample size, the balanced panel sample still show consistent treatment effects with similar magnitudes to the baseline. Again, this exercise confirms that our main findings cannot be explained by the sorting of polluting firms.

F. Spillover Effects

Spillover effects, i.e. water quality monitoring somehow also affecting downstream firms, would not be a concern in a market with perfect competition, where there are many firms and the output market is unaffected by local environmental regulation. However, in a market with imperfect competition or more complicated structures, spillovers can exist. In our empirical setting, both positive and negative spillovers can emerge, depending on how upstream firms and downstream firms interact.

If industries are highly concentrated and their major producers are geographically clustered near the water quality monitoring stations, then imposing tighter environmental regulations on upstream firms would cause positive spillovers to downstream firms. The reasons are twofold. First, because upstream and downstream firms are the main producers and competitors in the market, increased production costs for the upstream firms will raise the market price of their products. Competing downstream firms will thus benefit because of this change in market conditions, not just because of the environmental enforcement affecting their upstream counterparts. Second, tighter environmental regulations also may cause inputs, both labor and capital, to move toward the downstream firms. If more productive factors flow to downstream firms, their TFP will be higher for this reason as well.

Negative spillovers will emerge when clustered firms are collaborating instead of competing. This is particularly true if clustered firms are vertically integrated along the supply chain. If (geographically) upstream firms produce inputs for downstream firms, or *vice versa*, environmental regulations that increase upstream firms' marginal costs of production will also make downstream firms less competitive.

In the presence of spillovers, when we extrapolate these estimates to the whole country, the existence of a positive spillover effect will exaggerate the economic costs of regulation, while a negative effect will attenuate the estimated costs.

To assess whether or to what extent our findings are confounded by the potential spillover effects, we conduct a formal test using placebo downstream firms. Specifically, we first replace the actual downstream firms by their best matches from the sample of firms that are not in the neighbourhood of any monitoring stations, based on pre-2003 data. We then re-estimate the regression discontinuities for the matched firms using post-2003 data. These matched firms serve as placebo firms which are not affected by the potential spillovers between actual upstream and downstream firms. The intuition is that, if the spillover effects are insubstantial (downstream firms are not affected by monitoring stations), the placebo firms should have TFP similar to the actual downstream firms. Using placebo downstream firms should lead to results that are quantitatively similar to the baseline estimates.

In practice, we take the pre-2003 collapsed cross-sectional data and use a nearest neighbor matching strategy that finds the best matched firm from the pool of firms that are located outside the 10 km radius of the water monitoring stations for each downstream firm. These placebo downstream firms resemble the actual downstream firms in terms of TFP, industry type, and industrial output value before 2003. We then replace the actual downstream firms by the placebo firms in the post-2003 sample and estimate the regression discontinuities.

The results are reported in Table 6. Upstream firms have significantly lower TFP than placebo downstream firms, suggesting that the baseline findings are not driven by a positive spillover effect on the downstream firms. We focus our discussion on the RD estimates after station and industry fixed effects are absorbed, in Panel B. Compared with placebo downstream firms, upstream firms' log TFP is 0.48 to 0.61 units higher. These estimates are slightly larger than those in Table 3, but the differences are statistically indistinguishable. That implies that, if there may exist some spillover effect, this effect should be slightly negative. Consequently, the estimates in Tables 2 will only understate the economic costs of water pollution regulation.

G. Robustness to Different Specifications

We check the robustness of our findings in Table 7. In Panel A, we re-estimate our models using a method proposed by Calonico, Cattaneo, and Titiunik (2014) in which local linear regression estimates can be "bias-corrected" for biases resulting from choice of bandwidth. They also suggest an alternative method for calculating standard errors that is more conservative than the conventional procedure. Using these alternative methods, we generate results that are qualitatively similar to the results featured in our main analysis.

In Panel B, we use alternative bandwidth selectors. The bandwidth chosen in our main analysis is based on one common MSE (Mean Square Error)-optimal bandwidth selector for both sides across the cutoff. We supplement this analysis with five other bandwidth selectors: (1) MSE-two: two different MSE-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect estimator; (2) MSE-sum: one common MSE-optimal bandwidth selector for the sum of regression estimates (as opposed to the difference thereof); (3) CER (coverage error rate)-optimal: one common CER-optimal bandwidth selector for the RD treatment effect estimator; (4) CER-two: two different CER-optimal bandwidth selectors (below and above the cutoff) for the RD treatment effect estimator, and (5) CER-sum: one common CER-optimal bandwidth selector for the sum of regression estimates (as opposed to the difference thereof).[15] The results remain the same regardless of the bandwidth selector used.

In Panel C, we conduct a placebo test using "fake" monitoring stations. We move the original stations upstream or downstream by 5 km or 10 km and re-estimate the RD models. We find that the discontinuity in TFP is only evident at actual monitoring stations and not at the fake stations.

In the Appendix, we present more robustness checks. In Appendix Table S7, we report the RD estimates using the parametric approach, Equation (2). We find quantitatively similar results: water quality monitoring decreases polluting firms' TFP but has no impact on non-polluting firms. However, the estimates from the parametric approach are more sensitive to the choice of the polynomial function form and inclusion of different samples. In Appendix Table S8, we use an alternative TFP measure suggested by Ackerberg et al. (2015) as the outcome variable, and again the results remain unchanged.

V. Channels: What Happened to the Upstream Firms?

A. Regulation and Firm Production

How do firms respond to tighter environmental regulations? In this section, we examine the channels through which environmental regulation affects firms' TFP. To rationalize the baseline findings and guide the following discussions, we present a theoretical framework to illustrate how environmental regulation can negatively affect TFP in Appendix B. In this

[15] Please refer to Calonico, Cattaneo, Farrell (forthcoming) for technical details.

model, firms need to use extra labor and capital to clean up emissions and the government can enforce tighter environmental regulation by increasing the emission tax. Facing a higher emission tax, firms need to hire more labor and capital for emission abatement, but these extra inputs do not directly contribute to output production. As a result, tighter environmental regulation will lead to a reduction in firms' TFP.

In Table 8, we estimate the impacts of water quality monitoring on several key variables and test whether these findings are consistent with our theoretical predictions. Panel A of Table 8 summarizes the results for output-related measures: value added and profit. Although the effects of water monitoring are statistically insignificant, we see a tendency that downstream firms earn more profit despite not producing more product. In Panel B of Table 8, we focus on labor, capital, and intermediate input. Labor is measured by the number of employees. We find that upstream firms hire more employees (16%) but the effect is statically insignificant. Upstream firms also tend to use slightly more intermediate input for production. Capital stock is measured by two different methods for robustness, following Yang (2015) and Brandt at al. (2012) respectively. We see that upstream firms have significantly larger capital stock compared to their downstream counterparts and the effects are statistically significant using different estimators. These results are consistent with the ubiquitous anecdotal evidence that polluting firms have to buy expensive equipment or abatement facilities to cope with tighter environmental regulation standards.

The interpretation of "increased abatement capital expenses in the upstream" could be corroborated using the firm-level emission dataset, which has a variable documenting the amount of abatement equipment owned by the firm. As shown in Appendix Table S5, upstream polluting firms indeed own significantly more abatement equipment than their downstream counterparts, consistent with the capital stock results.

In Panel C of Table 8, we further present results for naïve (reduced-form) productivity measures, calculated by dividing firms' value-added by labor or capital stock. Similarly, both labor and capital productivity are higher in downstream firms, with the impact on capital productivity being statistically significant. These results show that our baseline findings are robust to the use of simpler and more transparent productivity measures and reflect a real loss in firm productivity, rather than being mechanical to specific procedures of TFP construction.

In Panel D of Table 8, we test the Porter Hypothesis. The outcome of interest is firms' investments in research and development. The results show that tighter environmental

regulation increases firms' investment in research and development, which is consistent with the Porter Hypothesis, but the result is statistically insignificant due to large standard errors.

The results in Table 8 suggest that the impacts of environmental regulation on TFP are manifested through multiple channels. The overall pattern confirms the predictions of our model: facing tighter environmental regulations, firms need to install expensive facilities (and potentially also hire more labor) to abate emission, leading to lower productivity.

B. Regulation and Emissions

The model in Appendix B also predicts that tighter environmental regulations will decrease both emission levels and emission intensity (emission per unit of output). In other words, upstream polluting firms are expected not only to reduce total emissions, but also to adopt cleaner technologies. This is consistent with the previous finding that upstream firms have larger capital stock. In this section, we formally examine the impacts of water quality monitoring on firm's emission and emission intensity.

We apply the same set of RD estimators to firm-level emission data (which is equivalent to polluting-source level data) from the ESR database. We examine four water pollution outcomes: (1) total amount of COD emitted, (2) COD emission intensity (total COD/total output value), (3) total amount of wastewater discharged, and (4) wastewater discharge intensity (total wastewater/total output value).

Table 10 reports the local linear RD estimates for the four outcomes. In Panel A, we can see that both COD emissions and COD emission intensity are higher for downstream firms, and most results are statistically significant at the 5% or 10% level. COD emissions by polluters immediately upstream from monitoring stations are 0.75-0.99 logarithmic units lower than those from firms immediately downstream. This implies that water quality monitoring reduces COD emission levels in upstream firms by 52.8%–62.8% ($e^{-0.75}$-1 to $e^{-0.99}$-1). For COD emission intensity, water quality monitoring reduces the COD emission intensity in upstream firms by 38.7%–49.3% ($e^{-0.49}$-1 to $e^{-0.68}$-1). In Panel B, we examine wastewater discharge. Downstream firms tend to discharge more wastewater but the results are statistically insignificant due to large standard errors. The results for wastewater discharge intensity, however, are statistically significant at the 5% or 10% level.

Combining both sets of results, we conclude that upstream firms emit less COD and wastewater overall and also produce fewer COD emissions or less wastewater per value of output (by adopting cleaner technologies), confirming the theoretical predictions.

Recall that the ESR database samples the most polluting firms in each county. Given that we focus on a small region around each monitoring station, many of the upstream and downstream firms are located within the same county. This causes a potential selective attrition problem because upstream firms facing tighter regulations tend to emit less and are thus less likely to be sampled in the ESR database compared to downstream firms. If such selection exists, our results in Table 10 will be underestimated, because the upstream firms that reduced the most pollution are no longer included in the sample. Thus, when we evaluate the environmental benefits of water monitoring, the estimates in Table 10 should be regarded as lower bounds.

C. Political Economy of Water Quality Monitoring

Our empirical analyses have shown that upstream firms' productivity is negatively affected by water quality monitoring. Our explanation is that, because water quality readings are politically important, local officials have incentives to enforce tighter regulations on upstream firms than on downstream firms. In this section, we explicitly explore such political-economic incentives behind water quality monitoring.

First, we document how upstream and downstream firms are treated differently by the government. In the ESR dataset, we find that upstream firms have substantially lower COD emission and wastewater discharge. In the ASIF dataset, we also have information on the waste discharge fees paid by each firm in 2004. If the government imposes a "fair" rule of punishing upstream and downstream firms for emissions, we should expect downstream firms to pay more than upstream firms, due to their higher emission levels. However, as shown in Panel A of Table 10, we find that upstream firms actually pay significantly more waste discharge fees to the government, despite that they emit much less. In other words, local governments are able to charge firms at differentiated emission fee rates, even though these firms are located close to each other and are within the same administrative jurisdiction. This double-standard phenomenon is not unique in our study and has been documented in some other settings as well. For example, Liu (2017) investigates China's tax reform and finds that local governments of China are able to collect additional informal taxes from certain firms for fiscal revenue. Fan et al. (2017) study China's value-added tax (VAT) system and show that firms located farther away from local tax agencies experience the largest increase in tax burden after VAT enforcement costs are brought down by a new information technology. Local governments in China have substantial discretion in the management of taxation and various fees.

Second, we examine the political incentives of local officials. As documented in the Chinese meritocracy literature, China has an implicit rule that a prefecture-level governor cannot be promoted to a higher level if his/her age reaches 57 (for example, Wang, 2016; Xi et al., 2017). This creates a discontinuous drop in political incentives at the age of 56. To test whether the TFP effects of water quality monitoring can be explained by political incentives, we digitize the résumés of every prefectural party secretary (the highest-ranked political leader in a prefectural city) between 2000 and 2007. We define a leader as "having strong political incentives" if he/she is younger than 56 in a given year, and "having weak political incentives" otherwise. We then assign a monitoring station either to an "incentivized" or "un-incentivized" party secretary in a given year and analyze two subsamples based on whether the monitoring station is under the governance of an "incentivized" leader in a particular year. The RD results are summarized in Panel B of Table 9.

We find that, when the prefectural city leader has strong political incentives, water quality monitoring has a statistically significant impact on upstream firms' TFP. The estimated impacts range from 0.57 to 0.66 using different kernel functions and are nearly twice as large as the baseline results in Table 2. In sharp contrast, when the prefecture city leader has weak promotion incentives, the TFP gap remains precisely at zero in all specifications. These results imply that the TFP discontinuity across the monitoring stations is driven by the political incentives of local officials.[16]

Third, despite the fact that state-controlled monitoring stations are established and run by the central government, it is still possible that local officials can exert their administrative powers to influence the water quality monitoring. Our concern is that, if local governments can manipulate water quality readings, they may be less incentivized to regulate upstream firms' emissions. These is evidence that air pollution data has been manipulated at the margin in some Chinese cities because air quality is important for political evaluation (Ghanem and Zhang, 2014).

To test this hypothesis, we estimate the RD separately for two types of monitoring stations: automatic stations and manual stations. Automatic stations conduct all water quality tests automatically and report the data directly to the central government, while manual stations

[16] In an alternative specification, we use the panel dataset, and exploit the age change from 56 to 57, holding the leader fixed. The main results still go through with this much more restrictive specification, as shown in Appendix Table S6.

require technicians to conduct the tests manually.[17] Because it is difficult for local governments to manipulate data from the automatic stations, we expect a larger TFP gap around automatic stations.

Panel C of Table 9 reports the findings: while we see an upstream-downstream TFP gap for both types of stations, this effect is much larger for automatic stations (almost three times larger). As the sample size shrinks substantially, most RD estimates are statistically significant only at the 10 percent level.

D. Enforcement in Practice

The results in the previous subsections are instrumental for the understanding of China's environmental regulations. In particular, we show that similar firms adjacent to each other can live in dramatically different regulatory environments, leading to different firm behaviors and diverging productivity.

To confirm these findings with qualitative materials, we conducted multiple field trips and interviewed dozens of firm managers and technicians working in the monitoring stations. Our discussions reveal that, not only do upstream firms bear larger burdens of emission fees, as shown in our empirical analyses, they also face a variety of command and control regulations that cannot be easily quantified in our data. For example, in one city, we learned that firms' production could be abruptly restricted or even suspended by the local government in order to improve water quality readings. In China's most polluted river basin, Huai River basin, environmental inspectors were placed in the polluting firms from time to time to ensure their compliance with environmental standards. These inspectors visited upstream firms more frequently because they knew these firms had large impacts on water quality readings. In some extreme cases, if certain firms do not comply with the regulation, electricity and natural gas supply could even be cut in order to meet the city's environmental abatement target.

Regulatory documents from local governments tell a similar story. Urgent orders were issued when local governments realized water quality readings might fail to meet higher-level government policy targets. A recent example of such an order, which attracted wide media attention in China, is presented in Appendix C. In this example, Kunshan city in Jiangsu Province required 270 manufacturing firms to suspend their production in order to improve

[17] Most stations were manual in the 1990s and early 2000s, but these were gradually replaced by automatic stations, in order to improve the accuracy of water quality reporting. Weekly water quality reports from the automatic stations are posted by the MEP at http://datacenter.mep.gov.cn/index and real-time water quality readings can be accessed at http://online.watertest.com.cn/help.aspx.

water quality. Sluice gates along the rivers were closed and the pumping facilities were shut down so that no wastewater could be discharged into the rivers, even after abatement treatment. Special investigators were sent to the plants to enforce the production suspension policy. While some of these command and control policies cannot be quantitatively analyzed in our empirical analysis, they do help explain why China's water quality monitoring can be so effective in reducing water pollution and has such a significant impact on upstream firms' productivity.

VI. Economic Significance

A. Economic Costs under Various Scenarios

Our baseline model estimates that water quality monitoring has caused an average loss in TFP of 0.31 logarithmic units for polluting firms (as shown in Panel B of Table 2), equivalent to a 26.7% drop. To translate this TFP loss into monetary value, one may ask what would happen if all of China enforced regulatory standards as stringent as those faced upstream. The total industrial output value (total revenue) from the polluting firms was about 11 trillion Chinese yuan (1,380 billion USD) in 2006.[18] If all these firms were subject to water quality monitoring regulations as stringent as those faced by the upstream firms in our empirical setting, the total annual loss in output value would exceed 4.0 trillion Chinese yuan (502 billion US dollars) based on 2006 industrial output value.[19]

However, the regulations faced by upstream firms may be too stringent to apply to all the other firms in the country. A more informative counterfactual would be to determine the TFP loss and economic costs associated with a given amount of emission abatement. Recall that all the firms in the ESR database together contribute 85% of China's total emissions, and all of them are local large emitters regardless of industry or revenues.

Because we are unable to match the ESR firms with ASIF firms, we cannot directly link the TFP estimates with COD estimates without imposing additional assumptions. The TFP and COD effects of water monitoring we estimated in previous tables essentially are the following:

$$\text{(5)} \qquad \mathrm{TFP_{ATE}}|\text{Revenue} \geq 5 \text{ million} = \mathrm{E}(\mathrm{TFP_1} - \mathrm{TFP_0}|\text{Revenue} \geq 5 \text{ million})$$

[18] We use the 2006 exchange rate of 1:7.97.

[19] We compute the difference between the counterfactual output of 14,973.7 billion Chinese yuan (calculated by 10975.7/(1-26.7%)) and the observed output of 10,975.7 billion Chinese yuan in the polluting industries in 2006. The calculations for other parts follow the same method.

(6) $\mathrm{COD_{ATE}|COD} \geq \mathrm{x} = \mathrm{E(COD_1 - COD_0|COD} \geq \mathrm{x})$

where $\mathrm{TFP_{ATE}|Revenue} \geq 5\ \mathrm{million}$ is the average treatment effect of water quality monitoring on TFP for firms with annual revenues over 5 million yuan, and $\mathrm{COD_{ATE}|COD} \geq \mathrm{x}$ is the average treatment effect of monitoring on emitters that produce COD pollution more than a given threshold x. $\mathrm{TFP_1}$ is the TFP for downstream firms, and $\mathrm{TFP_0}$ is the TFP for upstream firms.

The average treatment effects on TFP and COD over the entire distribution are:

(7) $\mathrm{TFP_{ATE}} = \mathrm{Prob(Revenue} \geq 5\ \mathrm{million)} \cdot \mathrm{TFP_{ATE}|Revenue} \geq 5\ \mathrm{million} + \mathrm{Prob(Revenue} < 5\ \mathrm{million)} \cdot \mathrm{TFP_{ATE}|Revenue} < 5\ \mathrm{million}$

(8) $\mathrm{COD_{ATE}} = \mathrm{Prob(COD} \geq \mathrm{x)} \cdot \mathrm{COD_{ATE}|COD} \geq \mathrm{x} + \mathrm{Prob(COD} < \mathrm{x)} \cdot \mathrm{COD_{ATE}|COD} < \mathrm{x}$

where the probabilities could be written as the share of firms appearing in each sample:

$$\mathrm{Prob(Revenue} \geq 5\ \mathrm{million)} = \frac{N_{ASIF}}{N},\ \mathrm{Prob(Revenue} < 5\ \mathrm{million)} = 1 - \frac{N_{ASIF}}{N};$$

$$\mathrm{Prob(COD} \geq \mathrm{x)} = \frac{N_{ESR}}{N},\ \mathrm{Prob(COD} < \mathrm{x)} = 1 - \frac{N_{ESR}}{N}.$$

While we cannot directly estimate "$\mathrm{TFP_{ATE}|Revenue} < 5\ \mathrm{million}$" and "$\mathrm{COD_{ATE}|COD} < \mathrm{x}$" in the data, we attempt to back them out by extrapolating the intra-sample heterogeneous treatment effects on TFP and COD.

In Table 11, we estimate the heterogeneous treatment effects of water quality monitoring on TFP with respect to firms' revenues, and the heterogeneous treatment effects of water quality monitoring on COD emission intensity with respect to firms' total COD emissions using the polynomial RD approach.[20] The revenue heterogeneity is estimated by using the polynomial RD approach with an interaction term between the downstream dummy and firms' revenue (log). We use the specification in Column 1 of Table S4 as our preferred parametric specification because it generates the closest RD estimates to the non-parametric RD estimates. To allow for non-linear heterogeneity, we also include quadratic and cubic interactions in the regressions. Based on the regression results, we then predict the estimated impacts at different levels of revenues and summarize the results in Panel A. We find that the TFP effect is substantially larger for larger firms and non-existent for smaller firms. The effects of water

[20] Ideally, we should also apply the non-parametric RD estimates to different sub-groups of firms and estimate the heterogeneity separately for each sub-group. However, doing so significantly reduces the sample size of each group and we would not have strong enough statistical power to make a reliable inference. In Appendix Table S6, we divide the sample into only two groups and find results that are largely consistent with Table 11: the impacts of water quality monitoring are primarily experienced by larger firms or emitters and are negligible for their smaller counterparts.

quality monitoring on TFP for the smallest 20% of firms (among all the firms with an annual revenue of above 5 million Chinese yuan) become negligible. The results are the same if we use quadratic or cubic heterogeneity. In Panel B, we conduct a similar analysis for COD emission intensity and check whether the effect of monitoring varies across different polluting sources. We find the same pattern: larger emitters are strongly affected by water quality monitoring, while the treatment effect becomes essentially zero for the smallest 20% of emitters in the ESR sample. These findings again confirm that the "Grasping the Large and Letting Go of the Small" strategy discussed in section IV B is applied to the context of water quality regulations.

In addition, there is an "exit" variable in the ASIF database documenting whether a firm is excluded from the sample in the following year. A firm that earns less than 5 million Chinese yuan in a particular year, based on the sampling criteria, is dropped from ("exits") the database the next year. This outcome provides additional information on whether water quality monitoring affects firms at the margin. In Appendix Table S7, we find that the probability of exiting the ASIF database is not affected by water quality monitoring. This finding again shows that monitoring does not affect smaller firms at the margin.

Given these findings, if we assume that water quality monitoring does not increase the TFP or emission levels of upstream firms, and that the size of the treatment effect on TFP or emissions is a well-behaved function with respect to revenue or emissions, then we can make the following extrapolations:

(9)
$$\mathrm{TFP_{ATE}}|\mathrm{Revenue} < 5\ \mathrm{million} = 0$$
$$\mathrm{COD_{ATE}}|\mathrm{COD} < \mathrm{x} = 0$$

Intuitively, as the smallest producers and emitters in our ASIF or ESR dataset already have zero treatment effects, the even smaller producers and emitters (those excluded from the ASIF/ESR dataset) also should have zero treatment effects. We can therefore simplify Equations 7 and 8 to the following:

(10)
$$MRS = \frac{TFP_{ATE}}{COD_{ATE}} = \frac{N_{ASIF}}{N_{ESR}} \cdot \frac{\mathrm{TFP_{ATE}}|\mathrm{Revenue} \geq 5\ \mathrm{million}}{\mathrm{COD_{ATE}}|\mathrm{COD} \geq \mathrm{x}}$$

The sample we use for estimation includes 6,581 firms in polluting industries from the ASIF database and 9,888 polluters from the ESR database. Using this equation, we can calculate the economic costs of water pollution abatement.

In Table 12, we compute the economic costs for various scenarios. For easy reference, Panel A reproduces the key results from Tables 2 and 10, and Panel B calculates the economic costs. We focus on the estimates in Column 1 because they produce modest values across all

specifications. We first focus on COD emissions. Water quality monitoring reduces COD emissions by 0.83 logarithmic units and decreases TFP by 0.31 logarithmic units. A 10% change in total COD emissions causes a 2.49% change in TFP levels in the polluting industries.[21] Using alternative specifications produces slightly different results, which are reported in Columns 2–4. Similar interpretations can also be applied to COD emission intensity. In Column 1, an upstream firm's COD emission intensity is about 0.55 logarithmic units lower than that of a downstream firm. This means a 10% change in COD emission intensity causes a drop in TFP by about 3.75%. Other combinations create slight variations, as summarized in Columns 2–4.

During China's 11th Five-Year Plan, total COD emissions were reduced by 12.45% from 2006 to 2010 (with the target being 10%). If we attribute the entire COD reduction from 2006 to 2010 to the polluting industries, then this 12.45% abatement in COD emissions would cause a total output loss worth 352 billion Chinese yuan (44.2 billion USD) in the polluting industries, based on 2006 industrial output values.[22] The annual reduction in COD emissions between 2006 and 2010 was roughly 2.5%, equivalent to an annual loss of 69 billion Chinese yuan (8.7 billion US dollars) in gross industrial output value per year using 2006 Chinese yuan.

In 2015, the gross output value (of firms above designated size) in China exceeded 110 trillion Chinese yuan, and about 35% of output value (38.8 trillion Chinese yuan) is contributed by the polluting industries. The central government aims to reduce COD emissions by another 10% during the 13th Five-Year Plan, from 2016 to 2020. Applying our estimates to the 2015 data, we can infer that the total output loss would be around 990 billion Chinese yuan (159 billion US dollars) under current monitoring and enforcement practices.[23] Other specifications generate slightly different estimates, ranging from 936 to 1,099 billion Chinese yuan (150.5 to 176.7 billion USD).

Note that we make two strong assumptions in calculating the economic costs. First, we implicitly assume the marginal cost of abatement is linear so that the large and small emission

[21] The way we interpret this relationship is analogous to the Wald estimator in the two-stage setting, except that we do not have a readily available tool to combine the two stages from two different samples non-parametrically and we need to adjust for sample size. Water quality monitoring reduced COD emissions by 0.83 logarithmic units and TFP by 0.31 logarithmic units, so a 10% change in COD emissions will lead to a (6,581/9,888)*(0.31/0.83)*10% (= 2.49%) change in TFP.

[22] We estimate that a 10% change in total COD emissions will cause a 2.49% change in TFP, which implies that a 12.5% change in total COD emissions will cause a 3.11% change in TFP. We then compute the difference between the counterfactual output of 11,328 (10,975.7/(1-3.11%)) billion and the observed output of 10975.7 billion in 2006. The calculations for other parts follow the same method.

[23] We use the 2015 exchange rate of 1 USD to 6.22 Chinese yuan.

reductions have proportional impacts on productivity. Second, we assume that the estimates in Table 10 are reliable and potential data manipulation would not significantly change our estimates.

B. Potential Sources of Bias

There are several reasons why the estimates in Table 12 may understate the true economic costs of China's water pollution controls. First, we use a conservative estimate of the effect of monitoring on TFP in our calculations. In fact, as shown in Table 3, the TFP loss due to water quality monitoring has increased from 0.31 to 0.40 since 2003. If we use these larger TFP estimates, the associated economic costs will increase.

Second, although we provide evidence that the smaller firms or emitters in our data are not affected by water quality monitoring, the assumption that water quality monitoring does not affect even smaller (unobserved) firms or emitters at all may still be violated. Shutting down very small polluters can be a feasible policy for some local governments to enforce tighter environmental standards. The TFP loss due to shutdown cannot be captured in our estimation.

Third, the distinction between polluting and non-polluting industries is based on two- to three-digit industrial codes. This distinction does not rule out the possibility that some firms in the non-polluting industries may also emit pollutants and are therefore regulated by local governments. If this is the case, the estimated TFP and economic loss are understated.

Fourth, some regions have a high density of monitoring stations as well as multiple tributaries along the main streams. These monitoring stations are excluded from our sample because we cannot credibly identify the upstream and downstream townships. If there are more monitoring stations in more polluted regions, some of the most polluted regions and firms are excluded from our sample; if environmental regulations are even more aggressively enforced in more polluted regions, the TFP loss in these regions could be even larger.

Finally, we only compute the direct economic costs caused by TFP loss. Previous research has shown that tighter environmental regulation can also cause unemployment, firm relocation, and worker migration, and can change the flow of investment. These indirect costs are non-trivial and should be considered when calculating the overall economic costs of environmental regulations.

VII. Conclusion

As the income levels of the Chinese people rise, the country starts to face a stark tradeoff between preserving high environmental quality and sustaining robust economic growth. This paper is the first study to credibly estimate the impacts of environmental regulations on Chinese manufacturing firms and provides a timely assessment of the economic costs of China's water pollution control policies. We exploit a regression discontinuity design based on the upstream-downstream relationship of water quality monitoring stations in China and find that tighter water quality regulations lead to significant TFP loss for firms located upstream from monitoring stations. This is the case for firms in polluting industries; such a discontinuity is not observed for firms in non-polluting industries.

We estimate that water quality monitoring reduces TFP levels by 26.7% in firms located immediately upstream from monitoring stations. This TFP loss is driven mainly by private Chinese firms instead of state-owned or foreign firms. A closer examination of the TFP effect by year reveals that the impacts of water quality controls have been greater in more recent years, consistent with the fact that environmental regulations in China have tightened over the past decade.

We also investigate the impacts of water quality monitoring on emissions. Using another firm-level dataset, we find that, at the extensive margin, upstream firms emit substantially (52.8%–62.8%) less COD and industrial wastewater than downstream firms; and, at the intensive margin, upstream firms adopt cleaner technology and emit less pollution per value of output (38.7%–49.3%).

Combining both sets of estimates, we calculate the economic costs of China's water pollution control policies. We estimate that a 10% abatement in COD emissions and COD emission intensity can lead to a 2.35%–2.75% and 3.43%–4.21% drop in a polluting firm's TFP respectively. These estimates imply that China's efforts in reducing COD emissions from 2016 to 2020 would cause a total loss in output of 936 to 1,099 billion Chinese yuan (150.5 to 176.7 billion US dollars) in the polluting industries, at least if current monitoring and enforcement practices remain unchanged.

Overall, our findings highlight the negative impacts of environmental regulations on productivity. The estimated efficiency loss is substantial; so high environmental quality comes at high economic cost. This is particularly salient for fast-growing economies that rely heavily on manufacturing.

Political incentives are fundamental to understanding China's environmental regulation. We show that the effect of environmental regulation depends on local officials' chances of promotion and local government's power to manipulate environmental data. Specifically, the TFP difference between upstream and downstream firms becomes twice as large when the city leader has a greater probability of promotion, and it approaches zero when the city leader cannot be promoted. The effect of water quality monitoring on TFP is substantially larger for stations that automatically test and report water quality to the central government.

Our findings also demonstrate that environmental regulations have profound distributional consequences. In the context of water quality monitoring, emission controls in upstream regions will improve the water environment in downstream regions. Upstream firms abate emissions and earn reduced profits, while downstream regions enjoy both higher environmental quality and more rapid economic growth. In the long run, these effects may imply a spatial redistribution of economic activity, population and social welfare.

Nevertheless, our findings do not answer the broader question of whether China's current environmental regulation standards are too aggressive or too lenient, because we do not know Chinese people's willingness to pay for cleaner surface water. After all, little research has been conducted on the socio-economic costs of water pollution in China.[24] To what extent environmental regulations should be designed and enforced, especially in developing countries that rely heavily on manufacturing industries, remains an underexplored research area.

We conclude by pointing out some limitations of this study and offering directions for future research. First, the estimates in this paper are derived in a partial equilibrium framework. We focus on a unique setting that affects only a small set of firms. Large-scale regulation will affect aggregate output and input markets, and our estimates should be interpreted with caution when used to evaluate large-scale environmental policies. Second, our sample covers a relatively short period of time, while firms might be able to better adjust investment and production in the long run. With the growing availability of firm-level longitudinal data, investigating how firms respond to regulation over long periods of time will be an important area for future research. Relatedly, sorting and its subsequent welfare implications are important for long-term impact assessment. Finally, with the expectation of increasingly tighter environmental regulations in China, entrepreneurs and investors may choose to develop businesses in non-

[24] Two exceptions are that (1) Ebenstein (2012) finds that China's surface water pollution has caused an increase in deaths from digestive cancers; and (2) He and Perloff (2016) find that a deterioration in surface water quality from Water Quality Grade Level I to Level III is associated with higher infant mortality.

polluting industries. Tighter environmental regulations in the polluting industries may create externalities affecting non-polluting industries, and there is a lack of rigorous empirical studies to quantify the impacts of such spillover effects on the economy.

REFERENCES

Acemoglu, Daron, and James A. Robinson. "Economics versus politics: Pitfalls of policy advice." *Journal of Economic Perspectives* 27, no. 2 (2013): 173-92.

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. Identification Properties of Recent Production Function Estimators. *Econometrica*, *83*(6), pp.2411-2451.

Ambec, Stefan, Mark A. Cohen, Stewart Elgie, and Paul Lanoie. 2013. "The Porter Hypothesis at 20: Can Environmental Regulation Enhance Innovation and Competitiveness?" *Review of Environmental Economics and Policy* 7 (1): 2-22.

Becker, Randy, and Vernon Henderson. 2000. "Effects of Air Quality Regulations on Polluting Industries." *Journal of Political Economy* 108 (2): 379-421.

Berman, Eli, and Linda TM Bui. 2001. "Environmental Regulation and Productivity: Evidence from Oil Refineries." *Review of Economics and Statistics* 83 (3): 498-510.

Burgess, R., Hansen, M., Olken, B.A., Potapov, P. and Sieber, S., 2012. "The political economy of deforestation in the tropics." *The Quarterly Journal of Economics*, *127*(4), pp.1707-1754.

Cai, Xiqian, Yi Lu, Mingqin Wu, and Linhui Yu. 2016. "Does Environmental Regulation Drive Away Inbound Foreign Direct Investment? Evidence from a Quasi-Natural Experiment in China." *Journal of Development Economics* 123: 73-85.

Calonico, Sebastian, Matias D. Cattaneo, and Max H. Farrell. Forthcoming. "On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference." *Journal of the American Statistical Association.*

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82 (6): 2295-326.

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

Cattaneo, M. D., Michael Jansson, and Xinwei Ma. 2017a. "Simple Local Polynomial Density Estimators." Working paper, University of Michigan.

Cattaneo, M. D., Michael Jansson, and Xinwei Ma. 2017b. "rddensity: Manipulation Testing based on Density Discontinuity." Stata Journal, forthcoming.

Cheng, Ming-Yen, Jianqing Fan, and James S. Marron. 1997. "On Automatic Boundary Corrections." *The Annals of Statistics* 25 (4): 1691-708.

Ebenstein, Avraham. 2012. "The Consequences of Industrialization: Evidence from Water Pollution and Digestive Cancers in China." *Review of Economics and Statistics* 94 (1): 186-201.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. "New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy." *Proceedings of the National Academy of Sciences* 114, no. 39 (2017): 10384-10389.

Fredriksson, Per G., John A. List, and Daniel L. Millimet. 2003. "Bureaucratic Corruption, Environmental Policy and Inbound US FDI: Theory and Evidence." *Journal of Public Economics* 87 (7): 1407-30.

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

Gelman, Andrew, and Guido Imbens. 2017. "Why High-Order Polynomials Should Not Be

Used in Regression Discontinuity Designs." *Journal of Business and Economic Statistics*.

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* 110 (6): 1175-219.

Greenstone, Michael, and B. Kelsey Jack. "Envirodevonomics: A research agenda for an emerging field." *Journal of Economic Literature* 53, no. 1 (2015): 5-42.

Greenstone, Michael, John A. List, and Chad Syverson. 2012. "The Effects of Environmental Regulation on the Competitiveness of US Manufacturing." *National Bureau of Economic Research Working Paper* 18392.

Fan, Haichao, Yu Liu, Larry Qiu, and Xiaoxue Zhao. 2017. "Export to Elude: Evidence from a Tax Enforcement Technology in China." Working Paper, School of Economics Fudan Univeristy.

Hanna, Rema. 2010. "US Environmental Regulation and FDI: Evidence from a Panel of US-Based Multinational Firms." *American Economic Journal: Applied Economics* 2 (3): 158-89.

He, Guojun, and Jeffrey M. Perloff. 2016. "Surface Water Quality and Infant Mortality in China." *Economic Development and Cultural Change* 65 (1): 119-39.

He, Guojun, and Shaoda Wang. "Do college graduates serving as village officials help rural China?." *American Economic Journal: Applied Economics* 9, no. 4 (2017): 186-215.

Henderson, J. Vernon. 1996. "Effects of Air Quality Regulation." *The American Economic Review* 86 (4): 789-813.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4): 1403-48.

Hsieh, Chang-Tai, and Zheng Song. 2015. "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China." *Brookings Papers on Economic Activity*: 328.

Huang, Zhangkai, Lixing Li, Guangrong Ma, and Lixin Colin Xu. Forthcoming. "Hayek, Local Information, and Commanding Heights: Decentralizing State-Owned Enterprises in China." *The American Economic Review.*

Jaffe, Adam B., Steven R. Peterson, Paul R. Portney, and Robert N. Stavins. 1995. "Environmental Regulation and the Competitiveness of US Manufacturing: What does the Evidence Tell Us?" *Journal of Economic Literature* 33 (1): 132-63.

Jia, Ruixue, 2017. "Pollution for promotion."

Kahn, Matthew E., and Erin T. Mansur. 2013. "Do local energy prices and regulation affect the geographic concentration of employment?" *Journal of Public Economics* 101: 105-114.

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

Keiser, David A., and Joseph S. Shapiro. 2017. "Consequences of the Clean Water Act and the Demand for Water Quality." *National Bureau of Economic Research Working Paper* w23070.

Lee, David S., and Thomas Lemieuxa. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281-355.

Lipscomb, Molly, and Ahmed Mushfiq Mobarak. 2017. "Decentralization and pollution spillovers: evidence from the re-drawing of county borders in Brazil." *The Review of Economic Studies* 84, no. 1 (2016): 464-502.

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." *Review of Economics and Statistics* 85 (4): 944-52.

List, J.A. and Sturm, D.M., 2006. "How elections matter: Theory and evidence from environmental policy." *The Quarterly Journal of Economics*, *121*(4), pp.1249-1281.

Liu, Yu. 2017. "Informal Taxation and Firm Performance: Evidence from China." Working Paper, School of Economics, Fudan University.

Martinez-Bravo, Monica. "The role of local officials in new democracies: Evidence from Indonesia." *American Economic Review* 104, no. 4 (2014): 1244-87.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698-714.

Munshi, Kaivan, and Mark Rosenzweig. *Insiders and outsiders: local ethnic politics and public goods provision*. No. w21720. National Bureau of Economic Research, 2015.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6): 1263-1297.

Porter, Michael. 1991. "America's Green Strategy." *Scientific American* 264 (4): 168.

Ryan, Stephen P. 2012. "The Costs of Environmental Regulation in a Concentrated Industry." *Econometrica* 80 (3): 1019-61.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing Like China." *The American Economic Review* 101 (1): 196-233.

Van Biesebroeck, Johannes. 2007. "Robustness of Productivity Estimates." *The Journal of Industrial Economics* 55 (3): 529-69.

Walker, W. Reed. 2011. "Environmental Regulation and Labor Reallocation: Evidence from the Clean Air Act." *The American Economic Review* 101 (3): 442-47.

Walker, W. Reed. 2013. "The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce." *The Quarterly Journal of Economics* 128 (4): 1787-835.

Wang, Shaoda. 2016. "Fiscal Competition and Coordination: Evidence from China." Working Paper, Department of Agricultural and Resource Economics, UC Berkeley.

World Bank. 2007. "Cost of Pollution in China: Economic Estimates of Physical Damages." Washington, DC: World Bank. http://documents.worldbank.org/curated/en/782171468027560055/Cost-of-pollution-in-China-economic-estimates-of-physical-damages.

Xi, Tianyang, Yang Yao, and Muyang Zhang. 2017. "Bureaucratic Capability and Political Opportunism: An Empirical Investigation of City Officials in China." Working Paper, National School of Development, Peking University.

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

Yasar, M., Raciborski, R. and Poi, B. 2008. "Production Function Estimation in Stata using the Olley and Pakes Method." *Stata Journal 8*(2): 221.

Yu, Miaojie. 2015. "Processing Trade, Tariff Reductions and Firm Productivity: Evidence from Chinese Firms." *The Economic Journal* 125 (585): 943-88.

Table 1. Covariate Balance Between Upstream Downstream Firms

	Mean		Mean Difference
	Downstream	Upstream	⩽10km
	(1)	(2)	(4)
Panel A. Time-Invariant Factors			
Year of Opening	1982.55	1984.16	-0.76
	(76.06)	(53.87)	(1.19)
SOE	0.24	0.27	0.02
(1=Yes, 0=Others)	(0.43)	(0.44)	(0.02)
Foreign	0.16	0.19	-0.00
(1=Yes, 0=Others)	(0.37)	(0.39)	(0.02)
Panel B. Pre-2003 Firm-Level Characteristics			
Value Added	8.42	8.48	0.10
(log)	(1.24)	(1.32)	(0.09)
Profit	1.19	2.13	0.30
(billion yuan)	(6.65)	(11.75)	(0.68)
# of Employee	5.06	5.15	0.10
(log)	(1.09)	(1.15)	(0.09)
Capital Stock 1#	8.96	9.04	0.17
(log)	(1.71)	(1.83)	(0.12)
Capital Stock 2#	9.00	9.07	0.15
(log)	(1.70)	(1.83)	(0.12)
Intermediate Input	9.38	9.48	0.11
(log)	(1.32)	(1.38)	(0.10)
Tax	2.22	2.26	0.11
(log)	(2.18)	(2.27)	(0.15)
Obs.	2,190	2,659	

Notes: The sample consists in ASIF firms from 2000–2002 (pre-regulation), that lie within 10km of a water monitoring station. Columns 1–2 report the means and standard deviations of firm characteristics. Column 3 tests the covariate balance between upstream and downstream firms. The difference coefficients are obtained by running OLS regressions of firm characteristics on an upstream dummy and water quality monitoring station and industry fixed effects. Standard errors reported in the parentheses are clustered at the water monitoring station level. # Capital Stock 1 measures the value of fixed assets at the end of year and Capital Stock 2 measures the annual average value of fixed assets.

Table 2. RD Estimates of the Impact of Water Quality Monitoring on TFP

	Polluting Industries			Non-Polluting Industries		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Water Quality Monitoring and TFP						
TFP (log) - Polluting Industries	0.36	0.38	0.43	-0.00	0.02	-0.05
	(0.23)	(0.24)	(0.28)	(0.14)	(0.15)	(0.14)
Bandwidth (km)	4.18	3.88	2.88	4.71	4.14	4.19
Panel B: Water Quality Monitoring and Residual TFP						
TFP (log) - Polluting Industries	0.25*	0.25**	0.33**	-0.01	0.00	0.02
(Station FE Absorbed)	(0.14)	(0.13)	(0.15)	(0.09)	(0.10)	(0.11)
Bandwidth (km)	5.80	5.98	4.82	6.02	5.48	4.26
Panel C: Water Quality Monitoring and Residual TFP						
TFP (log) - Polluting Industries	0.31**	0.31**	0.35**	0.02	0.03	0.03
(Station and Industry FE Absorbed)	(0.15)	(0.15)	(0.16)	(0.08)	(0.08)	(0.09)
Bandwidth (km)	6.56	6.54	5.41	5.553	4.918	4.329
Obs.	6,582	6,582	6,582	12,422	12,422	12,422
Kernel	Triangle	Epanech.	Uniform	Triangle	Epanech.	Uniform

Notes: Each cell in the table represents a separate regression. TFP is estimated using Olley and Pakes (1996) method. The discontinuities at monitoring stations are estimated using local linear regressions and MSE-optimal bandwidth proposed by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods. Standard errors clustered at the monitoring station level are reported below the estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table 3. Heterogeneous Impacts of the Impact of Water Quality Monitoring on TFP

	Residual TFP – Polluting Industries			Residual TFP – Non-Polluting Industries		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: By Ownership						
Private Firms	0.34**	0.37**	0.31*	0.04	0.04	0.03
	(0.17)	(0.18)	(0.18)	(0.08)	(0.08)	(0.09)
Obs.	5,636	5,636	5,636	10,084	10,084	10,084
BW	5.965	5.590	5.087	6.052	6.059	5.537
SOEs	-0.31	-0.16	0.23	-0.13	-0.10	-0.01
	(0.52)	(0.54)	(0.50)	(0.25)	(0.25)	(0.27)
Obs.	635	635	635	1,357	1,357	1,357
BW	4.282	4.474	4.407	4.724	4.545	3.955
Foreign Firms	-0.06	-0.07	-0.11	-0.12	-0.15	0.02
	(0.27)	(0.28)	(0.31)	(0.40)	(0.42)	(0.25)
Obs.	1,104	1,104	1,104	2,427	2,427	2,427
BW	6.927	6.541	5.479	3.287	3.070	4.286
Panel B: By Size						
Small Firm	0.16	0.16	0.14	-0.13	-0.12	-0.14
(Empl<50)	(0.29)	(0.24)	(0.25)	(0.15)	(0.14)	(0.13)
Obs.	1,998	1,998	1,998	4,357	4,357	4,357
BW	7.400	8.096	6.273	3.823	3.985	3.758
Large Firm	0.41***	0.42***	0.40**	-0.01	0.01	0.04
(Empl≥50)	(0.14)	(0.15)	(0.17)	(0.09)	(0.10)	(0.11)
Obs.	5,369	5,369	5,369	9,691	9,691	9,691
BW	4.825	4.738	4.520	4.610	4.674	4.513
Panel C: By Firm Age						
Old Firms	0.33*	0.39**	0.45**	0.05	0.05	0.04
	(0.17)	(0.19)	(0.21)	(0.09)	(0.09)	(0.09)
Obs.	4,481	4,481	4,481	8,373	8,373	8,373
BW	6.695	5.881	4.624	5.432	5.199	4.526
Young Firms	0.48**	0.51**	0.39	-0.03	-0.00	0.07
	(0.19)	(0.21)	(0.26)	(0.16)	(0.18)	(0.20)
Obs.	1,438	1,438	1,438	2,627	2,627	2,627
BW	3.768	3.537	3.798	5.768	5.084	4.357
Kernel	Triangle	Epanech.	Uniform	Triangle	Epanech.	Uniform

Notes: Each cell in the table represents a separate regression. Monitoring station and industry fixed effects are absorbed before estimating regression discontinuity. In columns 1–3, we report the estimated discontinuity for polluting industries, and in columns 4–6, we report the estimated discontinuity for non-polluting industries. Local linear regression and MSE-optimal bandwidth proposed by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods are used for the estimation. Conventional local linear regression discontinuity standard errors clustered at the monitoring station level are reported below the estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table 4. RD Estimates of the Impact of Water Quality Monitoring on TFP by Year

	Residual TFP – Polluting Industries			Residual TFP – Non-Polluting Industries		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Before and After 2003						
Before 2003	0.09	0.10	0.11	0.01	0.01	0.06
	(0.19)	(0.20)	(0.24)	(0.12)	(0.13)	(0.15)
Obs.	2,570	2,570	2,570	4,565	4,565	4,565
After 2003	0.36**	0.35**	0.40**	0.03	0.04	0.07
	(0.16)	(0.16)	(0.17)	(0.08)	(0.09)	(0.10)
Obs.	5,916	5,916	5,916	10,992	10,992	10,992
Panel B. by Year						
Year 2000	-0.09	-0.04	-0.00	-0.22	-0.21	-0.11
	(0.34)	(0.32)	(0.36)	(0.17)	(0.18)	(0.16)
Obs.	1,411	1,411	1,411	2,428	2,428	2,428
Year 2001	-0.02	-0.01	-0.04	-0.07	-0.05	-0.19
	(0.21)	(0.21)	(0.24)	(0.17)	(0.18)	(0.17)
Obs.	1,411	1,411	1,411	2,428	2,428	2,428
Year 2002	0.04	0.09	0.05	0.03	0.01	-0.02
	(0.20)	(0.20)	(0.25)	(0.13)	(0.13)	(0.12)
Obs.	2,106	2,106	2,106	3,644	3,644	3,644
Year 2003	0.30	0.34	0.37*	0.04	0.04	0.04
	(0.29)	(0.29)	(0.21)	(0.16)	(0.16)	(0.15)
Obs.	2,367	2,367	2,367	3,888	3,888	3,888
Year 2004	0.12	0.14	0.21	0.08	0.06	0.06
	(0.30)	(0.32)	(0.31)	(0.11)	(0.11)	(0.11)
Obs.	3,288	3,288	3,288	5,509	5,509	5,509
Year 2005	0.31	0.35	0.35	-0.04	-0.05	-0.06
	(0.24)	(0.25)	(0.26)	(0.15)	(0.15)	(0.15)
Obs.	3,750	3,750	3,750	6,296	6,296	6,296
Year 2006	0.48**	0.52**	0.61**	0.01	0.01	0.03
	(0.22)	(0.25)	(0.27)	(0.14)	(0.15)	(0.16)
Obs.	3,981	3,981	3,981	6,969	6,969	6,969
Year 2007	0.37**	0.38*	0.42*	0.14	0.15	0.17*
	(0.19)	(0.20)	(0.22)	(0.09)	(0.09)	(0.10)
Obs.	4,460	4,460	4,460	8,103	8,103	8,103
Kernel	Triangle	Epanech.	Uniform	Triangle	Epanech.	Uniform

Notes: Each cell in the table represents a separate regression. Monitoring station and industry fixed effects are absorbed before estimating regression discontinuity. In columns 1-3, we report the estimated discontinuity for polluting industries, and in columns 4-6, we report the estimated discontinuity for non-polluting industries. Local linear regression and MSE-optimal bandwidth proposed by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods are used for the estimation. Conventional local linear regression discontinuity standard errors clustered at the monitoring station level are reported below the estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table 5. Instrumental Variable Estimation using Hydrological Stations

	Polluting Industries		Non–Polluting Industries	
	Upstream	TFP (log)	Upstream	TFP (log)
	(1)	(2)	(3)	(4)
Upstream Hydrological Station	0.38**		0.30**	
	(0.18)		(0.14)	
Upstream Monitoring Station		-0.35**		-0.08
		(0.15)		(0.16)
Specification	1st Stage	2SLS	1st Stage	2SLS
Station FE	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y
Observations	4,445	4,462	8,976	8,981
F Statistic	10.48	0.03	22.82	1.18

Notes: Each column in the table represents a separate regression. We define "upstream monitoring station" as a dummy indicator for whether a firm is upstream from a monitoring station within a 10 km range, and similarly, we define "upstream hydrological station" as a dummy indicator for whether a firm is upstream from a hydrological station within a 10 km range. Our outcome of interest is firm-level TFP estimated using Olley and Pakes (1996) method, our endogenous variable is "upstream monitoring station", and our instrumental variable is "upstream hydrological station". We present first-stage results and IV 2SLS results separately for firms in polluting industries (columns 1 and 2) and firms in non-polluting industries (columns 3 and 4). Monitoring station fixed effects are controlled for in all specifications. Standard errors are clustered at the monitoring station level. * significant at 10% ** significant at 5% *** significant at 1%.

Table 6. RD Estimates using Placebo Downstream Firms

	Polluting Industries			Non-Polluting Industries		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Water Quality Monitoring and TFP						
TFP (log) - Polluting Industries	0.36	0.44*	0.26	-0.18	-0.20	-0.13
	(0.23)	(0.26)	(0.29)	(0.16)	(0.15)	(0.18)
Panel B: Water Quality Monitoring and Residual TFP						
TFP (log) - Polluting Industries	0.48**	0.52**	0.61***	0.13	0.11	0.14
(Station and Industry FE Absorbed)	(0.20)	(0.21)	(0.23)	(0.13)	(0.11)	(0.12)
Obs.	4,435	4,435	4,435	8,001	8,001	8,001
Kernel	Triangle	Epanech.	Uniform	Triangle	Epanech.	Uniform

Notes: Each cell in the table represents a separate regression, where each control firm is replaced by its best match in the whole sample from a pre-2003 nearest neighbour matching (based on TFP, industry, and other basic characteristics). TFP is estimated using Olley and Pakes (1996) method. The discontinuities at monitoring stations are estimated using local linear regression and MSE-optimal bandwidth proposed by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods. Standard errors clustered at the monitoring station level are reported below the coefficients in columns 1–5 and conventional local linear regression discontinuity standard errors clustered at the monitoring station level are reported in columns 6-8. * significant at 10% ** significant at 5% *** significant at 1%.

Table 7. Robustness Checks: Impact of Water Quality Monitoring on TFP

	TFP – Polluting Industries		
	(1)	(2)	(3)
Panel A. Alternative Ways to Estimate RD and Standard Errors			
Bias-corrected RD Estimates	0.35**	0.34**	0.38**
	(0.15)	(0.15)	(0.16)
Bias-corrected Robust Estimates	0.35*	0.34*	0.38**
	(0.19)	(0.19)	(0.19)
Panel B. Alternative Ways to Choose Optimal Bandwidth			
Bandwidth Chosen by MSE-Two Selector	0.30**	0.29*	0.25
	(0.15)	(0.15)	(0.17)
Bandwidth Chosen by MSE-Sum Selector	0.31**	0.30**	0.34**
	(0.15)	(0.15)	(0.16)
Bandwidth Chosen by CER-D Selector	0.38**	0.40**	0.43**
	(0.19)	(0.19)	(0.20)
Bandwidth Chosen by CER-Two Selector	0.35**	0.39**	0.48**
	(0.17)	(0.17)	(0.20)
Bandwidth Chosen by CER-Sum Selector	0.37**	0.39**	0.44**
	(0.18)	(0.19)	(0.20)
Panel C. Placebo Tests			
Move Monitoring Stations Upstream by 5km	0.12	0.13	0.11
	(0.16)	(0.16)	(0.16)
Move Monitoring Stations Upstream by 10km	-0.08	-0.09	-0.08
	(0.11)	(0.11)	(0.12)
Move Monitoring Stations Downstream by 5km	0.13	0.15	0.11
	(0.09)	(0.09)	(0.11)
Move Monitoring Stations Downstream by 10km	0.03	0.05	0.07
	(0.16)	(0.15)	(0.17)
Kernel	Triangle	Epanech.	Uniform

Notes: Each cell in the table represents a separate regression. Monitoring station and industry fixed effects are absorbed before estimating regression discontinuity. Conventional local linear regression discontinuity standard errors clustered at the monitoring station level are reported below the estimates (except Panel A). Local linear regression and MSE-optimal bandwidth selected by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods are used for the estimation (except Panel B). Monitoring station and industry fixed effects are absorbed before estimating regression discontinuity. * significant at 10% ** significant at 5% *** significant at 1%.

Table 8. Channels: RD Estimates on other Measures

	Conventional Local RD		Bias-Corrected RD		Bias-Corrected Robust	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Output Related						
Value-Added	0.10	0.09	0.04	0.04	0.04	0.04
(log)	(0.20)	(0.21)	(0.20)	(0.21)	(0.26)	(0.27)
Profit	1.07	1.18	1.34*	1.49*	1.34	1.49
(10 million yuan)	(0.74)	(0.80)	(0.74)	(0.80)	(0.89)	(0.95)
Panel B. Input Related						
Employees	-0.11	-0.11	-0.16	-0.13	-0.16	-0.13
(log)	(0.15)	(0.14)	(0.15)	(0.14)	(0.19)	(0.19)
Intermediate Input	-0.07	-0.04	-0.14	-0.11	-0.14	-0.11
(log)	(0.21)	(0.23)	(0.21)	(0.23)	(0.29)	(0.30)
Capital Stock 1 #	-0.44	-0.42	-0.55**	-0.54**	-0.55	-0.54
(log)	(0.27)	(0.27)	(0.27)	(0.27)	(0.34)	(0.34)
Capital Stock 2 #	-0.50*	-0.48*	-0.63**	-0.61**	-0.63*	-0.61*
(log)	(0.28)	(0.28)	(0.28)	(0.28)	(0.36)	(0.35)
Panel C. Naïve Labor and Capital Productivity						
VA/Employee	0.23	0.31	0.26	0.36	0.26	0.36
(log)	(0.20)	(0.22)	(0.20)	(0.22)	(0.25)	(0.27)
VA/Cap Stock 1	0.27***	0.28***	0.26**	0.29***	0.26**	0.29**
(log)	(0.10)	(0.11)	(0.10)	(0.11)	(0.12)	(0.13)
VA/Cap Stock 2	0.23**	0.26**	0.23**	0.27**	0.23*	0.27**
(log)	(0.11)	(0.11)	(0.11)	(0.11)	(0.14)	(0.14)
Panel D. Porter Hypothesis						
R&D	-0.09	-0.10	-0.07	-0.10	-0.07	-0.10
(log)	(0.35)	(0.37)	(0.35)	(0.37)	(0.55)	(0.57)
Kernel	Triangle	Epanech.	Triangle	Epanech.	Triangle	Epanech.

Notes: Each cell in the table represents a separate regression. Monitoring station and industry fixed effects are absorbed before estimating regression discontinuity. We use post-2003 collapsed data to estimate the regression discontinuities. Local linear regression and MSE-optimal bandwidth proposed by Calnoico et al. (2014) and Calonico (2017) for different kernel weighting methods are used for the estimation. Station and industry fixed effects are controlled in all regressions. Standard errors are clustered at the monitoring station level, and reported below the estimates. # Capital Stock 1 is measured by the value of fixed assets at the end of each year and Capital Stock 2 is the annual average value of fixed assets. * significant at 10% ** significant at 5% *** significant at 1%.

Table 9. Political Economy of Water Quality Monitoring

	Conventional Local RD		Bias-Corrected RD		Bias-Corrected Robust	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. "Double Standard"						
Waste Discharge Fee	-1.15**	-1.07**	-1.41***	-1.32**	-1.41**	-1.32**
(log)	(0.51)	(0.53)	(0.51)	(0.53)	(0.57)	(0.60)
Panel B. Strong vs. Weak Political Incentives						
TFP (log) - Strong Incentive	0.57***	0.59***	0.63***	0.66***	0.63***	0.66***
	(0.19)	(0.20)	(0.19)	(0.20)	(0.21)	(0.23)
TFP (log) - Weak Incentive	0.01	0.08	0.00	0.07	0.00	0.07
	(0.23)	(0.24)	(0.23)	(0.24)	(0.29)	(0.31)
Panel C. Automatic vs. Manual Monitoring Stations						
TFP (log) - Automatic Stations	0.92	1.01*	1.11*	1.22**	1.11	1.22*
	(0.59)	(0.57)	(0.59)	(0.57)	(0.74)	(0.71)
TFP (log) - Manual Stations	0.26*	0.26*	0.27*	0.27*	0.27	0.27
	(0.15)	(0.15)	(0.15)	(0.15)	(0.18)	(0.18)
Kernel	Triangle	Epanech.	Triangle	Epanech.	Triangle	Epanech.

Notes: Each cell in the table represents a separate regression. Monitoring station and industry fixed effects are absorbed before estimating regression discontinuity. We focus on polluting firms and use post-2003 collapsed data to estimate the regression discontinuities. Local linear regression and MSE-optimal bandwidth proposed by Calnoico et al. (2014) and Calonico (2017) for different kernel weighting methods are used for the estimation. Panel A examines how tax and waste discharge fee collected by the government differ between upstream and downstream firms. Panel B estimates the discontinuities separately using the subsamples where the Prefecture Party Secretary has or does not have strong promotion incentives (age<=57 vs. age >57). Panel C estimates the discontinuities separately for automatic and manual monitoring stations. * significant at 10% ** significant at 5% *** significant at 1%.

Table 10. RD Estimates of the Impact of Water Quality Monitoring on Emissions

	Conventional Local RD		Bias-Corrected		Bias-Corrected Robust	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: COD Emission						
COD Emission (log)	0.83*	0.75*	0.99**	0.92**	0.99**	0.92*
	(0.44)	(0.42)	(0.44)	(0.42)	(0.49)	(0.47)
COD Emission Intensity (log)	0.55**	0.49*	0.68**	0.62**	0.68**	0.62**
	(0.27)	(0.26)	(0.27)	(0.26)	(0.32)	(0.31)
Panel B: Wastewater Discharge						
Waste Water Discharge (log)	0.39	0.39	0.49	0.50	0.49	0.50
	(0.33)	(0.35)	(0.33)	(0.35)	(0.40)	(0.42)
Waste Water Discharge Intensity (log)	0.34*	0.33*	0.42**	0.41**	0.42*	0.41*
	(0.20)	(0.20)	(0.20)	(0.20)	(0.23)	(0.22)
Bandwidth Selector	MSE	MSE	MSE	MSE	MSE	MSE
Obs.	9,888	9,888	9,888	9,888	9,888	9,888
Kernel	Triangle	Epanech.	Triangle	Epanech.	Triangle	Epanech.

Notes: Each cell in the table represents a separate regression. Monitoring station fixed effects are absorbed before estimating regression discontinuity. Local linear regression and MSE-optimal bandwidth selected by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods are used for the estimation. Conventional local linear regression discontinuity standard errors clustered at the monitoring station level are reported below the estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table 11. Predicted Effects of Water Quality Monitoring on TFP and COD

	Model 1	Model 2	Model 3
	(1)	(2)	(3)
Panel A. TFP Effects for Large and Small Firms (Measured by Industry Output Value)			
20% (log Rev ~ 9.01)	0.14	0.10	0.11
	(0.24)	(0.23)	(0.23)
40% (log Rev ~ 9.58)	0.27	0.32	0.32
	(0.23)	(0.22)	-0.22
60% (log Rev ~ 10.16)	0.42*	0.50**	0.49**
	(0.22)	(0.21)	0.21
80% (log Rev ~ 10.92)	0.62***	0.66***	0.65***
	(0.22)	(0.20)	(0.20)
Panel B. COD Effect for Large and Small emitters (Measured by COD Emissions)			
20% (log COD ~ 5.97)	0.14	0.12	0.19
	0.44	(0.43)	(0.44)
40% (log COD ~ 7.46)	0.93**	0.96**	0.99**
	(0.43)	(0.43)	(0.44)
60% (log COD ~ 8.70)	1.59***	1.62***	1.59***
	(0.43)	(0.43)	(0.43)
80% (log COD ~ 10.18)	2.38***	2.34***	2.27***
	(0.43)	(0.43)	(0.42)
Heterogeneity Specification	Linear	Quadratic	Cubic

Notes: This table reports the predicted effects of water quality monitoring on TFP and COD emission intensity. Monitoring station and industry fixed effects are absorbed before estimating regression discontinuity. In Panel A, we explore the TFP heterogeneity at different revenue levels, and in Panel B, we explore the COD intensity heterogeneity at different COD emission levels. We use parametric RD to estimate the heterogeneous effects using different heterogeneity functional forms. We choose the polynomial RD specifications that generate the closest estimates to the non-parametric estimates reported in Table 2 and Table 8 as the baselines. We then include interactions to test the heterogeneity. "Linear" means that we use a linear interaction between the downstream dummy and log revenue (or log COD), and "quadratic" means we interact the downstream dummy with a quadratic function of log revenue (or log COD). Panel A shows that the monitoring effect is only significant for large firms, and Panel B shows that the monitoring effect is only significant for large emitters.

Table 12. Economic Costs of COD Abatement

	Conventional		Bias-Corrected	
	(1)	(2)	(3)	(4)
Panel A. Estimated Effect of Water Quality Monitoring				
Effect on log TFP	0.31**	0.31**	0.35**	0.34**
	(0.15)	(0.15)	(0.15)	(0.15)
Effect on log COD Emission	0.83*	0.75*	0.99**	0.92**
	(0.44)	(0.42)	(0.44)	(0.42)
Effect on log COD Emission Intensity	0.55**	0.49*	0.68**	0.62**
	(0.27)	(0.26)	(0.27)	(0.26)
Panel B. Estimated Economic Costs Estimates:				
TFP Loss if all Polluting Firms are Monitored	26.66%	26.66%	29.53%	28.82%
TFP Loss per 10% COD Emission Abatement	2.49%	2.75%	2.35%	2.46%
TFP Loss per 10% COD Emission Intensity Reduction	3.75%	4.21%	3.43%	3.65%
Total Output Loss if all Polluting Firms are Monitored (billion CNY)	3988.9	3988.9	4599.6	4444.6
Total Output Loss in the Polluting Industry during the 11th Five-Year Plan (billion CNY), A	351.98	390.86	332.60	348.16
Total Output Loss in the Polluting Industry per 2.5% COD Abatement (billion CNY), A	68.64	76.01	64.95	67.91
Total Output Loss in the Polluting Industry per 10% COD Abatement (billion CNY), A	279.79	310.48	264.48	276.77
Total Output Loss in the Polluting Industry per 2.5% COD Abatement (billion CNY), B	242.91	269.00	229.85	240.34
Total Output Loss in the Polluting Industry per 10% COD Abatement (billion CNY), B	990.2	1098.8	936.0	979.5
Kernel	Triangle	Epanech.	Triangle	Epanech.
Gross Output Value in the Polluting Industry in 2006 (billion CNY), A	10975.7			
Gross Output Value in the Polluting Industry in 2015 (billion CNY), B	38844.9			

Notes: The gross output values were obtained from the website of the National Bureau of Statistics. A: calculation is based on gross output value (of industries above designated size) in 2006; B: calculation is based on gross output value (of industries above designated size) in 2015.

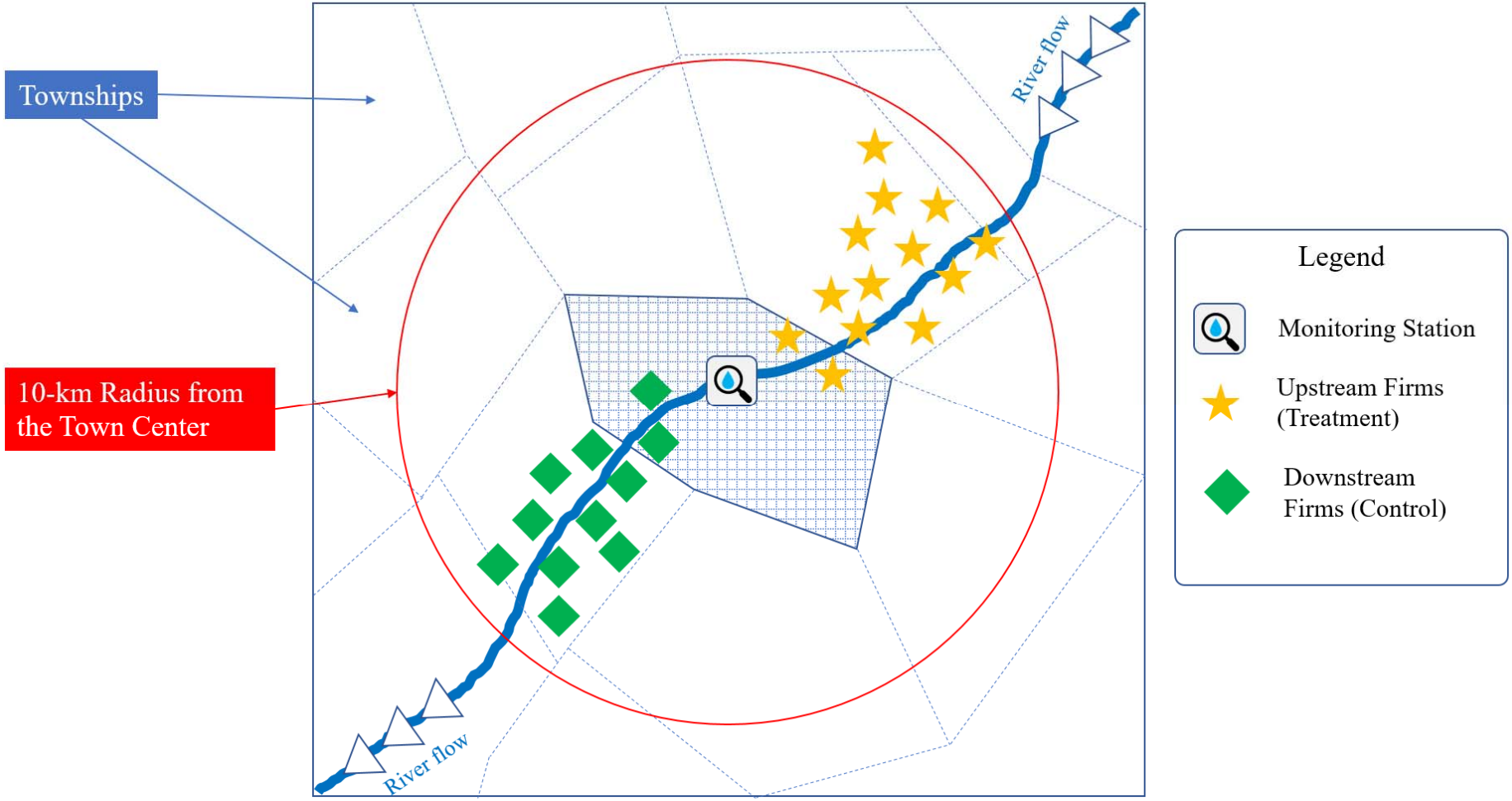


Figure 1. Illustrating the Identification Strategy

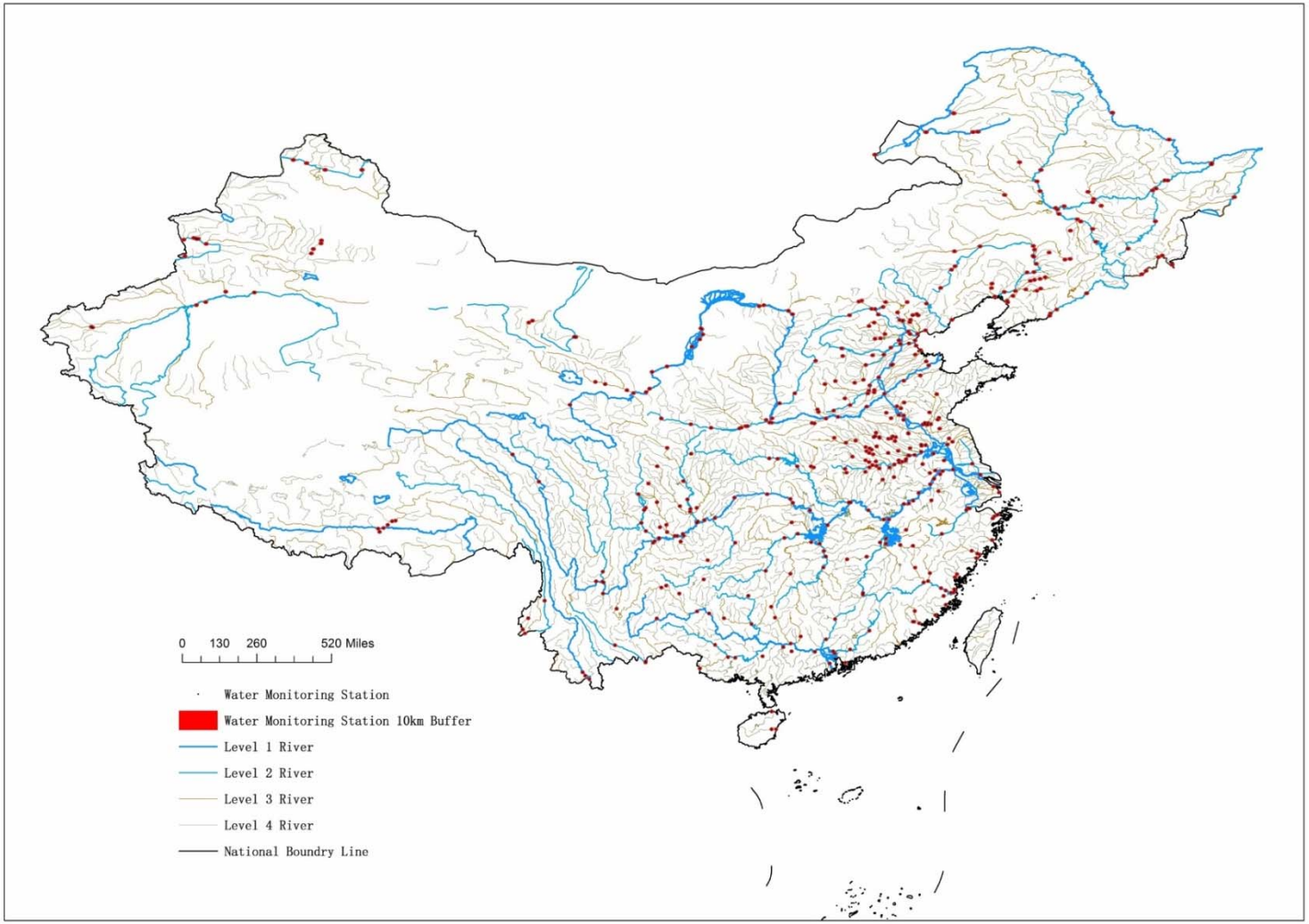
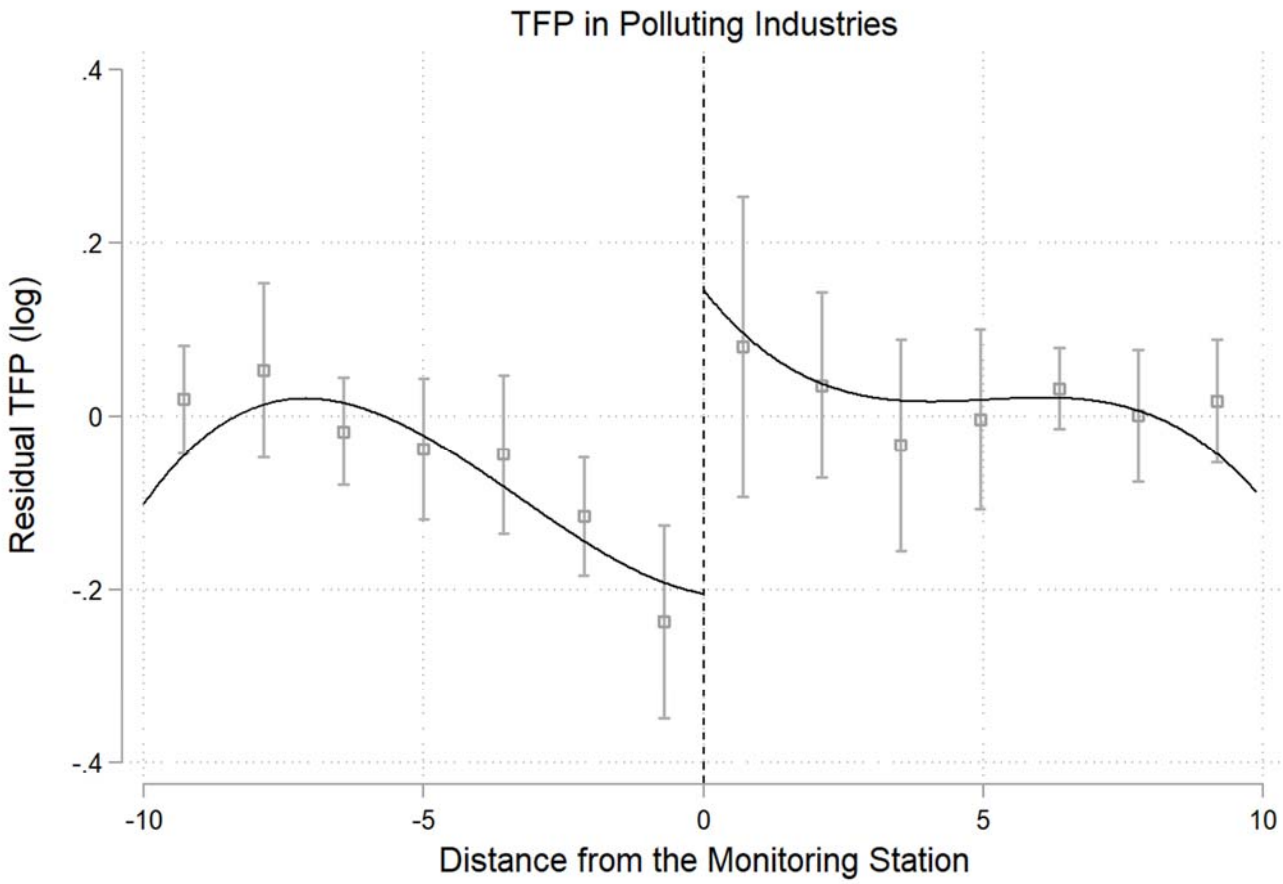


Figure 2. Distribution of Surface Water Quality Monitoring Stations

Panel A: TFP in Polluting Industries



Panel B: TFP in Non-Polluting Industries

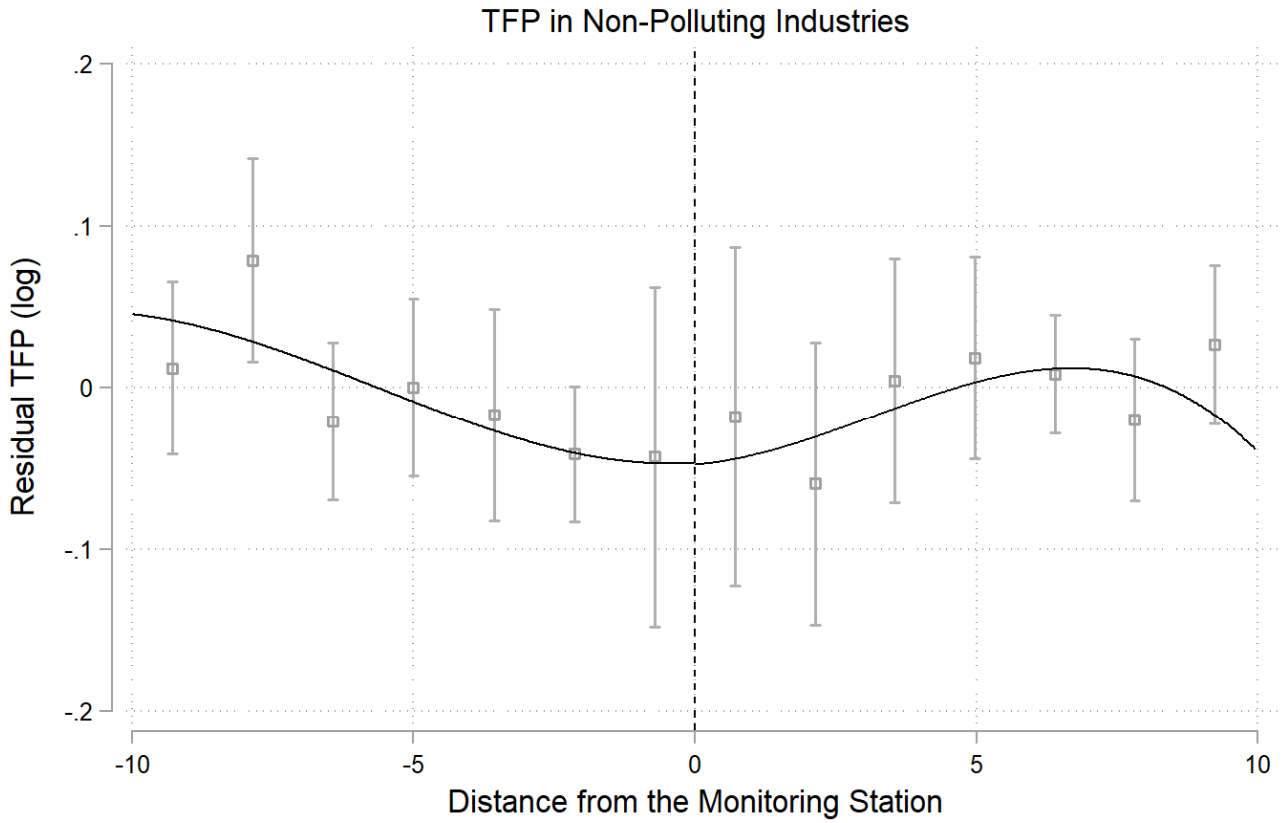


Figure 3. RD Plot: Effects of Water Quality Monitoring on TFP

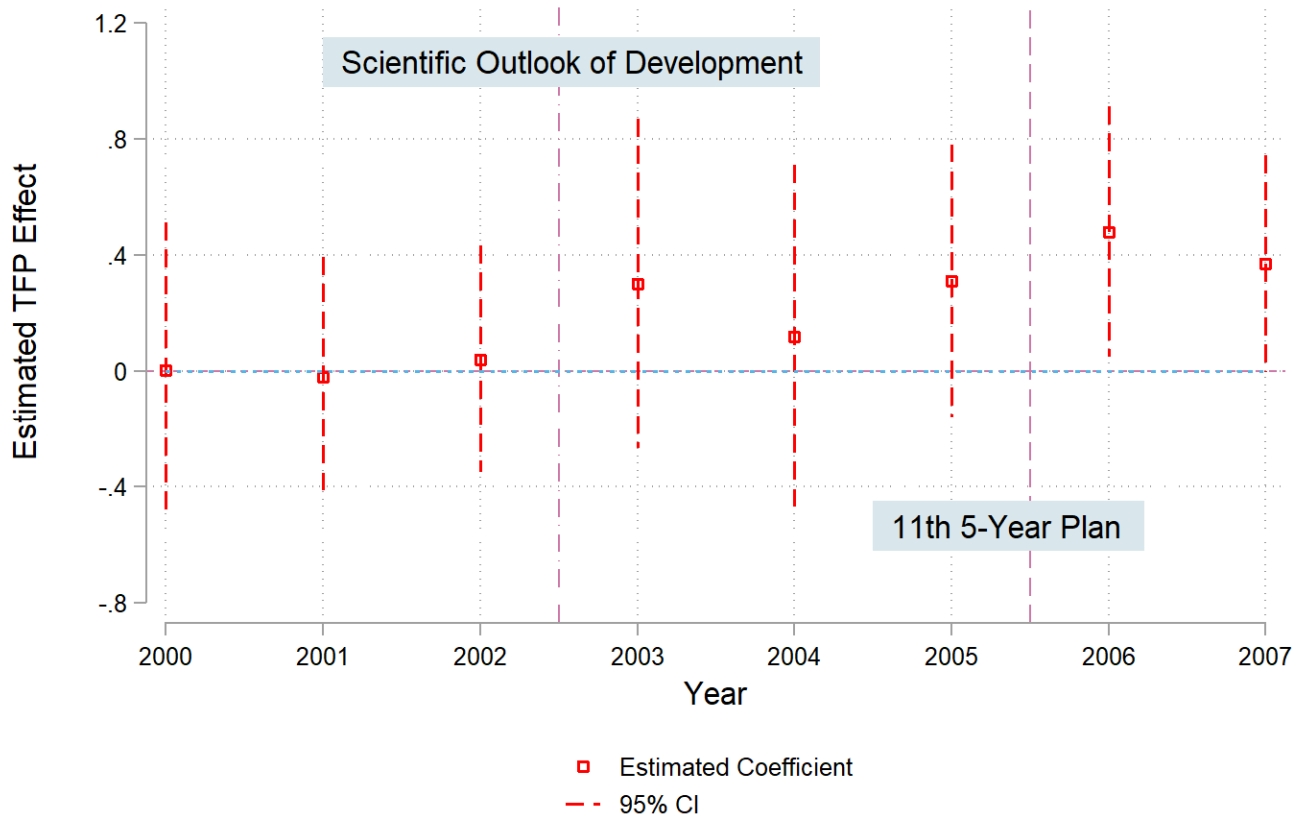


Figure 4. RD Estimates by Year

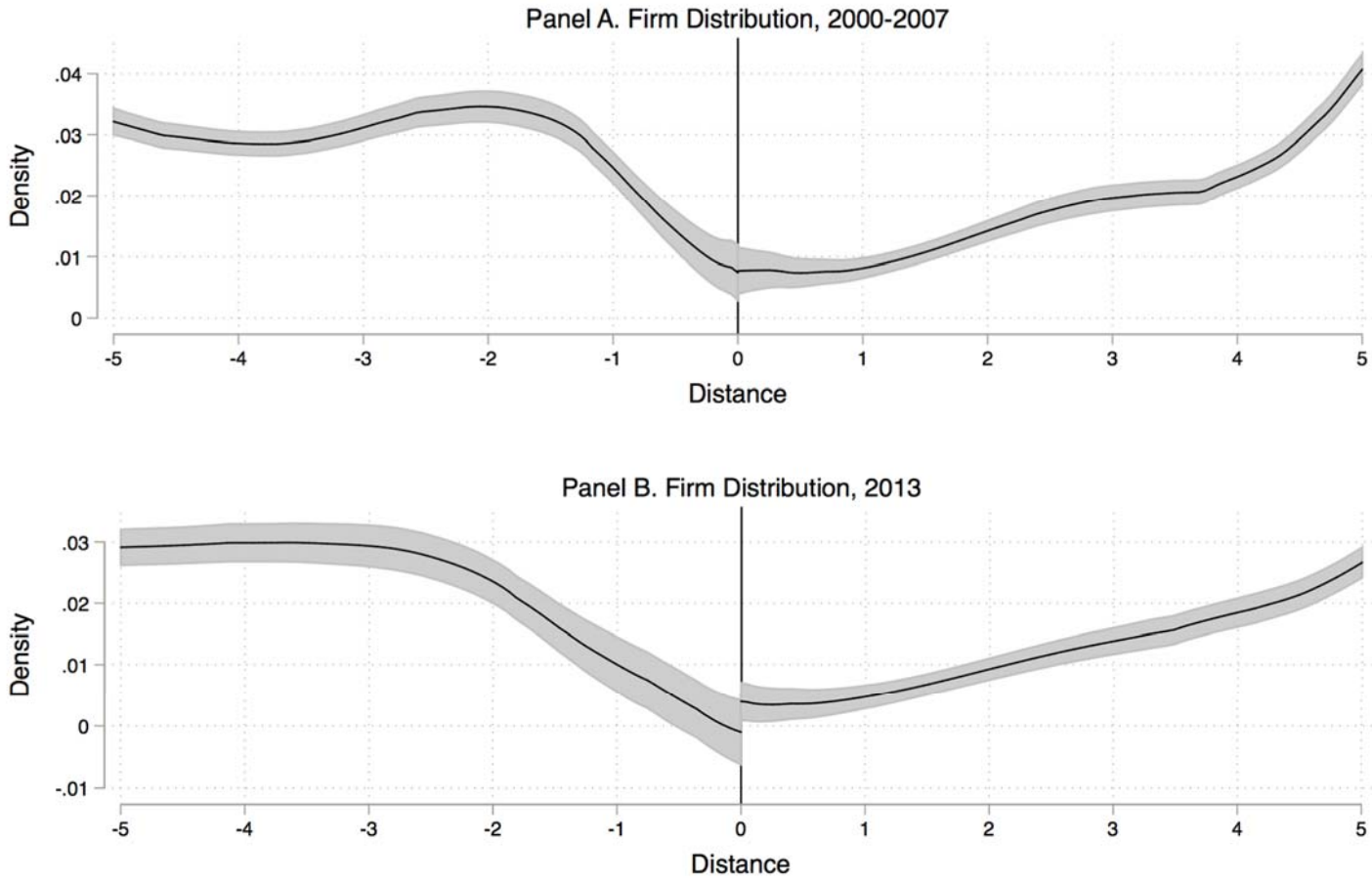


Figure 5. Distribution of Firms

Online Appendix to
"Environmental Regulation and Firm Productivity in China: Estimates from a Regression Discontinuity Design"

Appendix A. Estimation of TFP using Olley-Pakes Method

Our Olley-Pakes TFP measure is constructed based on Brandt et al. (2012) using the Annual Survey of Industrial Firms (ASIF) dataset from 2000 to 2007. We made slight changes to the estimations of some key parameters to improve the accuracy of productivity measurement in the ASIF dataset, as suggested by Yang (2015). We explain in detail how we construct these key parameters.

Gross Output

Following the literature, we use production value, instead of sales, as the gross output measure. Production value and sales differ slightly due to the change in inventories. The former is more closely related to input and productivity, and thus more relevant for TFP estimation.

When constructing output deflators, we follow Yang (2015) by using output price indexes for every 2-digit industry in each year from the "Urban Price Yearbook 2011" published by the National Bureau of Statistics. Because those price indexes are linked across different years, we can use them to deflate yearly nominal output to real output in 2000.

Value Added

When constructing real value added, we subtract from the aforementioned real output the goods purchased for resale, indirect taxes, and material inputs.

We construct input deflators from National Input-Output tables in 1997, 2002, and 2007, to take into account the dynamics of input price in different sectors, as suggested by Yang (2015). By doing so, we are able to deflate nominal inputs in each sector in each year to the real values in 2000.

Employment and Wages

The ASIF dataset contains information on the number of employees and the compensation for labor, including wages, employee supplementary benefits, and insurance. We follow Brandt et al. (2012) to sum up wages, benefits, and insurance as a proxy for total labor compensation.

Capital Stock and Investment

In the ASIF dataset, firms report the value of their fixed capital stock at original purchase prices, as well as capital stock at original purchased prices less accumulated depreciation. Because these values are the sum of nominal values in all the past years, they cannot be taken directly to proxy for real capital stock. To back out the real capital stock and construct real investment from this variable, we follow the approach suggested by Yang (2015).

For each year after the first period, we first take the difference between "current capital stock" and "capital stock in the previous period," then deflate it according to the previously calculated price indexes for this period. For observations in the first period of the panel, we assume that, from the firm's establishment until this first period, it had on average the same increasing trend in investment rate as the 2-digit sector average value, which can be collected from the yearbooks published by the National Bureau of Statistics. Under this assumption, together with the nominal capital stock in the first period, nominal capital stock when established, and relevant deflators, we are able recover the real investment and real capital stock in the first period as well.

TFP Estimation

With the key variables constructed, we follow the literature and use the Olley and Pakes (1996) approach to estimate TFP. This approach addresses both simultaneity and selection problems that are salient in the traditional Solow-residual type TFP estimates. For implementation, we use the Stata package provided by Yasar et al. (2008); please refer to their manual for the details of the estimation.

Appendix B. Conceptual Framework

We provide a conceptual framework that helps to explain the empirical findings. We focus on firms' production decisions and address how environmental regulations can affect their TFP. We assume that firms produce homogeneous goods, with a Hicks-neutral continuously differentiable production function$Q(K, L)$, where K represents capital, L represents labor, and $Q_k, Q_l > 0; Q_{kk}, Q_{ll} < 0$.

When a firm produces output Q, emissions are generated as a by-product and are an increasing function of output Q. The firm can reduce its emissions by employing extra (non-productive) labor L_E and/or capital K_E. The final emission level is therefore a continuously differentiable function $E(Q, K_E, L_E)$. We assume that $E_1 > 0, E_{11} > 0; E_2 > 0, E_{22} < 0; E_3 > 0, E_{33} < 0$ and $E_{23} = E_{32} = 0$.

We model the government's environmental regulations as a unit tax (fine), t, on firm's emissions E. A firm maximizes its profit by setting K, L, K_E, L_E as follows:

(1) $$\max_{K,L,K_E,L_E} \pi = p \cdot Q(K, L) - r \cdot (K + K_E) - w \cdot (L + L_E) - t \cdot E(Q, K_E, L_E)$$

where p represents the market output price, r represents the capital price or interest rate, and w represents wages.

The first order conditions for the firm's profit maximization problem are therefore:

(2) $$\frac{\partial \pi}{\partial K} = p \cdot Q_k - r - t \cdot E_1 \cdot Q_k = 0$$

(3) $$\frac{\partial \pi}{\partial L} = p \cdot Q_l - w - t \cdot E_1 \cdot Q_l = 0$$

(4) $$\frac{\partial \pi}{\partial K_E} = -r - t \cdot E_2 = 0$$

(5) $$\frac{\partial \pi}{\partial L_E} = -w - t \cdot E_3 = 0$$

Applying the implicit function theorem, we can prove the following:

(6) $$\frac{\partial K}{\partial t} < 0, \frac{\partial L}{\partial t} < 0; \ \frac{\partial K_E}{\partial t} > 0, \frac{\partial L_E}{\partial t} > 0;$$

(7) $$\frac{\partial E/\partial t}{E} < \frac{\partial Q/\partial t}{Q} < 0;$$

(8) $$\frac{\partial Q/\partial t}{Q} < \frac{\partial (K+K_E)/\partial t}{(K+K_E)}; \ \frac{\partial Q/\partial t}{Q} < \frac{\partial (L+L_E)/\partial t}{(L+L_E)}.$$

Proposition 1. An increase in the emissions tax reduces TFP.

Proof. By definition, TFP $= \frac{p \cdot Q(K,L)}{r \cdot (K+K_E)+w \cdot (L+L_E)}$; and we therefore have the following:

$$\text{(11)} \qquad \frac{\partial \text{TFP}}{\partial t} = \frac{\partial [\frac{p \cdot Q(K,L)}{r \cdot (K+K_E)+w \cdot (L+L_E)}]}{\partial t} = p \cdot \frac{r\left[\frac{\partial Q}{\partial t} \cdot (K+K_E) - Q \cdot \frac{\partial (K+K_E)}{\partial t}\right] + w\left[\frac{\partial Q}{\partial t} \cdot (L+L_E) - Q \cdot \frac{\partial (L+L_E)}{\partial t}\right]}{[r \cdot (K+K_E)+w \cdot (L+L_E)]^2} < 0$$

where the inequality follows from Equation (8).

Proposition 2. An increase in the emission tax t reduces the emission level E and emission intensity $\frac{E(Q,K_E,L_E)}{Q}$.

Proof. Taking the derivative of emissions with respect to the emission tax, we have:

$$\text{(9)} \qquad \frac{\partial E}{\partial t} = E_1 \cdot \frac{\partial E}{\partial t} + E_2 \cdot \frac{\partial K_E}{\partial t} + E_3 \cdot \frac{\partial L_E}{\partial t} < 0;$$

where the inequality follows from Equations (6) and (7).

For emission intensity, we have:

$$\text{(10)} \qquad \frac{\partial (E/Q)}{\partial t} = \frac{\frac{\partial E}{\partial t} \cdot Q - \frac{\partial E}{\partial t} \cdot E}{E^2} < 0$$

where the inequality follows from Equation (7).

In this model, we implicitly assume that production has no effect on the market price. This assumption is likely to hold in our empirical setting because we focus on a small set of firms concentrated in a small geographical area. On the one hand, these firms face the same market because they are located close to each other; on the other hand, as there are many other firms and buyers in the market, local water quality regulations cannot affect the output market prices. This is important because we cannot directly measure output quantity Q in our firm-level production data. Instead, we can only measure revenue $p \cdot Q(K, L)$. Because firms are price-takers in our setting, we can translate the effects of environmental regulation on revenue-based TFP to real (output-based) TFP. In the case where prices depend on marginal cost, we will underestimate the true TFP effect because the price increases as marginal cost of production increases.

Appendix C1. An Order Issued by Kunshan Government to Improve Water Quality around the Monitoring Stations (Scan Copy)

昆山市两减六治三提升专项行动领导小组办公室

昆 263 办〔2017〕186 号

关于对吴淞江赵屯（石浦）等 3 个断面所属流域工业企业实施全面停产的紧急通知

昆山开发区、昆山高新区、花桥经济开发区管委会，各镇人民政府，市水利局、环保局，水务集团：

近期我市自动监测数据显示，我市国省考断面水质较差，达标形势严峻，尤其是赵屯断面（劣V类）、振东渡口断面（劣 V 类）、千灯浦口断面（V类）均无法达到国省考要求。

为确保我市国省考断面达到国家下达 2018 年度考核目标要求，决定对吴淞江赵屯（石浦）等 3 个断面所属流域工业企业（企业名单附后）自 2017 年 12 月 25 日起至 2018 年 1 月 10 日期间实施全面停产，到期视水质情况，决定是否延期；请相关区镇通知相关企

—1—

措施是否到位；对吴淞江、太仓塘、千灯浦沿线闸门全部关闭，泵站禁止排水；实施停产时、停产期间和复工复产时，请各区镇和相关部门督促企业落实好安全生产各项措施。

停产期间实施日报告制度，市水利局每日对站闸关闭情况、泵站排水情况进行检查，每日下午 4:30 前，将检查结果报市 263 办公室邮箱；各相关区镇每日对相关企业停产、网格员驻厂情况、所辖污水厂排放情况进行检查，每日下午 4:30 前将检查结果报市 263 办公室邮箱。（ks263bgs@163.com）

特此通知，请遵照执行。

附件：吴淞江赵屯（石浦）等 3 个断面所属流域停产企业名单

昆山市“两减六治三提升”专项行动领导小组办公室

2017 年 12 月 24 日

抄报：市委办、市府办。

抄送：市安监局、消防大队。

—2—

Appendix C2. An Order Issued by Kunshan Government to Improve Water Quality around the Monitoring Stations (Authors' Translation)

Kunshan 263 Special Action Team Office

Kunshan 263 Office【2017】#186

An Urgent Order on Suspending Production of Industrial Enterprises Located in the River Basin of Wusong River Zhaotun (Shipu) Water Quality Monitoring Station and Two Other River Basins

To Kunshan Development Zone, Kunshan High-tech Zone, Huaqiao Economic Development Zone Administrative Board, People's Governments of Townships, Municipal Water Resources Bureau, Municipal Environmental Protection Bureau and Municipal Water Affairs Group:

According to the recent data from the automatic water quality monitoring stations in Kunshan city, the water quality of several river segments used for national and provincial assessment is relatively poor. The situation is particularly severe for the Zhaotun water quality monitoring station (worse than Grade V), Zhengdong ferry station (worse than Grade V), and Qian Deng Pu Kou station (Grade V), all of which may fail to meet the national and provincial assessment requirements.

To ensure that the water quality in these river segments meets the annual national assessment requirement in 2018, we have decided to suspend the production of industrial enterprises (list attached) located near the Wusong River Basin Zhaotun (Shipu) water quality monitoring station and two other river basins near water quality monitoring stations, effective from December 25, 2017 to January 10, 2018. The suspension of production may be further extended, depending on the conditions of water quality readings. Relevant district and township governments should inform the enterprises about the decision. The inspection teams should supervise and take production cessation measures. Special investigators should be placed in the plants to ensure full compliance. Sluice gates along Wusong River, Taicang Embankment and Qian Deng Pu must be closed, and the pumping facilities need to be shut down and stop discharging wastewater during this period. District and township governments and relevant departments shall ensure that enterprises take proper safety measures in the process of suspending and resuming production.

During the production suspension period, a daily reporting system will be adopted. The Municipal Water Resources Bureau shall inspect the status of all sluice gates and pumping facilities and report the inspection results to the City's "263" Office before 4:30pm every day. District and township governments, special investigators based in the plants, and wastewater treatment plants shall check whether there are violations of the production suspension order and report the results to the Office (ks263bgs@163.com) before 4:30 PM every day.

Hereby noticed, and please follow the order.

Appendix: List of industrial enterprises that shall suspend production

The "263" Office of Kunshan Government

24th December, 2017

cc. Municipal Office of Kunshan, Government Office of Kunshan

cc. Kunshan Safety Supervision Bureau, Kunshan City Fire Brigade

Table S1. 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	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 Manufac	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 Manufactu	34
		General Purpose Machinery Manufact	35
		Special purpose Machinery Manufactu	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

Notes: Industrial classification for national economic activities (GB/T 4754—2002). The division between polluting Industries and non-polluting Industries is according to the Ministry of Environmental Protection (http://wfs.mep.gov.cn/gywrfz/hbhc/zcfg/201009/t20100914_194483.htm).

Table S2. Covariate Balance Between Upstream Townships and Downstream Townships

	Mean		Mean Difference
	Downstream	Upstream	≤10km
	(1)	(2)	(3)
Panel A. Basic Township Characteristics			
Town Area	7,636	7,345	302.02
(Mu)	-4,210	-4,233	(1,467.30)
Arable Area	3,308	2,734	-58.29
(Mu)	-2,005	-1,880	(803.44)
Distance to County Center	2.34	2.45	0.52
(KM)	(0.94)	(1.04)	(0.70)
Old-Region Town	0.21	0.16	-0.14
(1=Old-Region Town)	(0.41)	(0.36)	(0.18)
Minority Town	0.01	0.02	-0.00*
(1=Minority Town)	(0.12)	(0.15)	(0.00)
No. of Residents Communities	1.65	1.29	-2.52
	(4.34)	(2.72)	(4.86)
No. of Villages	25.94	22.03	-1.62
	(16.77)	(14.57)	(5.33)
Panel B. Basic Infrastructure			
Road Length	53.55	44.09	-3.83
(KM)	(46.37)	(43.40)	(11.57)
# of Villages with Paved Road	24.73	21.27	-0.57
	(16.51)	(13.92)	(5.51)
# of Villages with Electricity	25.94	22.03	-1.62
	(16.77)	(14.57)	(5.33)
# of Villages with Tap Water	14.24	10.07	0.69
	(15.65)	(13.60)	(6.02)
Panel C. Human Capital			
No. of Primary School	18.08	17.02	2.10
	(9.38)	(9.45)	(3.99)
No. of Primary School Students	7,271	5,936	-309.76
	-4,824	-3,790	(2,329.78)
Obs.	187	143	104

Notes: Data are collected from the Township Conditions Survey in 2002. Columns 1–2 report the means and standard deviations of township covariates. In columns 3, we test the covariate balance between upstream and downstream towns. The difference coefficients are obtained by running OLS regressions of township variables on an upstream dummy and a set of water quality monitoring station fixed effects. Standard errors reported in the parentheses are clustered at the water monitoring station level.

Table S3. Density Tests for Sorting Using Local Polynomial Density Estimation

	(1)	(2)	(3)	(4)
Panel A. Firms 2000-2007, Obs = 6582				
T	0.09	0.73	-12.07	0.09
P>|T|	0.93	0.47	0.00	0.93
Bandwidth Left	2.47	1.97	6.17	2.47
Bandwidth Right	2.01	1.97	6.17	2.01
Panel B. Firms in 2013, Obs = 2562				
T	1.60	1.44	1.26	1.60
P>|T|	0.11	0.15	0.21	0.11
Bandwidth Left	2.93	2.70	3.60	2.93
Bandwidth Right	3.29	2.70	3.60	3.29
Bandwidth Selector	Each	Diff	Sum	Comb

Notes: This table reports RD manipulating tests using the local polynomial density estimators proposed by Cattaneo et al. (2017a, 2007b). We use four different bandwidth selectors to check the robustness of the results. "Each" means we use two distinct bandwidths based on MSE of each density separately for upstream and downstream firms, which is our preferred test. "Diff" bandwidth selection is based on MSE of difference of densities with one common bandwidth. "Sum" bandwidth selection is based on MSE of sum of densities with one common bandwidth. "Sum" selects the largest bandwidth in our tests and creates large noises in the testing results. "Comb" bandwidth is selected as the median of "Each", "Diff" and "Sum". Technical explanations of different bandwidth selectors can be found in Cattaneo et al. (2017a, 2007b).

Table S4. Balanced Panel Results

	Polluting Industries			Non-Polluting Industries		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Water Quality Monitoring and TFP						
TFP (log)	0.24	0.28	0.47**	-0.20	-0.22	-0.14
	(0.30)	(0.30)	(0.24)	(0.23)	(0.23)	(0.23)
Bandwidth (km)	4.89	4.55	5.23	3.64	3.29	3.02
Panel B: Water Quality Monitoring and Residual TFP						
TFP (log)	0.38*	0.41**	0.63***	0.05	0.05	0.01
(Station FE Absorbed)	(0.22)	(0.21)	(0.22)	(0.22)	(0.22)	(0.20)
Bandwidth (km)	6.41	6.37	5.63	3.45	3.17	3.42
Panel C: Water Quality Monitoring and Residual TFP						
TFP (log)	0.40*	0.46**	0.32*	0.13	0.12	0.12
(Station and Industry FE Absorbed)	(0.21)	(0.21)	(0.18)	(0.18)	(0.19)	(0.19)
Bandwidth (km)	6.24	5.99	6.12	3.69	3.59	2.99
Obs.	1281	1281	1281	2896	2896	2896
Kernel	Triangle	Epanech.	Uniform	Triangle	Epanech.	Uniform

Notes: This table replicates the baseline results in Table 2 with only firms in the balanced panel. Each cell in the table represents a separate regression. TFP is estimated using Olley and Pakes (1996) method. The discontinuities at monitoring stations are estimated using local linear regressions and MSE-optimal bandwidth proposed by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods. Standard errors clustered at the monitoring station level are reported below the estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table S5. Abatement Facilities

	(1)	(2)	(3)
Panel A: Water Quality Monitoring and Abatement Facilities			
Abatement Facilities	-0.83**	-0.82**	-0.98**
	(0.32)	(0.33)	(0.39)
Bandwidth (km)	3.28	3.29	2.97
Panel B: Residual Abatement Facilities			
Abatement Facilities	-1.09*	-1.03*	-1.26*
(Station FE Absorbed)	(0.60)	(0.58)	(0.69)
Bandwidth (km)	6.09	6.16	4.23
Panel C: Residual Abatement Facilities			
Abatement Facilities	-0.98*	-0.92	-1.17*
(Station and Industry FE Absorbed)	(0.59)	(0.57)	(0.66)
Bandwidth (km)	6.16	6.23	4.40
Obs.	9041	9041	9041
Kernel	Triangle	Epanech.	Uniform

Note: Each cell in the table represents a separate regression. The outcome variable is the amount of abatement facilities owned by the firm. The discontinuities at monitoring stations are estimated using local linear regressions and MSE-optimal bandwidth proposed by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods. Standard errors clustered at the monitoring station level are reported below the estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table S6. Political Incentive Results with Politician FE

VARIABLES	(1) Log TFP	(2) Log TFP
Downstream	0.35***	0.12
	(0.05)	(0.08)
Distance	-0.01**	-0.01**
	(0.00)	(0.00)
Downstream*Distance	-0.00	-0.01
	(0.01)	(0.01)
Incentive		-0.03
		(0.07)
Downstream*Incentive		0.31**
		(0.11)
Mean of Outcome Variable	2.58	2.58
Station FE	Yes	Yes
Industry FE	Yes	Yes
Leader FE	Yes	Yes
Year FE	Yes	Yes
Observations	19,572	19,572
R-squared	0.361	0.364

Note: This table reports the robustness of the "promotion incentive" results to the inclusion of politician fixed effects. Column 1 presents the baseline coefficients under this specification, column 2 adds the political incentive dummy and its interaction with the downstream dummy. Standard errors clustered at the monitoring station level are reported below the estimates. * significant at 10% ** significant at 5% *** significant at 1%.

Table S7. Polynomial RD Estimates of the Impact of Water Quality Monitoring on TFP

	Polluting Industries					Non-Polluting Industries				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Control for Station FE										
TFP (log)	0.30**	0.33**	0.22	0.38***	0.65***	0.03	0.07	0.02	0.15*	0.19
	(0.12)	(0.16)	(0.21)	(0.12)	(0.19)	(0.07)	(0.10)	(0.15)	(0.09)	(0.14)
R-Square	0.14	0.14	0.14	0.17	0.16	0.08	0.08	0.08	0.09	0.10
Panel B: Control for Station and Industry FE										
TFP (log)	0.22**	0.29*	0.30	0.32***	0.47**	0.00	0.10	0.17	0.13*	0.16
	(0.09)	(0.14)	(0.20)	(0.10)	(0.19)	(0.07)	(0.09)	(0.13)	(0.07)	(0.14)
R-Square	0.26	0.26	0.26	0.28	0.29	0.27	0.27	0.27	0.29	0.27
Obs.	6,582	6,582	6,582	4,462	1,474	12,422	12,422	12,422	8,981	3,260
Polynomial Function	Linear	Quadratic	Cubic	Linear	Linear	Linear	Quadratic	Cubic	Linear	Linear
Sample	20km	20km	20km	10km	5km	20km	20km	20km	10km	5km

Notes: Each cell in the table represents a separate regression. TFP is estimated using Olley and Pakes (1996) method. We report OLS estimates of the coefficient on a "downstream" dummy after controlling for polynomial functions in distance from the water quality monitoring stations interacted with a downstream dummy. Standard errors clustered at the monitoring station level are reported below the coefficients. * significant at 10% ** significant at 5% *** significant at 1%.

Table S8. RD Estimates of the Impact of Water Quality Monitoring on Alternative TFP

	Polluting Industries			Non-Polluting Industries		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Water Quality Monitoring and TFP						
TFP (log) - Polluting Industries	0.61*	0.67*	0.83**	-0.06	-0.10	-0.05
	(0.35)	(0.37)	(0.39)	(0.21)	(0.22)	(0.25)
Panel B: Water Quality Monitoring and Residual TFP						
TFP (log) - Polluting Industries	0.55**	0.54**	0.69**	0.07	0.06	0.00
(Station FE Absorbed)	(0.27)	(0.27)	(0.31)	(0.13)	(0.13)	(0.14)
Panel C: Water Quality Monitoring and Residual TFP						
TFP (log) - Polluting Industries	0.31	0.35	0.57*	0.15	0.14	0.08
(Station and Industry FE Absorbed)	(0.23)	(0.24)	(0.30)	(0.12)	(0.13)	(0.12)
Obs.	6,039	6,039	6,039	11,440	11,440	11,440
Kernel	Triangle	Epanech.	Uniform	Triangle	Epanech.	Uniform

Notes: Each cell in the table represents a separate regression. TFP is estimated using method proposed by Ackerberg et al. (2015). The discontinuities at monitoring stations are estimated using local linear regression and MSE-optimal bandwidth proposed by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods. * significant at 10% ** significant at 5% *** significant at 1%.

Table S9. Effect of Monitoring on TFP and Emission by Size

	Small Firms/Emitters			Large Firms/Emitters		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: RD Estimates for TFP (log)						
Downstream	0.11	0.11	0.24	0.31**	0.32**	0.30
	(0.17)	(0.18)	(0.17)	(0.14)	(0.16)	(0.19)
Obs.	3,038	3,038	3,038	3,538	3,538	3,538
Panel B: RD estimates for COD Emission Intensity (log)						
Downstream	0.49	0.44	0.25	0.92*	0.80*	0.48
	(0.40)	(0.36)	(0.28)	(0.47)	(0.45)	(0.44)
Obs.	4,901	4,901	4,901	4,906	4,906	4,906
Kernel	Triangle	Epanech.	Uniform	Triangle	Epanech.	Uniform

Notes: Each cell in the table represents a separate regression. Monitoring station and industry fixed effects are absorbed before estimating regression discontinuity. In columns 1–3, we report the estimated discontinuity for smaller firms or emitters, and in columns 4–6, we report the estimated discontinuity for large firms or emitters. Local linear regression and MSE-optimal bandwidth proposed by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods are used for the estimation. * significant at 10% ** significant at 5% *** significant at 1%.

Table S10. Effect of Water Quality Monitoring on Firm Exit

	Exit – Polluting Industries			Exit – Non-Polluting Industries		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Conventional RD Estimates						
Downstream	0.07	0.07	0.11	-0.02	-0.02	-0.04
	(0.09)	(0.09)	(0.10)	(0.05)	(0.05)	(0.05)
Panel B: Bias-Corrected Robust RD Estimates						
Downstream	0.10	0.10	0.15	-0.01	-0.02	-0.04
	(0.10)	(0.10)	(0.11)	(0.07)	(0.07)	(0.06)
Obs.	6,581	6,581	6,581	12,422	12,422	12,422
Kernel	Triangle	Epanech.	Uniform	Triangle	Epanech.	Uniform

Notes: Each cell in the table represents a separate regression. Monitoring station and industry fixed effects are absorbed before estimating regression discontinuity. In columns 1–3, we report the estimated discontinuity for polluting industries, and in columns 4–6, we report the estimated discontinuity for non-polluting industries. Local linear regression and MSE-optimal bandwidth proposed by Calonico et al. (2014) and Calonico (2017) for different kernel weighting methods are used for the estimation. Conventional local linear regression discontinuity standard errors clustered at the monitoring station level are reported below the estimates in Panel A; bias-corrected RD estimates and robust standard errors are reported in Panel B. * significant at 10% ** significant at 5% *** significant at 1%.

APPENDIX REFERENCES

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. Identification Properties of Recent Production Function Estimators. *Econometrica*, *83*(6), pp.2411-2451.

Brandt, L., Van Biesebroeck, J. and Zhang, Y. 2012. "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing." *Journal of Development Economics*, *97*(2), pp.339-351.

Calonico, Sebastian, Matias D. Cattaneo, and Max H. Farrell. Forthcoming. "On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference." *Journal of the American Statistical Association*.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression - Discontinuity Designs." *Econometrica* 82 (6): 2295-326.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6): 1263-1297.

Yang, Rudai. 2015. "Study on the Total Factor Productivity of Chinese Manufacturing Enterprises." *Economic Research Journal* (Chinese), *2*, p.006.