Environmental Regulation and Firm Productivity in China: Estimates from a Regression Discontinuity Design

BY GUOJUN HE, SHAODA WANG AND BING ZHANG

(This Version: Dec. 2017)

This paper estimates the causal effect of environmental regulation on firm productivity in a regression discontinuity design implicit in China's surface water quality monitoring system. Since water quality readings are important for political evaluation, and such readings can only capture emissions from the upstream, local governments are incentivized to enforce tighter regulations on firms in the immediate upstream of a monitoring station, rather than those in the immediate downstream, despite their spatial adjacency. Exploiting the discontinuity in the regulation stringency across the monitoring stations, we find that upstream polluting firms' TFP are 27% lower than downstream firms, which corresponds to a 56% difference in COD emissions. These estimates imply that China's COD abatement target between 2016 and 2020 would lead to a roughly one trillion-Yuan loss in total industrial output.

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

JEL: Q56, Q58, O13, O44, D24

* He, Division of Social Science, Division of Environment and Sustainability, and Department of Economics, the Hong Kong University of Science and Technology; Email: gjhe@me.com; Wang, Department of Agricultural and Resource Economics, University of California, Berkeley; Email: shaoda@berkeley.edu; Zhang, Center for Environmental Management and Policy Analysis, School of Environment, Nanjing University; Email: zhangb@nju.edu.cn.

I. Introduction

The question of whether environmental regulation hinders economic efficiency has long been important and controversial. Today, especially in many rapidly-growing economies, this debate attracts much attention and entails significant policy ramifications. On the one hand, neoclassical models suggest that environmental regulations will increase production cost, create unemployment, and thus reduce the competitiveness of an economy. On the other hand, environmentalists and other proponents of environmental preservation argue that more stringent regulations can provide incentives for polluters to develop cleaner and less costly technologies to reduce pollution, which can in turn be beneficial to productivity.[1]

In this study, we estimate the causal effect of environmental regulation on firm productivity in a novel spatial regression discontinuity setting. We exploit China's surface water quality monitoring system and investigate how tighter water emission controls affect the total factor productivity (TFP) of Chinese manufacturing firms. We argue that, because water quality monitors can only pick up pollution information from its upstream regions, local governments have strong incentives to require upstream firms to abate emissions in order to improve water quality measured at the monitoring station. As a result, within a small neighborhood around a water quality monitoring station, upstream firms face tighter environmental regulations than downstream firms. By focusing on a narrow geographic band that only stretches from a few townships upstream and downstream to each surface water monitoring station, we are able to isolate the impacts of water quality controls on industrial firms' productivity from potential confounding factors. Our analysis shows that upstream firms in polluting industries have significantly lower TFP as compared to downstream firms.

Our identification relies on the assumption that upstream and downstream firms should be *ex ante* identical in the absence of any pollution controls. This identifying assumption likely holds because the locations of surface water quality monitoring stations were mainly determined by hydrological factors (such as water flow and river width) rather than socio-economic factors. Particularly, the Chinese government explicitly requires that water quality monitoring stations should be established close to existing hydrological stations so that they can share certain facilities to realize economy of scale, and combine water quality and

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

hydrological data. Since the hydrological stations were established to monitor hydrological conditions rather than industrial conditions, and most of them were established between the 1950s and 1970s, during which China had relatively little industrial pollution, the location choices of hydrological monitoring stations should be orthogonal to firms' economic or environmental performances today.

In addition to the qualitative arguments, we document four primary patterns in the data to support the validity of our research design. First, we show that time-invariant socio-economic conditions and basic infrastructure measures are well balanced between upstream and downstream townships. Second, we find that only firms in the polluting industries are affected by water quality monitoring, while firms in the non-polluting industries are unaffected. Third, we analyze the data by year and show that the spatial discontinuity in firms' TFP in the polluting industries only became evident after 2003, when the new political regime of President Hu Jintao started to emphasize achieving a balance between economic growth and sustainability. Finally, exploiting the fact that many monitoring stations are intentionally located adjacent to hydrological stations, we use whether a firm is in the adjacent upstream of a hydrological station as an instrumental variable for whether a firm is in the adjacent upstream of a monitoring station, and find similar results.

We further investigate the impacts of water quality monitoring on TFP using subsamples of private Chinese firms, state-owned firms and foreign-owned firms and find that private firms suffer most from tighter water pollution controls. We conduct similar analyses separately for old and young firms, and find that water quality monitoring affects both old and young firms.

Using a different firm-level dataset that records polluting firms' environmental performances, we show that water quality monitoring indeed leads to less pollution. Both chemical oxygen demand (COD) emissions and COD emission intensity (emissions per unit of output) are higher in downstream regions, suggesting that the total reduction in emissions comes not only from the reduction in output but also from the adoption of cleaner technologies. We also find similar patterns for waste water discharge and waste water discharge intensity (emissions per unit of output).

This study contributes to the ongoing debate on the economic costs of environmental regulations in several important ways. First, although the topic is of tremendous policy relevance in developing countries, 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 study, we investigate China, the largest developing country, manufacturer and emitter in the world, and highlight the potential significant economic costs of environmental regulation in the context of a rapidly-growing economy.

Second, in addition to adopting a credible identification strategy, another appealing feature of our research design is that we can isolate the TFP effect of regulation from the price effect of regulation. Like most studies using firm-level production data, we use revenue-based TFP measures because firm-level production data often only contains total revenue and lacks price information. As a result, changes in revenue-based TFP measures often capture a combined effect of changes in price and changes in total output.[2] In this study, our upstream and downstream firms are located so close to each other that they essentially serve the same market. Meanwhile, these firms jointly only produce a small proportion of total national industrial output, so local water quality monitoring is unlikely to affect aggregated output market prices. This implies that the TFP effects estimated in this study, given that output price remains the same for upstream and downstream firms, are not plagued by the price effect.

Third, our findings on the heterogeneous effects shed light on how different types of firms in China respond to environmental regulations. Particularly, 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. We also find that the TFP loss is substantially larger for older firms than for younger firms, and younger firms are more likely to be borne in downstream regions. In the long run, these findings imply a redistribution of production, income, environmental quality and social welfare between upstream and downstream regions. Our findings therefore also speak 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 FDI (Fredriksson, List and Millimet 2003; Hanna, 2010; Cai et al. 2016).

Fourth, we also reveal the channels and political economy behind China's environmental regulations. We find that upstream firms need to pay more emission fees and taxes than downstream firms do, even though they do not produce more emissions or outputs. Upstream

[2] Please refer to Greenstone, List, and Syverson (2012) for a discussion on how price mismeasurement may bias the estimates.

firms also tend to hire more labor and make more capital invest, possibly to abate emissions. We find no evidence that tighter regulation encourages research and development investment. These patterns seem consistent with the neoclassical explanations instead of the Porter hypothesis (Porter, 1991). On the government's side, we show that political incentives matter. When city leader has higher promotion probability and thus strong political incentives, the impacts of water quality monitoring are twice larger. Meanwhile, for monitoring stations that are less susceptible to local political influence, the impacts are more salient.

Finally, understanding the economic costs of environmental regulations is critical for optimal policy design. Applying the estimates to China's national abatement plans, we can calculate the economic costs of tightening water pollution regulations. We estimate that a 10% reduction in COD emissions leads to a 2.49% decrease in TFP, and that 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). These estimates have important policy implications and can help the Chinese government irmpove environmental policy designs.

The remainder of this paper is structured as follows. Section II describes the institutional background, research design and empirical strategy. Section III describes the data and presents descriptive statistics. Section IV presents the estimation results and discuss the findings. Section V examines the channels 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 was polluted and contained water deemed unsafe for human consumption. Poor surface water quality has driven policymakers to propose regulations to protect the water bodies and reverse the process of degradation. A national

water quality monitoring system was established to monitor surface water quality in major river segments, lakes and reservoirs.

The Ministry of Environmental Protection (MEP) started to monitor surface water quality in the 1990s. Initially known as the Bureau of Environmental Protection, the MEP issued the "National Environmental Quality Monitoring Network-Surface Water Monitoring System" (NEQMN-SWMS) in 1993, announcing the establishment of a national surface water quality monitoring system. However, the established monitoring system was mainly intended for scientific rather than regulatory purposes at the initial stage, and most of the station-level monitoring data collected were kept confidential by the government. No strict emission abatement targets were set by the Chinese government at the time as economic growth was considered the country's priority. 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 over from Jiang Zemin, and held office till 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), which sought integrated sets of solutions to economic, environmental and social problems in China, opening an era of environmental regulation.[3]

Responding to the SOD slogan, the MEP increased its efforts to resolve the issue of water pollution. In 2003, it issued an updated version of NEQMN-SWMS and the "Technical Specification Requirements for Monitoring of Surface Water and Waste Water". These new policy documents signaled an expansion of the national surface water quality monitoring system and made water pollution control an important political task in the governance system. During this period, the total number of state-controlled surface water quality monitoring stations in China's seven major river systems increased from 419 to 574,[4] and a significant amount of resources were allocated to controlling water pollution. Various environmental yearbooks started to publish station-level water quality readings in 2003.

---

[3] SOD is sometimes translated as either the "Scientific Development Concept," or the "Scientific Development Perspective."

[4] The most recent expansion of the system, under the administration of President Xi Jinping, further increased the number of state-controlled monitoring stations to 972 (with 766 for major rivers, and 206 for lakes and reservoirs) in 2015.

During President Hu's political regime, the central government adopted a target-based abatement system is 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 Standard 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 then assigned abatement mandates to prefectures and counties and used local environmental performance along with other criteria to assess and promote local government officials.

Since rivers flow from areas of higher elevation towards areas of lower elevation, water quality monitoring stations can only capture emissions from their upstream areas, but not 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. Since 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 before 2003, and become stronger afterwards.

B. Location Choice of Water Quality Monitoring Stations

The primary target of the national surface 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, important lakes, and reservoirs. A water quality monitoring station is required to be spatially representative of its neighborhood water body and can properly reflect changes in water environmental pollutants over time. As a result, the

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

locations of the monitoring stations were chosen mainly based on scientific or hydrological considerations.

According to the MEP, the monitoring stations must be placed on 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 explicitly demands that monitoring stations should be established to serve for "long-term" purposes, ensuring that short-term needs (such as avoiding pollution from a specific region or a specific firm) cannot be accommodated.

Another important requirement of station placement is that they should be built close to hydrological stations whenever possible to enable 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 were established and supervised by the MEP and the State Council of China. The water quality readings from the state-controlled stations are directly reported to the MEP to ensure data quality. The yearly average water quality readings from the state-controlled stations are reported in various environmental yearbooks and used 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 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 to the monitoring station serves as the running variable and we examine whether firms located immediately upstream from the monitoring station have lower productivity than adjacent downstream firms. The empirical strategy in this study is similar in spirit to recent work that also uses the flow of pollution along rivers as a source of identifying variation (Kaiser and Shapiro, 2017; Lipscomb and Mobarak, 2017).

The identifying assumption of our research design is that due to spatial adjacency, firms located immediately upstream and downstream to monitoring stations should be balanced *ex*

ante along various dimensions, but will differ from each other when 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 polynomial and has some other undesirable statistical properties. As a result, estimators based on local linear regression or other smooth functions are often preferred, as they can assign larger weights on observations closer to the threshold and produce more accurate estimates. We thus focus on the non-parametric RD approach, which can be estimated by the following equation:

$$TFP_i = \alpha_1 Down_i + \alpha_2 Dist_i + \alpha_3 Down_{ijk} Dist_i + \varepsilon_i \tag{1}$$

$$s.t. \ -h \leq Dist_i \leq h$$

where TFP_i is the total factor productivity of firm i around a monitoring station. $Down_i$ is an indicator variable that equals 1 if firm i is downstream from a monitoring station, and 0 otherwise. $Dist_i$ measures the distance between firm i and the monitoring station, and h is the bandwidth length (i.e., the acceptable distance from the discontinuity for sample inclusion). The choice of h involves balancing the conflicting goals of focusing comparisons near the monitoring stations where the identification assumption is strongest and providing a large enough sample for reliable estimation. In this study, we rely on a MSE-optimal bandwidth h proposed by Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, Farrell (forthcoming) and experiment with different kernel weighting functions.

To account for location-specific and industry-specific TFP determinants in the non-parametric estimations, we first absorb station fixed effects and industry fixed effects by running an OLS regression of TFP on a set of station-specific and industry-specific dummies, and then apply the non-parametric estimations on the residual TFP obtained from OLS estimation. This approach is suggested by Lee and Lemieux (2010), who 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.

As a way to check the robustness, we also estimate the RD estimates using the parametric approach:

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

where TFP_{ijk} is the total factor productivity of firm i around a monitoring station j in industry k. $f(Dist_{ijk})$ is a polynomial in distance between firm i and monitoring station j in industry 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 δ_k are station and industry fixed effects.

III. Data and Summary Statistics

A. Data

Our analysis is based on several data files that together provide comprehensive information on the socioeconomic conditions of townships, the production and performance of 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 input, output, 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) as the main outcome variable. The

Olley-Pakes method addresses the simultaneity and selection biases in estimating TFP and is widely used in empirical research. The details of estimating TFP using the Olley-Pakes method are discussed in Appendix A. For robustness checks, we also estimate a recently-developed TFP measure that is proposed by Ackerberg et al. (2015).

The ASIF data were used in several previous studies and a well-known issue that the data contain outliers. We follow standard procedures documented in the literature to clean the data.[6] 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; current depreciation larger than accumulative depreciation. Finally, we trim the data by dropping observations with values of key variables outside the range of 0.5^{th} to 99.5^{th} percentile.

The ASIF data have detailed address data 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.

Since our research design is fundamentally cross-sectional, despite having multiple observations over time for some firms, we collapse the data into cross-sectional data and apply the RD estimators to them. The interpretation of the coefficients is therefore an average effect that persists for years. To fully utilize the panel 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 run 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

[6] 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).

plants and urban sewage treatment plants. In this study, we only keep the ESR firms that are in the same polluting industries as the ASIF firms.

The sampling criteria in the ESR is based on the cumulative distribution of emissions in each county. Polluting sources are ranked based on their emission levels 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 for the ASIF database.

During our sample period, the "criteria pollutants" changed overtime. 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.[7] 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 waste water 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.

Township-level Socioeconomic 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-

[7] 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 target and conduct performance appraisal.

economic data for all the townships in China. We have obtained 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 have 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 matching process involves several steps. 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.

Our research design is illustrated in Figure 1. We first match water quality monitoring stations with China's water basin system. In some regions, the distribution of monitoring stations is very dense and multiple tributaries or branch rivers merge into the trunk streams, making it difficult to identify upstream and downstream relationships (for example, an adjacent upstream township for a monitoring station can be in the adjacent downstream of another monitoring station). We therefore exclude these water monitoring stations from our dataset. About a quarter of the monitoring stations are located on lakes or reservoirs, and we drop them as well. 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 water quality station identified, we then determine in which township it is located and then draw a circle with a radius of 10 km from the town center. Based on elevation data and river flow directions, we then identify the upstream and downstream townships in which the river passes by.[8] This process generates our sampled regions, which include 544 townships.

[8] Because some townships only barely intersect with this circle, the firms located in these townships can be more than 10 km away from the monitoring station. We keep these firms in the baseline 10 km sample and also show that our results are robust to using much smaller bands (5km).

The final step is to overlay the coordinates of all ASIF or ESR firms on the map of identified townships and keep only the firms within these townships and calculate their distances to the monitoring stations.

In the end, we are able to assemble a geo-coded data set that includes township-level socioeconomic conditions, firm-level production and performance, and firm-level emissions. Our sample includes 12,422 unique ASIF firms and 9,888 ESR firms for 161 water quality monitoring stations.

We attempted to match firms in the ASIF database with firms in the ESR database. However, because these two datasets use very different sampling criteria and are managed by different government agencies, we were only able to match 10% of ASIF firms with ESR firms. The matched sample is too small for us to draw any credible statistical inference. As a result, in subsequent analysis, we analyze these two datasets separately.

C. Balance Check

The underlying assumption for our RD design is that except for environmental regulation, other determinants of TFP change smoothly around the monitoring stations. At the firm level, environmental regulations may affect many production decisions. It is thus difficult to test this assumption using firm-level data, which primarily contain time-varying variables (arguably, even firm type and firm age can be affected by environmental regulations). In contrast, our township level data include a rich set of variables that are important for firm production and can be informative about whether these firms face the same market environment.

Since our sample only includes townships located around monitoring stations, we expect that the township-level statistics are largely balanced. In Table 1, we summarize comparisons between upstream and downstream townships. We examine three sets of covariates. Panel A reports the results for basic township characteristics, Panel B summarizes comparisons of local infrastructure, and Panel C further compares human capital measures. We present the means of these variables separately for upstream and downstream townships and then test the mean differences using different bandwidths.

Basic township characteristics in Panel A include township area, arable area, distance to county center, whether the township is an old-region town, whether it is a minority town, the number of resident's communities, and the number of administrative villages. We cannot

reject the null hypothesis that upstream townships and downstream townships are balanced in every measure at the conventional significance levels.

Infrastructure is important for production. In Panel B, we test whether basic infrastructure measures are similar between upstream and downstream townships. We have data on the length of roads, number of villages with road access, number of villages with electricity access, and number of tap water access. Again, we find that upstream and downstream townships are similar along these dimensions.

Finally, production requires labor and we examine whether human capital differs significantly between upstream and downstream townships. 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 are encouraging, as they indicate that upstream and downstream townships are very similar. While it is, of course, impossible to rule out the presence of unobserved factors affecting firm's productivity discontinuously, these balance checks lend additional credibility to our research design.

IV. Results

A. Effects of Water Quality Monitoring on TFP

We begin the analysis by graphically presenting our main findings. Applying the Olley-Pakes method, we estimate the log TFP for each sampled firm. Figure 3 plots log TFP (or residual log TFP) against distance to monitoring stations. Each dot represents the average log TFP for firms within a bin of distance and their 95% confidence intervals are also presented. A quartic polynomial function is then overlaid on the graph to illustrate the discontinuity at the monitoring stations.

We divide the firms in ASIF into two categories: the polluting industries and the non-polluting industries. This categorization is based on the definition of polluting industries used by the MEP.[9]

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

In Panel A, we present the RD plot for log TFP in the polluting industries; and in Panel B, we show the RD plot for residual log TFP in the polluting industries. The difference is that in the residual log TFP panel, monitor station fixed effects are absorbed. In both panels, we see a sharp change in TFP at precisely the locations where the water quality controls take effect. The TFP for upstream firms is significantly lower than for downstream firms in the polluting industries. In contrast, in Panels C and D, we do not observe similar discontinuities in TFP in the 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 differences for the non-polluting firms. The estimates are not statistically significant because of larger 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 noises for the statistical inference. To address this issue, we first absorb station fixed effects and industry fixed effects in the regressions and then estimate the RD using the residual TFP in Panels B and 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. We see that after controlling for station and industry fixed effects, the RD estimates for the polluting industries become statistically significant.

In our preferred specifications in which both station and industry fixed effects are accounted for, the estimated increase in log TFP for downstream monitoring stations ranges from 0.31 to 0.35 for the polluting industries. The 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).

The estimates for the non-polluting industries are close to zero and none of them are statistically significant. Comparing the RD estimates in Panels B and C to Panel A, we see that the magnitudes of the estimated impacts are remarkably close. This is important because it suggests that station and industry specific determinants of TFP levels are uncorrelated with the treatment status. However, since location and industry-specific factors can substantially explain the variation in TFP, including them can significantly reduce the estimated standard errors of the treatment effects. The RD estimates are also robust to different choices of kernel functions.

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

In Table 3, we explore whether the effect of water quality monitoring on TFP varies by ownership, firm age and year. 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 mainly driven 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, which are believed to have greater bargaining power over local governments. Another possible explanation is that SOEs and foreign firms generally have superior *ex-ante* environmental performance compared to private Chinese firms and are therefore 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.

The stringency of water quality regulations has changed substantially over the course of 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. Starting from 2006, COD abatement further became a key indicator in evaluating local environmental performance.

We thus hypothesize that the TFP effect of water monitoring should be larger in later years than in earlier years, and this change should occur in or after 2003. To test this, we estimate the RD separately for samples before 2003 and after 2003, and summarize the results in Panel B. As expected, upstream and downstream polluting firms had similar TFP levels before 2003, but upstream firms became significantly less productive than their downstream counterparts after 2003.

In Figure 4, we further provide RD estimates separately for each year. We find that the TFP differences between upstream and downstream firms match exactly to the policy changes we have discussed. Particularly, the estimate is close to zero in 2000 to 2002, and becomes larger in 2003, the first year President Hu resumed his duty. The effect becomes statistically significant from year 2006, the starting year of the 11th Five Year Plan. The corresponding by-year RD estimates are summarized in Appendix Table S2.

The monitoring effect was close to zero and statistically insignificant prior to 2003 is important because it supports 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.

In Panel C, we compare the TFP loss by firm age. We are interested in whether old firms and young firms respond to water quality monitoring differently. We define new firms as firms born in or after 2003, starting from which 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 are statistically significant for both old and young firms. This finding is not consistent with the "grandfathering" phenomenon which describes that 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 generally higher than that of building new sources with cleaner technology. It turns out in the context of China's water quality monitoring, all upstream firms are under tighter regulations.

C. IV Results Using Hydrological Stations

Our qualitative discussions on the rules of setting up monitoring stations, the balance tests of 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 in its upstream, rather than downstream. 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 to allow data, equipment and technicians to be shared, so that there could be 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. Since hydrological stations were set up between the 1950s and 1970s, and their locations were chosen purely based on hydrological considerations, these locations should be orthogonal to the future socioeconomic conditions of their neighborhoods. All the hydrological stations were built and supervised by the Ministry of Water Resources (MWR).

As a result, whether a firm is in the near upstream area of a hydrological station can be a valid IV for whether a firm is in the near upstream area of a monitoring station. The exclusion restriction will most likely hold because within a small bandwidth, a hydrological station in the near downstream should affect a polluting firm only if it brings a monitoring station close to it; otherwise, the downstream hydrological station should have no direct influence on the polluting firm's productivity.

Empirically, we estimate the following first-stage regression:

$$UpMoni_{ij} = \alpha \cdot UpHydro_i + \lambda_j + \epsilon_{ij} \tag{3}$$

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

$$TFP_{ij} = \alpha \cdot \widehat{UpMoni}_{ij} + \lambda_j + \epsilon_{ij} \tag{4}$$

where TFP_{ij} is the TFP of firm i in the neighborhood of monitoring station j; $\widehat{UpMoni}_{ij}$ is the predicted value from the first stage regression; λ_j is the monitoring site fixed effects, and ϵ_{ij} is the error term.

The regression results are presented in Table 4. 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 could indeed strongly predict the locations of water quality monitoring stations (columns 1 and 3). The IV 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 does not affect the productivity of non-polluting firms (column 4).

Note that the regression results in Table 4 are not readily comparable to these in Table 2. In essence, 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 the compliers. Nevertheless, the closeness of the magnitudes of the estimates between the two approaches (0.31 versus 0.35), and the consistent findings in both sets of results, suggest that the relationship between water quality monitoring and firm TFP is likely causal.

D. Sorting

Environmental policies can affect firms' production plans and their location choices. In particular, the pollution heaven hypothesis (PHH) posits that polluting capital would flow from places with more stringent environmental regulations to places with less stringent regulations. At the same time, in our empirical setting, local government officials have an incentive to limit the number of new polluting firms in upstream areas given the importance of water quality to their political careers. This issue is important because it can affect the interpretation of the RD estimates. For example, if water quality monitoring causes more polluting firms to relocate to downstream areas and if these firms have higher TFP, our RD estimates will be biased upwards. Alternatively, if polluting firms emerging in the downstream tend to have lower TFP (which are unlikely to survive in the upstream), our RD estimates will be biased downwards.

However, we believe sorting is not a serious concern in our research for three reasons. First, firms in ASIF sample are large manufacturing firms, which are difficult to relocate. Our discussions with policy makers and firm owners suggest that relocating firms are costly and sometimes politically infeasible. Second, recall that in Table 3, water quality monitoring affects both old and young firms. Presumably, only young firms are able to endogenously choose their locations, thus if sorting were serious, we would observe the impacts of water quality monitoring on news firms' TFP differ. However, the estimated impacts of water quality monitoring are similar between old and young firms. Third, in Figure 5 and Table 5, we plot the location distribution of firms in the polluting industries and conduct a formal density manipulation test on firm location distribution. We observe that polluting firms are continuously distributed across the monitoring stations and cannot reject the null hypothesis using various tests proposed by Cattaneo et al. (2017a, 2007b). The results remain the same if we just focus on the new firms that were born in or after 2003.

E. Spillover Effects

In a perfect competitive market, as there are many firms and output price is unaffected by local environmental regulation, spillover effect, i.e. water quality monitoring somehow also affects downstream firms, is not a concern. 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 firms and downstream firms are the main producers and competitors on the market, increased production costs for the upstream firms will raise the market price of their products. Downstream competing firms will thus benefit because of improved market conditions. Second, tighter environmental regulations may also cause inputs, both labor and capital, to move toward the downstream firms. If more productive factors flow to downstream firms, their TFP will also be higher.

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, regardless of the sign, our baseline RD estimates can still be (properly) interpreted as the partial equilibrium effects of environmental regulation on productivity. However, when we extrapolate these estimates to the whole country, the existence of a positive spillover effect will exaggerate the economic costs of regulation and a negative one will attenuate the estimated costs.

To assess whether or to what extent our findings are confounded by the potential spillover effects, we conduct a placebo test using replaced 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 the monitoring stations based on pre-2003 data; and 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. The intuition is that if the spillover effects are insubstantial (downstream firms are not affected by

monitoring), the placebo firms should have similar TFPs 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 10km 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 emphasize the RD estimates after station and industry fixed effects are absorbed in Panel B. Compared with placebo downstream firms, upstream firms' log TFP 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, if there may exist some spillover effect, this effect should be negative. Consequently, the estimates in Tables 2 and 3 will only understate the economic costs of water pollution regulation.

F. 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 conventional standard errors. 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 difference thereof).[10] The results remain the same regardless of the bandwidth selector used.

In Panel D, 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 themselves and not at the fake stations.

In the appendix, we conduct more robustness checks. In Appendix Table S3, we report the RD estimates using the parametric approach, Equation (2). We find similar quantitatively results: water quality monitoring decreases polluting firms' TFP but has no impact on the 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 S4, we use an alternative TFP measure suggested by Ackerberg et al. (2015) as the outcome variable, and again the results are unchanged.

VI. Channels: What Happened to the Upstream Firms?

A. Regulation and Production

How do firms respond to tighter environmental regulations? We examine the channels through which tighter environmental regulation affects firms' TFP in this section. To guide the empirical investigation, we present a theoretical framework to illustrate how environmental regulation can negatively affect TFP in Appendix B. In this 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 related to the calculation of TFP and test whether these findings are consistent with our theoretical predictions.

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

In Panel A, we examine output-related measures: total revenue, value added, and profit. Although the effects of water monitoring are statistically insignificant on all the outcomes, we see a tendency that downstream firms can earn more profit despite that they generate less revenue.

In Pane B, we focus on input-related measures, including number of employees, net investment, and use of intermediate input. We find that upstream firms need to slightly hire more employees, make somewhat larger investments and use slightly more intermediate input.

In Panel C, we present results for policy related outcomes and firm's research and development investment. We find that upstream firms need to pay larger amounts of taxes and fees than their downstream counterparts and these effects are statistically significant.

In Panel D, we test the Porter Hypothesis. The outcome of interest is firms' investments in research and development. The results show that tighter environmental regulation reduces firm's investment in research and development (statistically insignificant), which contradicts the Porter Hypothesis.

The results in Table 8 suggest that the impacts of environmental regulation on TFP are manifested through multiple channels. Even though many outcomes are statistically insignificant on their own, but the overall pattern is consistent and informative: facing tighter environmental regulations, firms need to hire more labor and/or make more capital investment to reduce emissions. Upstream firms, despite that they do not produce more outputs than downstream firms, have to pay more to the government and earn less profit.

The results also imply that the local governments are able to charge firms with differential tax rates and emission fees, despite that firms are spatially adjacent and are located within the same administrative jurisdiction. In practice, local officials can strategically target upstream firms by inspecting more frequently their environmental performances and increasing the penalty of violations. They can also extract more taxes through informal regulations that are only applied to certain types of firms. This phenomenon has been documented in some other studies as well. For example, Liu (2017) investigates China's income tax reform and find that local governments of China are able to collect more taxes for medium-sized firms after the reform. Fan et al. (2017) study China's value-added tax (VAT) reform and show that firms located further away from local tax agencies experience the largest increase in tax burden after the VAT enforcement costs are brought down by a new information technology.

B. Political Economy of Water Quality Monitoring

Our empirical analysis shows that upstream firms are negatively affected by water quality monitoring. Our explanation is that because water quality readings are politically important, local officials have strong incentives to enforce tighter regulations on upstream firms than on downstream. In this section, we dig into this story and explore the political economy behind water quality monitoring.

First, we examine the political incentives of local officials. As documented in the Chinese meritocracy literature, there exists an implicit rule that a prefecture-level governor cannot be promoted to a higher level (provincial) if his/her age reaches to 57 (for example, Wang, 2016; Xi et al., 2017). This creates a discontinuous drop in political incentives at the age of 56. To test if the TFP effects of water quality monitoring can be explained by political incentives, we digitize the resumes of every Chinese Prefecture Party Secretary (the highest-ranked political leader in a prefecture) between 2000 and 2007, then define a leader as "having strong political incentives" if he/she is younger than 56 in a year, and "having weak political incentives" otherwise. In doing so, we can assign each monitoring station to either a "incentivized" governor or "un-incentivized" party secretary in a given year. The panel data are divided into two subsamples based on whether in a particular year the firm is under the governance of an "incentivized" leader or not. We conduct similar RD analysis using these two subsamples and report the findings in Panels A and B of Table 9.

We find that, when the prefecture city leader has strong political incentives, water quality monitoring has a large and statistically significant impact on upstream firms' TFP. The estimated impacts range from 0.57 to 0.66 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 a precisely estimated zero in all specifications. These results confirm our explanation that the TFP discontinuity across the monitoring stations are driven by political incentives of local officials.

Second, despite 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 and influence the water quality monitoring. In the environmental economics literature, scholars find that some Chinese cities manipulate air pollution data (Ghanem and Zhang, 2014) because air quality is important for political evaluation. Our concern is that if local governments also manipulate water quality readings, they may have different incentives to regulate upstream firms' emissions.

To examine this story, we estimate the RD separately for two types of monitoring stations, i.e., 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 require technicians to conduct the tests manually.[11] As it is difficult for local governments to manipulate data from the automatic stations, we expect a larger TFP-gap around automatic stations.

Panels C and D of Table 9 report the findings: while we see an upstream-downstream TFP gap for both types of stations, this effect is much more salient for automatic stations (almost three times larger). However, as the sample size shrinks substantially for both sub-groups, most of them RD estimates are statistically insignificant at 5 percent level.

C. 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 should not only reduce total emissions, but also adopt cleaner technologies. In this section, we examine the impacts of water quality monitoring on firms' emissions.

Ideally, we would like to examine emissions for the same set of firms covered in our ASIF sample, so that we can directly link the reduction in emissions to the reduction in TFP. However, the ASIF sample does not include information on emissions. Instead, we use the ESR data, which document various types of pollutant emissions for all the major polluting firms in each county.

We apply the same set of RD estimators to firm (polluting-source)-level emission 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. Station fixed effects are absorbed before estimation. Different RD estimates are reported, including conventional

[11] Most stations in the 1990s and early 2000s were manual, but these were gradually replaced by automatic stations 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.

local linear RD estimates, bias-corrected estimates, and bias-corrected estimates adjusted with robust standard errors.

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 5% or 10% level. COD emissions in polluters immediately upstream from monitoring stations are 0.75-0.99 logarithmic units lower than that in 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 5% or 10% level.

Combining both sets of results, we conclude that upstream firms, on the one hand, emit less COD and wastewater overall; and on the other hand, produce fewer COD emissions or less wastewater per dollar 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.

Linking the emission results to findings in Table 8 points to an especially striking pattern: while downstream firms emit 50% more than their upstream counterparts, they actually pay lower levels of emission fees to the government. That implies, local governments use double standards in environmental regulation, which are consistent with our explanation of the research design and findings.

VII. 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 (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 the entire China enforced regulation standards as stringent as those faced by the upstream. The total industrial output value (total revenue) from the polluting firms was about 11 trillion Chinese yuan (1,380 billion US dollars) in 2006;[12] 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.[13]

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 to 85% of China's total emissions, and all of them are local large emitters regardless of industry or revenues. Since 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:

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

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

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

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

[13] 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.

than a given threshold x. TFP_1 is the TFP for downstream firms, and TFP_0 is the TFP for upstream firms.

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

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

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

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

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

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

While we cannot directly estimate "$\mathrm{TFP}_{\mathrm{ATE}}|\mathrm{Revenue} < 5 \text{ million}$" and "$\mathrm{COD}_{\mathrm{ATE}}|\mathrm{COD} < \mathrm{x}$" 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.[14] 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 S3 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 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

[14] 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 number of observations in each group so that there is no enough statistical power to make a reliable inference. In Appendix Table S5, we only divide the sample into two groups and find largely consistent results as in Table 11: the impacts of water quality monitoring are primarily experienced by larger firms or emitters and are negligible for their smaller counterparts.

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.

In China, the government adopts a policy strategy called "Grasping the Large and Letting Go of the Small" ("*Zhua Da Fang Xiao*" strategy). "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 various policy implementations (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.[15] Our findings seem suggest that the "Grasping the Large and Letting Go of the Small" strategy applies to the context of water quality regulations, too.

In addition, there is an "exit" variable in the ASIF database documenting whether a firm will be 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 S6, 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)
$$\mathrm{TFP}_{\mathrm{ATE}}|\mathrm{Revenue} < 5\ \mathrm{million} = 0$$
$$\mathrm{COD}_{\mathrm{ATE}}|\mathrm{COD} < \mathrm{x} = 0$$

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) should also have zero treatment effects. We can therefore simplify equations 7 and 8 to the following:

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

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

The sample we use for estimation includes 6,581 firms in the polluting industry 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 in Table 2 and Table 10, and Panel B calculates the economic costs. We emphasize the estimates in column 1 for subsequent discussion as 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.[16] Using alternative specifications produces slightly different results and 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 (44.2billion US dollars) in the polluting industries based on 2006 industrial output values.[17] 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 industry above designated size) in China exceeded 110 trillion Chinese yuan, and about 35% of output value (38.8 trillion Chinese yuan) is

[16] 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.

[17] 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.

contributed by the polluting industries. The central government aims to reduce COD emissions by another 10% during 13th Five-Year Plan, from 2016 to 2020. Applying our estimates to the 2015 data, we can infer that the total output loss will be around 990 billion Chinese yuan (159 billion US dollars).[18] Other specifications generate slightly different estimates, ranging from 936 to 1099 billion Chinese yuan (150.5to 176.7 billion US dollars).

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 overly strong. 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 the polluting and the 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 their 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; and if environmental regulations are more aggressive in more polluted regions, the TFP loss in these regions can thus 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

[18] We use the 2015 exchange rate of 1 US dollar to 6.22 Chinese yuan.

relocation, and worker migration, and change the flow of investment. These indirect costs are non-trivial and should be considered when calculating the overall economic costs of environmental regulations.

We are only aware of one potential bias that may attenuate our estimated costs of tightening water pollution controls. In China, the quality of environmental data has been widely questioned because of the possibility of manipulation by polluting firms or local governments. In our setting, polluting firms and local governments may be incentivized to underreport emission levels in upstream regions. If this were true, the effect of water quality monitoring on emissions would be overestimated, and thus the economic costs would also be overstated. Through our extensive discussions with policy experts and local government agencies, we learn that despite that the efforts in reducing water pollution are real and substantial, it is difficult to rule out the possibility of data manipulation. To what extent the Chinese government would manipulate water quality and emission data and how it affects our interpretations remain an unknown and future research is warranted on these issues.

VIII. Conclusion

As the income levels of Chinese people rise, the country has started to face a stark tradeoff between preserving high environmental quality and sustaining robust economic growth. This paper is the first large-scale study to estimate the impacts of environmental regulations on productivity in the Chinese manufacturing sector and provides a timely assessment on 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 in the polluting industries, while such a discontinuity is not observed for firms in the non-polluting industries.

We estimate that water quality monitoring reduces TFP levels in firms located immediately upstream from stations by 26.7%. This TFP loss is mainly driven by private Chinese firms instead of state-owned or foreign firms. We also find the TFP effect is larger for older firms than for younger 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 significantly (52.8%–62.8%) less COD and industrial wastewater than downstream firms do; and at the intensive margin, upstream firms adopt cleaner technology and emit less pollution per dollar 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 1099 billion Chinese yuan (150.5to 176.7 billion US dollars) in the polluting industries.

Overall, our findings highlight the negative impacts of environmental regulations on productivity, and the estimated efficiency loss is substantial. In other words, high environmental quality comes at high economic costs, this is particularly salient for fast-growing economies that rely heavily on manufacturing.

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

Nevertheless, our findings do not answer the larger question of whether China's current environmental regulation standards are too aggressive or not, as 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.[19] 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.

[19] Two exceptions are that (1) Ebenstein (2012) finds that China's surface water pollution has caused an increase in deaths from digestive cancers; and that (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.

We conclude by pointing out some limitations of this study and offer directions for future research. First, the estimates in this paper are derived in a partial equilibrium framework. We focus on a unique setting that only affects a small set of firms. Large-scale regulations 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 regulations over long periods of time will be an important area for future research. Finally, with the expectation of increasingly tight environmental regulations in China, entrepreneurs and investors may choose to develop businesses in the non-polluting industries. Tighter environmental regulations in the polluting industries may create externalities on non-polluting industries, and there lack rigorous empirical studies to quantify the impacts of such spillover effects on the economy.

REFERENCES

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.

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.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2016**.** "Rddensity: Manipulation Testing Based on Density Discontinuity." *The Stata Journal*: 1-18.

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.

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, no. 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, 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.

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.
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.
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.
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.
Liu, Yu. 2017. "Informal Taxation and Firm Performance: Evidence from China." Working Paper.
McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698-714.
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.
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.

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

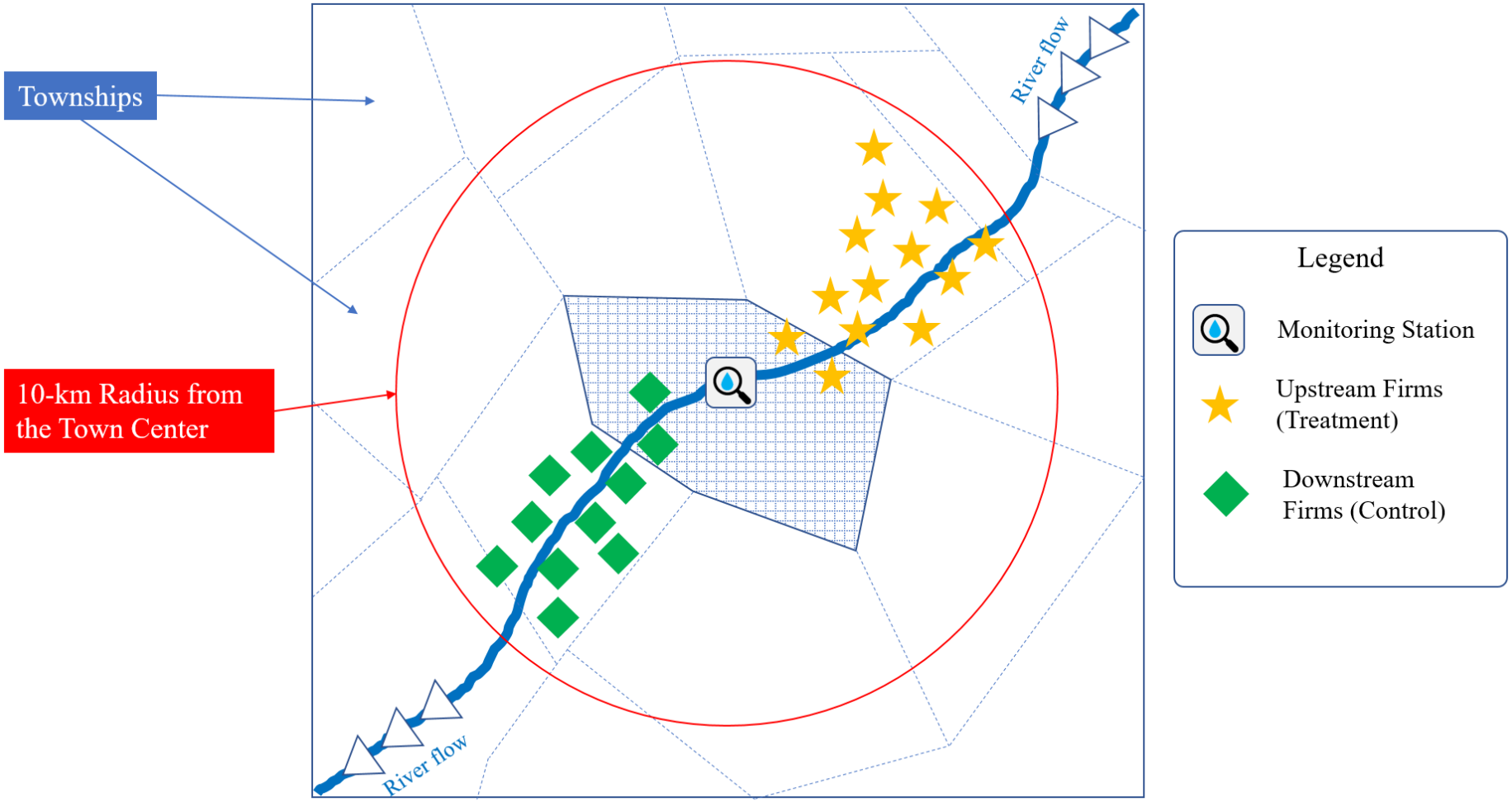

Figure 1. Illustrating the Identification Strategy

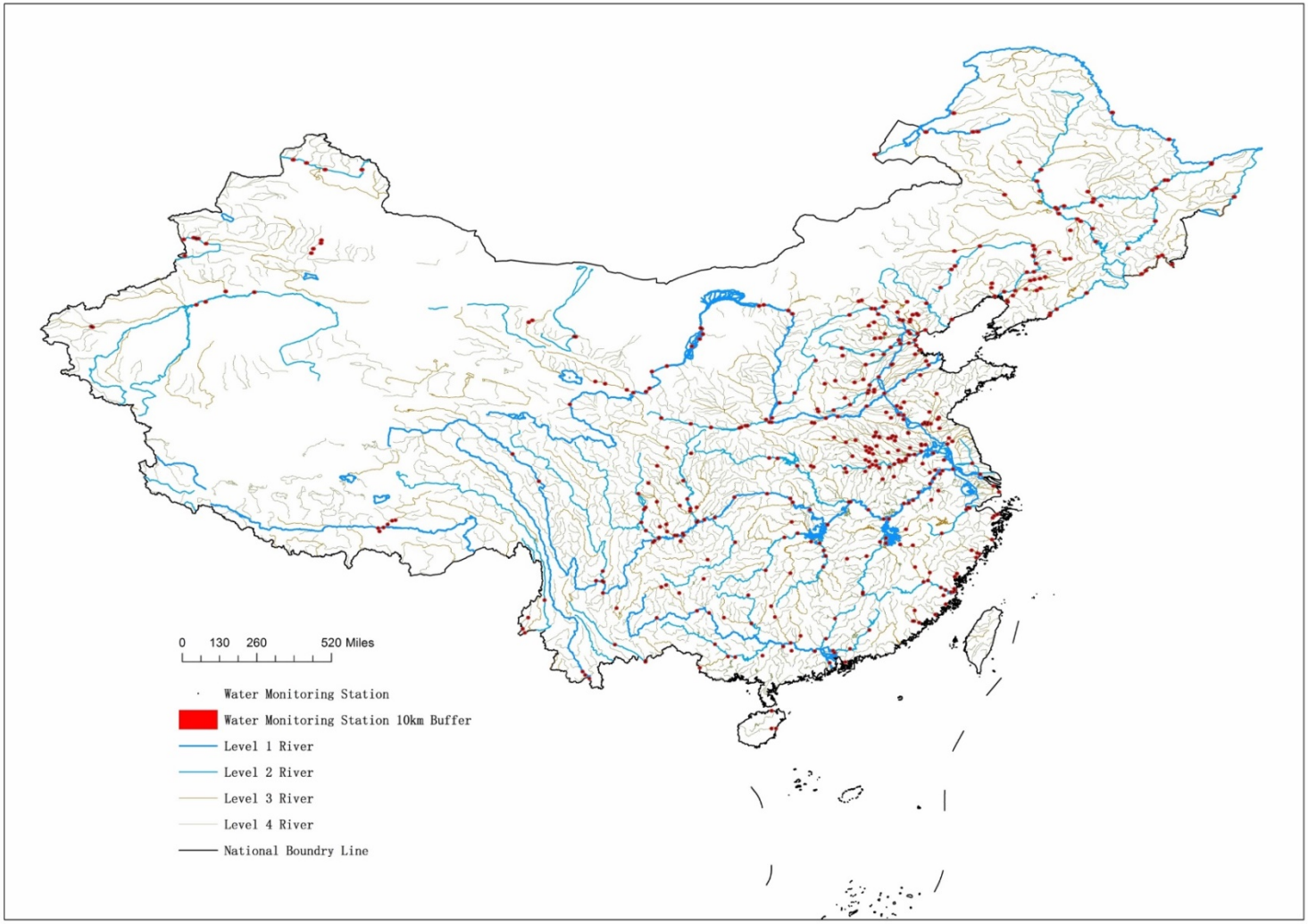

Figure 2. Distribution of Surface Water Quality Monitoring Stations

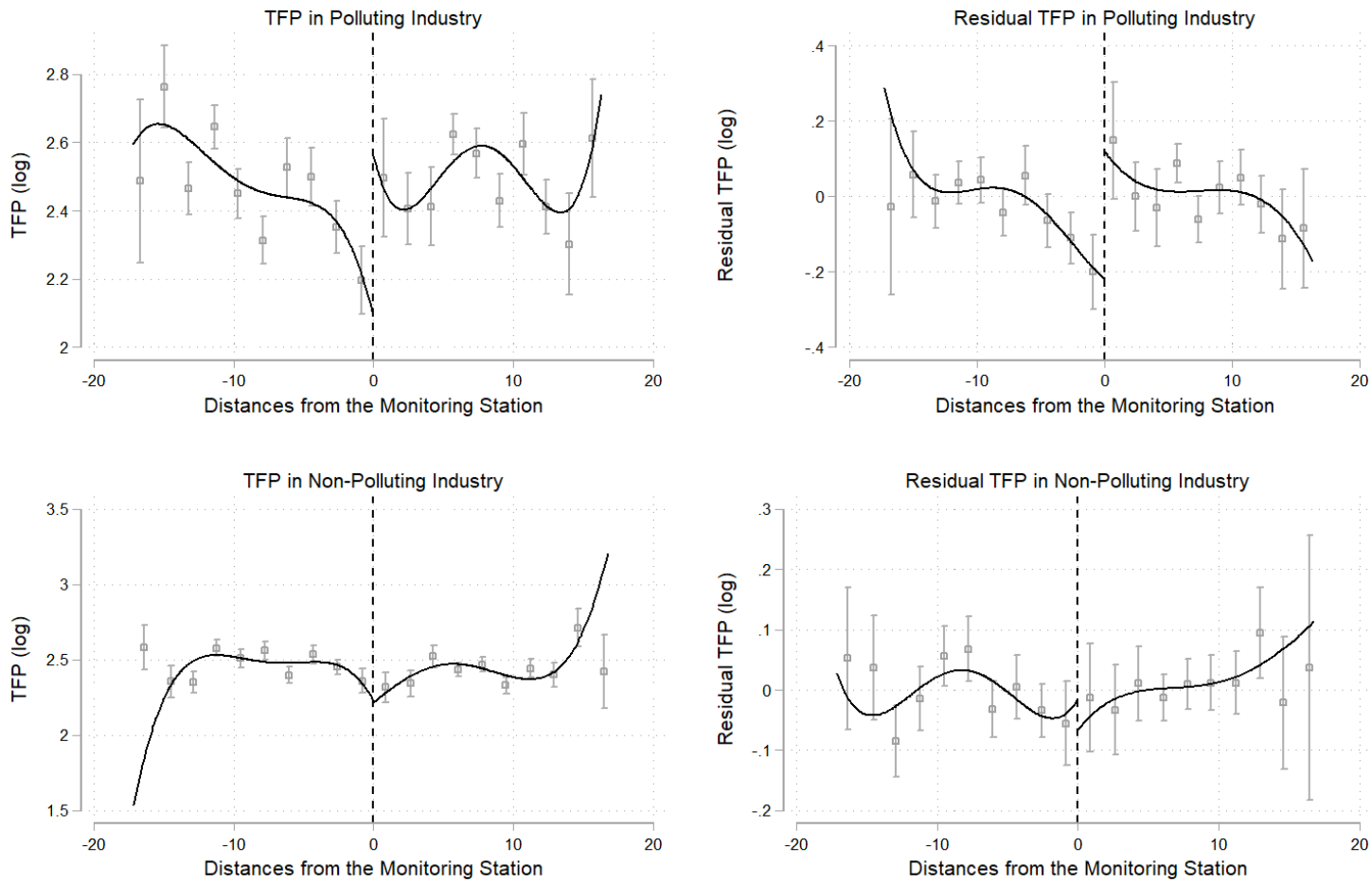

Figure 3. RD Plot: Effects of Water Quality Monitoring on TFP

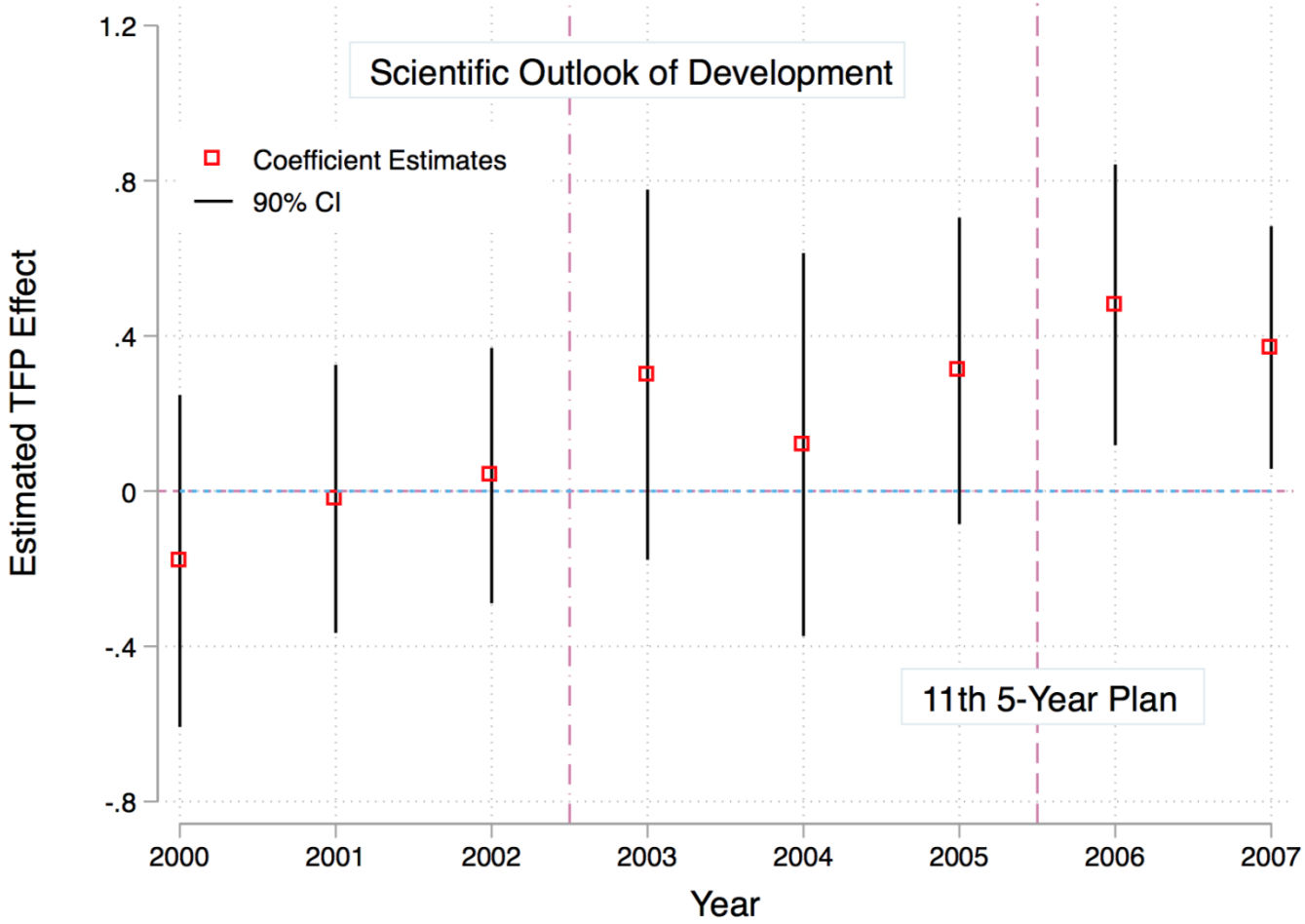

Figure 4. RD Estimates by Year

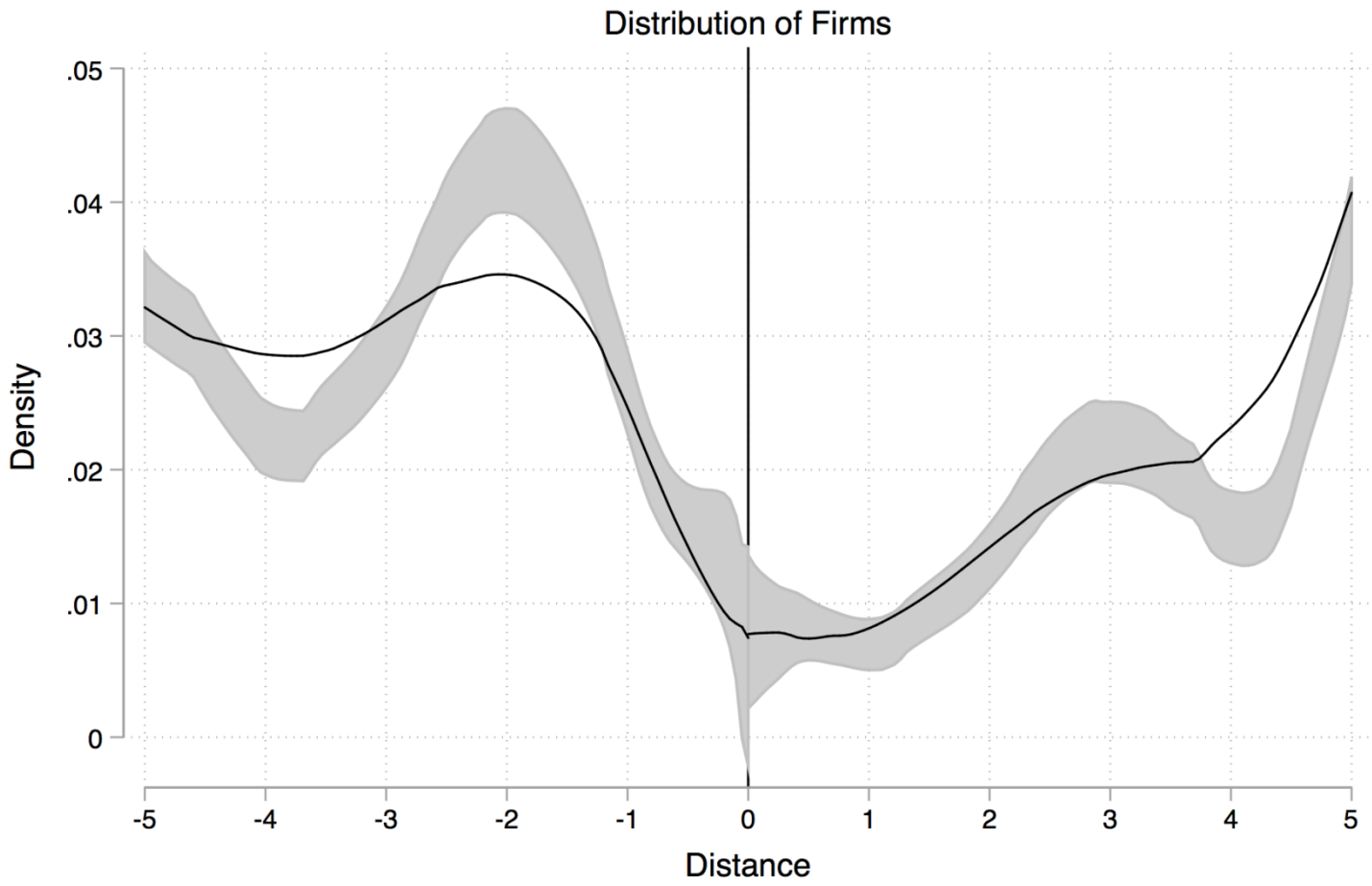

Figure 5. Distribution of Firms

Table 1. Covariate Balance Between Upstream Townships and Downstream Townships

	Mean (within 15km)		Mean Difference		
	Downstream	Upstream	≤15km	≤10km	≤5km
	(1)	(2)	(3)	(4)	(5)
Panel A. Basic Township Characteristics					
Town Area	7,662	7,684	-361.34	-298.63	302.02
(Mu)	(4,267)	(4,171)	(337.78)	(378.11)	(1,467.30)
Arable Area	3,220	2,779	-334.37**	-356.97	-58.29
(Mu)	(1,858)	(1,869)	(154.18)	(217.43)	(803.44)
Distance to County Center	2.39	2.49	0.04	0.10	0.52
(KM)	(0.98)	(1.06)	(0.13)	(0.17)	(0.70)
Old-Region Town	0.22	0.18	-0.02	-0.03	-0.14
(1=Old-Region Town)	(0.41)	(0.39)	(0.03)	(0.04)	(0.18)
Minority Town	0.01	0.02	0.00	-0.01	-0.00*
(1=Minority Town)	(0.09)	(0.14)	(0.02)	(0.02)	(0.00)
No. of Residents Communities	1.96	1.51	-0.16	-0.58	-2.52
	(5.78)	(3.47)	(0.65)	(0.63)	(4.86)
No. of Villages	25.30	23.07	-2.09	-2.12	-1.62
	(16.65)	(14.71)	(1.42)	(1.57)	(5.33)
Panel B. Basic Infrastructure					
Road Length	52.92	47.66	-6.11	-4.10	-3.83
(KM)	(47.05)	(45.38)	(4.55)	(6.52)	(11.57)
# of Villages with Paved Road	24.01	22.38	-1.96	-2.24	-0.57
	(16.09)	(14.38)	(1.39)	(1.47)	(5.51)
# of Villages with Electricity	25.30	23.03	-2.15	-2.12	-1.62
	(16.65)	(14.74)	(1.42)	(1.57)	(5.33)
# of Villages with Tap Water	12.95	10.76	-1.58	-1.39	0.69
	(16.17)	(13.57)	(1.28)	(1.75)	(6.02)
Panel C. Human Capital					
No. of Primary School	18.11	17.53	-0.76	-0.56	2.10
	(9.10)	(9.46)	(0.85)	(1.13)	(3.99)
No. of Primary School Students	7,100	6,186	-629.30	-966.96	-309.76
	(4,611)	(4,312)	(464.15)	(680.70)	(2,329.78)
Obs.	237	307			

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 3–5, we restrict our sample to 15km, 10km and 5km from water quality monitoring stations and 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 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
Bandwidth	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
Bandwidth	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
Bandwidth	6.927	6.541	5.479	3.287	3.070	4.286
Panel B: By Year						
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
Bandwidth	5.722	5.211	3.359	4.375	4.323	3.533
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
Bandwidth	6.223	6.287	5.159	6.302	5.926	5.050
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
Bandwidth	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
Bandwidth	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. 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**		0.31**	
	(0.18)		(0.14)	
Upstream Monitoring Station		-0.35**		-0.00
		(0.16)		(0.17)
Specification	1st Stage	2SLS	1st Stage	2SLS
Station FE	Y	Y	Y	Y
Observations	4,445	4,462	8,976	8,981
F Statistic	10.48	0.03	22.82	1.18
R-squared	0.47	0.16	0.44	0.09

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 5. Density Tests for Sorting Using Local Polynomial Density Estimation

	(1)	(2)	(3)	(4)
Panel A. All Firms, Obs = 6582				
T	0.36	0.21	-2.37	0.36
P>|T|	0.72	0.83	0.02	0.72
Bandwidth Left	2.47	1.97	6.17	2.47
Bandwidth Right	2.01	1.97	6.17	2.01
Panel B, Young Firms, Obs = 2825				
T	0.73	1.30	-1.02	0.08
P>|T|	0.47	0.19	0.31	0.93
Bandwidth Left	2.68	2.00	3.81	2.68
Bandwitdth Right	1.94	2.00	3.81	2.00
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. "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. "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 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)
Monitoring Station FE	Y	Y	Y
Industry FE	Y	Y	Y
Kernel	Triangle	Epanech.	Uniform

Notes: Each cell in the table represents a separate regression. 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						
Revenue	-0.09	-0.09	-0.19	-0.18	-0.19	-0.18
(log)	(0.25)	(0.28)	(0.25)	(0.28)	(0.32)	(0.34)
Value-Added	0.04	0.02	-0.05	-0.07	-0.05	-0.07
(log)	(0.22)	(0.23)	(0.22)	(0.23)	(0.28)	(0.29)
Profit	4.47	4.36	5.48	5.36	5.48	5.36
(10 million yuan)	(4.07)	(3.88)	(4.07)	(3.88)	(5.58)	(5.37)
Panel B. Input Related						
Employees	-0.16	-0.14	-0.22	-0.19	-0.22	-0.19
(log)	(0.16)	(0.16)	(0.16)	(0.16)	(0.21)	(0.21)
Investment	-1.72	-0.60	-2.03	-0.34	-2.03	-0.34
(10 million yuan)	(1.77)	(2.06)	(1.77)	(2.06)	(2.13)	(2.46)
Intermediate Input	-0.18	-0.16	-0.28	-0.26	-0.28	-0.26
(log)	(0.25)	(0.27)	(0.25)	(0.27)	(0.33)	(0.33)
Panel C. Policy-Related and RD						
Tax	-0.59	-0.70	-0.76	-0.87*	-0.76	-0.87
(log)	(0.49)	(0.53)	(0.49)	(0.53)	(0.61)	(0.63)
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 D. Porter Hypothesis						
R&D	-0.06	-0.17	-0.16	-0.28	-0.16	-0.28
(log)	(0.39)	(0.46)	(0.39)	(0.46)	(0.58)	(0.62)
Kernel	Triangle	Epanech.	Triangle	Epanech.	Triangle	Epanech.

Notes: Each cell in the table represents a separate regression. 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. Investment is calculated by subtracting previous year's value of fixed assets from this year's value of fixed assets after decpreciation. Standard errors are clustered at the monitoring station level, and reported below the estimates. * 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. When City Leader Has Strong Political Incenitives						
TFP (log)	0.57***	0.59***	0.63***	0.66***	0.63***	0.66***
	(0.19)	(0.20)	(0.19)	(0.20)	(0.21)	(0.23)
Panel B. When City Leader Has Weak Political Incenitives						
TFP (log) - Polluting Industries	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 Monitoring Stations						
TFP (log)	0.92	1.01*	1.11*	1.22**	1.11	1.22*
	(0.59)	(0.57)	(0.59)	(0.57)	(0.74)	(0.71)
Panel D. Manual Monitoring Stations						
TFP (log)	0.26*	0.26*	0.27*	0.27*	0.27	0.27
	(0.15)	(0.15)	(0.15)	(0.15)	(0.18)	(0.18)
Station FE	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y
Kernel	Triangle	Epanech.	Triangle	Epanech.	Triangle	Epanech.

Notes: Each cell in the table represents a separate regression. 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 uses the subsample where the Prefecture Party Secretary has strong promotion incentives (age<=56). Panel B uses the subsample where the Prefecture Party Secretary has weak promotion incentives (age>56). Panel C uses the subsample of automatic monitoring stations and Panel D uses that of 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*
(Station and Industry FE Absorbed)	(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**
(Station and Industry FE Absorbed)	(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
(Station and Industry FE Absorbed)	(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*
(Station and Industry FE Absorbed)	(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. 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.

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

(Available Soon)

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, which has 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) $$\frac{\partial Q/\partial t}{Q} < \frac{\partial(K+K_E)/\partial t}{(K+K_E)}; \quad \frac{\partial Q/\partial t}{Q} < \frac{\partial(L+L_E)/\partial t}{(L+L_E)}.$$

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:

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

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:

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

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

For emission intensity, we have:

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

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 $p\cdot Q(K,L)$. Since 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 that prices depend on marginal costs, we will underestimate the true TFP effect because price increases as marginal cost of production increases.

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 Manufact	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 Manufactur	34
		General Purpose Machinery Manufactu	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. 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)
Year 2000	-0.18	-0.06	-0.15	-0.22	-0.21	-0.11
	(0.26)	(0.20)	(0.28)	(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 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 S3. 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
Station FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
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
Station FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
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 S4. 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 S5. 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 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 S6. 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 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%.