Journal of Environmental Economics and Management 76 (2016) 86–104

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

Polluting thy neighbor: Unintended consequences of China's pollution reduction mandates

Hongbin Cai [a,*], Yuyu Chen [a], Qing Gong [b]

[a] Guanghua School of Management and IEPR, Peking University, Beijing 100871, China
[b] Department of Economics, University of Pennsylvania, United States

ARTICLE INFO

Article history:
Received 7 October 2012
Available online 29 January 2015

JEL classification:
D62
H77
Q53
Q58

Keywords:
Water pollution
Abatement
Environmental spillover
China

ABSTRACT

This paper studies how the pollution reduction mandates imposed by China's central government in 2001 triggered unanticipated responses from its provinces. We apply the difference-in-differences-in-differences (DDD) method to a unique dataset on industry-level activities in counties along 24 major rivers in China from 1998 through 2008. We find that the most downstream county of a province has up to 20 percent more water-polluting activities than otherwise identical counties since 2001. Moreover, we find that the enforcement of pollution fee collection is more lenient in the most downstream county of a province, and that private firms contribute more to the downstream effect than state-owned enterprises and foreign firms. These findings are consistent with the hypothesis that the provincial governments respond to the pollution reduction mandates by shifting their enforcement efforts away from the most downstream county.

Introduction

China's rapid growth over the past three decades has been accompanied by severe environmental pollution. Deteriorating water quality, pollution-related disputes and accidents, and haze has frequently plagued Chinese cities, raising serious concerns among both the public and the central government (Chan and Yao, 2008; vanRooij, 2010). River pollution is a particularly serious problem. A mere 28 percent of the country's 500 monitored river sections report drinkable water quality, and one-third are so contaminated that the water is unsuitable for drinking, agriculture, or any other common uses (World Bank, 2006). China's economic losses from water pollution are estimated to be around 150 billion yuan per year, and losses of health and life associated with water pollution are enormous but impossible to estimate (World Bank, 2007).

As a country with one of the lowest per capita fresh water availability rates in the world,[1] the Chinese central government became alarmed at the severe river pollution in recent years. In its Tenth Five-Year Plan, released in 2001, the central government for the first time added environmental protection and pollution reduction to its list of “national strategic goals” and set a target to reduce pollutant discharges by 10 percent by the end of 2005 (State Council, 2001). Each province was assigned a specific target, and the provincial government officials were to be evaluated on, among other things, how well these targets were met. Despite the central government's resolution, China's water quality saw almost no improvement over the 15 years between 1991 and 2005 (World Bank, 2006).

* Corresponding author. Fax: +86 10 62751470.
E-mail address: hbcai@gsm.pku.edu.cn (H. Cai).

[1] China's per capita renewable water resource availability was 2156 m^3/year in 2007, one-fourth of the world average.

http://dx.doi.org/10.1016/j.jeem.2015.01.002
0095-0696/

In this paper, we investigate the effects and consequences of the 2001 policy change by the Chinese central government that imposed pollution reduction mandates on its provinces. We identify 24 major rivers in China and study the polluting activities and location choices of industrial firms along the provincial borders. We aggregate firm information from the Annual Survey of Above-Scale Industrial Firms in China from 1998 through 2008 to the county level and the two-digit industry level; we also collect information about county characteristics from other sources. To control for confounding factors in firm location and production choices, such as transportation, geographical features, and industry characteristics, we take the difference-in-differences-in-differences (DDD) approach, constructing control groups using non-water-polluting industries and non-riverside counties. These control groups help us to eliminate many unobserved county and industry heterogeneities. In addition, we control for a range of county socioeconomic, demographic, and other characteristics, as well as a rich set of potentially time-variant two-digit industry effects and county group effects in order to mitigate selection based on unobservable variables.

Using the DDD method, we find strong evidence of the *downstream effect*. That is, all else being equal, the most downstream county of a province has up to 20 percent more water-polluting activities than otherwise identical counties. This leads to the phenomenon of "polluting thy neighbor" in that provinces concentrate water-polluting activities in the most downstream counties, thus shifting the burden of water pollution to their downstream neighbor provinces.

The downstream effect sheds light on why water quality has not improved even though the central government of China has been emphasizing environmental protection since 2001. At the time of the policy change, the central government set pollution reduction targets for each province (see section "Institutional background" for details) but failed to anticipate the provincial governments' responses in how they would meet the targets. Under the pressure from the central government to curb river pollution, growth-driven provincial governments responded by optimally allocating enforcement efforts among their counties: given the externalities inherent in river pollution, the provinces cannot reap the full benefits of pollution reduction, especially in the most downstream counties. At the same time, the crude pollution monitoring technology adopted by the central government gave the provincial governments considerable power over the enforcement of environmental regulations. Therefore, provinces tend to exert the least enforcement efforts in the most downstream counties, resulting in the increase of water-polluting activities at the downstream provincial border.

Our empirical analysis finds strong evidence that is consistent with the mechanism proposed above. Specifically, we find that (i) the downstream effect is absent among interior counties, and is much weaker before 2001; (ii) the enforcement of pollution fee collection is more lenient in the most downstream county of a province than in other counties; and (iii) private firms, which are more sensitive to the enforcement of environmental regulation than state-owned enterprises (SOEs) and foreign firms, contribute the most to the downstream effect. These findings suggest that the downstream effect is due to the *strategic polluting* of provincial governments in response to the central government's pollution reduction mandates.

Our paper is closely related to the literature on river pollution (Sigman, 2002, 2005; Bernauer and Kuhn, 2010; Lipscomb and Mobarak, 2013). Recognizing the unidirectional externalities in river pollution, Sigman (2002) uses cross-border comparisons to show that pollution levels are higher upstream of national borders in many countries. Sigman (2005) uses variations of when states were authorized to issue pollution permits in the U.S. to identify strategic polluting across state borders. Making use of county border changes in Brazil, Lipscomb and Mobarak (2013) study river pollution spillovers across riverside counties and investigate the overall effect of decentralization on water quality.

Our paper differs from previous studies in three ways. First, our identification strategy using the DDD approach offers a new way of controlling for unobserved county and industry characteristics. Second, unlike existing papers that focus on pollution outcomes or water quality (e.g., chemical oxygen demand (COD)) measured at monitoring stations, we use indirect measures, the amount of pollution-generating production activities (the industrial value added and the number of firms in water-polluting industries) and firm location choices (the number of new firms in water-polluting industries), as our dependent variables. This is not by choice, as we do not have water quality data from monitoring stations for the years we are studying. Nevertheless, our measures not only complement the use of water quality data but they also have several advantages. One advantage is that our measures are all county-specific, making it easy to perform cross-county comparisons. On the contrary, point observations of water quality at monitoring stations do not directly reflect the pollution level in the counties they belong to, as the water quality at any point depends on the cumulative effects of all polluting activities upstream. One would need very specific assumptions about the pollution decay function and about industry distributions to deduce the contribution of pollutants from the upstream counties. Another advantage of our measures is that they are less subject to misreporting than direct measures of water quality would be in China. Previous researchers (Sigman, 2002; Bernauer and Kuhn, 2010) have warned of the strategic reporting of water quality in other countries. Third, our paper goes beyond identifying the downstream effect. The existing literature on pollution spillovers usually alludes to the strategic choice of enforcement efforts by local governments as the reason behind negative spillovers. But the connection is often not made explicit.[2] In this paper, we use information on pollution fees and take advantage of a policy change to make explicit the provincial governments' incentives as the driving force behind the downstream effect that we document.

[2] An exception is Konisky and Woods (2012), who use the number of environmental inspection visits to polluting facilities as a proxy for enforcement efforts. However, they find that there are *more* inspection visits in counties along the state borders in the U.S. than in interior counties.

Our paper also adds to the literature that tries to empirically identify transboundary environmental pollution and pollution spillovers. A branch of the existing literature focuses on the "border effect" by examining differences in pollution levels or health outcomes between border and interior jurisdictions (Helland and Whitford, 2003; Kahn, 2004; Gray and Shadbegian, 2004). We also find such border effects in that there are significant differences between counties on provincial borders and interior counties in the amount of water-polluting activities and in new entry into water-polluting industries (see section "DID analysis"). However, since riverside counties on provincial borders can be inherently different from their interior counterparts, it is hard to attribute border effects to strategic polluting. Thus in this paper we use the aforementioned DDD identification strategy to help control for unobservable county characteristics and identify the downstream effect.

The remainder of this paper is organized as follows. The section "Institutional background" provides the institutional background of China's water pollution and regulatory structure of environmental protection. The section "Empirical strategy" discusses our empirical strategy. The section "Data" describes the data we use. The section "Estimation results" presents the main empirical results. The section "Mechanism of strategic polluting" proposes a mechanism for the downstream effect and examines its four testable hypotheses. The section "Concluding remarks" concludes.

Institutional background

River pollution in China

Environmental deterioration is one of the most undesirable byproducts of China's rapid economic growth in the past three decades. Each year 750,000 people die prematurely as a result of air or water pollution (World Bank, 2007). Moreover, the poor are disproportionately affected by water pollution. 300 to 500 million of China's rural populace lack safe drinking water; and diarrheal diseases and digestive system cancers due to polluted drinking water cost 1.9 percent of rural GDP every year. The Chinese central government has been alarmed at the severe river pollution in recent years. Yet there has been little improvement, if any, in overall water quality (World Bank, 2007).

Industrial activities are the most significant source of river pollution in China, producing more pollutants than agricultural and domestic sources. According to the First Census of Polluting Sources in China (Ministry of Environmental Protection, 2010), industrial activities generate more than 35 percent of pollutant discharges into Chinese rivers, but their overall environmental damage is much higher because the concentrated, highly toxic industrial pollutants overburden the self-cleaning capacities of rivers and exacerbate the deterioration of water quality (Wu et al., 2000). The following seven industries are the major water polluters in China: agricultural products and byproducts, textile manufacturing, garments manufacturing, pulp and paper manufacturing, petroleum and nuclear fuel processing, chemical manufacturing and processing, and non-ferrous metals smelting and pressing. They contribute 70 percent of total industrial ammonia nitrogen discharges (Ministry of Environmental Protection, 2010). These seven industries will be labeled *water-polluting* industries in our analysis that follows, while others will be called *non-water-polluting* industries.

Regulatory structure of environmental protection

Until 2008, the regulatory agency of environmental protection in China was the Bureau of Environmental Protection (BEP).[3] The central BEP's main responsibilities include setting national policies and regulations of environmental protection and monitoring enforcement activities of local BEPs. BEPs at each level of local government (province, prefecture and county) are in charge of enforcing the environmental protection regulations in their own localities. They mainly use two types of instruments, the ex ante permit system for industrial projects and ex post monitoring and punitive measures. The permit system requires that all industrial projects obtain approval from the local BEPs before production to ensure that new projects meet the basic environmental standards. The ex post punitive measures include warnings, fines, suspension of business licenses, and legal action.

One important feature of China's environmental protection regulatory structure is that local BEPs are controlled by the same-level local governments. Local governments in China are given strong incentives to pursue economic growth and tax revenue, while other social objectives such as environmental protection are secondary concerns at best (Zhou, 2008). Given the local governments' objectives, local BEPs are often instructed to be lax on regulation enforcement. New industrial projects that increase GDP and fill up local coffers are approved even if they do not meet the environmental standards.[4] Moreover, local BEPs are often encouraged by government officials to turn a blind eye to environmental violations in order to create a "good business environment" (Zhang, 2012).[5]

[3] In 2008, it was elevated to the Ministry of Environmental Protection (MEP) to signal the central government's increased commitment to environmental protection.

[4] For example, the Chinese Academy of Environmental Planning blamed local governments for insufficient efforts in pollution control and for violating central government policies by allowing heavy polluters like pulp and paper mills to operate (Liu, 2006).

[5] For example, the bureau of environmental protection of Hebei Province processed 20 "serious" environmental violations in 2003, and the highest fine was only 100,000 yuan (about 12,000 US dollars).

Environmental protection policy after 2001

The 2001 pollution reduction mandate

Economic growth has been the dominant policy objective for the Chinese government since the 1980s, as Deng Xiaoping once said "Development is the top priority" (Deng, 1992). Environmental protection was not an important concern until pollution became too severe to ignore. The major policy shift came in 2001 with the release of the Tenth Five-Year Plan, in which protecting the environment and reducing pollution became the national strategic goals for the first time (State Council, 2001). Originating from the central planning era, the Five-Year Plans, which set goals for the country's development in the upcoming 5 years, are still the most important policy instruments and the highlights of major policy change in China. Therefore, when the Tenth Five-Year Plan sets the target of reducing pollutant discharges by 10 percent by the end of 2005, it was viewed as a clear signal of policy changes in China's environmental protection.

The National Development and Reform Committee (NDRC) complemented this overly general target with a follow-up document, detailing almost every aspect of the pollution reduction mandate and its implementation (National Development and Reform Committee, 2002). Specifically, the NDRC first listed targeted air, water, and solid waste pollutants, where the major water pollutants considered were chemical oxygen demand (COD) and ammonia nitrogen (NO_x). The NDRC then decomposed the overall pollution reduction target across major rivers. This is also why we focus on the 24 major rivers, and why we choose "non-riverside" counties, i.e. those not located on a *major* river, as a control group. Targets for the rivers are roughly the same, varying only slightly around the 10 percent mark. Lastly, the NDRC further decomposed the targets by province along each major river. The provincial targets again vary little. Each province, facing pollution reduction mandates imposed by the central government, has discretion over how to further delegate pollution reduction targets to lower level governments, such as prefectures and counties (Wang, 2002).

Monitoring pollution reduction

The monitoring technology for pollutant discharges adopted by China's central government, however, was very crude in that the pollutant discharges were not directly measured (State Council, 2003). Instead, discharges were estimated using an industry-specific formula that converted the production activities of each firm to pollutants. Although there were mandates to install pollutant-measuring devices that could automatically transmit data to the central BEP, such mandates only apply to 4000 "major polluting establishments" selected by the central BEP, and only required installation to be completed by the end of 2008. Water quality data, although available at river monitoring stations, could not identify the pollutant sources, hence were not used as an index for performance.

The estimation of pollutant discharges was implemented at the county level, and then reported to the provincial governments for aggregation and review. Each province then reported its total estimated discharges to the State Council and the central BEP by the end of every March. The reports were jointly audited by the central BEP, the NDRC, the National Bureau of Statistics, and the National Audit Office by the end of May. The degree to which these pollution reduction targets were met was used, among other things, to evaluate local government officials. Provinces that failed to meet their target needed to report to the State Council within 30 days and submit a detailed plan for reducing future pollution.

One feature of the pollution reduction mandate in 2001 was that it did not include any mechanism for coordination across provinces. Even though environmental protection is a public good, the NDRC simply decomposed the targets, and did not mention how local governments could work together to meet these targets more effectively. The central BEP was also silent on this issue in its policy guidance for implementing the pollution reduction mandates (Bureau of Environmental Protection, 2002). This lack of coordination was later acknowledged by the central BEP as one reason for the failure of pollution reduction (Chinese Academy of Environmental Planning, 2006).

Another feature of the mandate is that it potentially allowed the provinces to manipulate their polluting activities. As discussed in section "Regulatory structure of environmental protection", the provincial governments can greatly interfere in the operation of local BEPs. In addition, they can also manipulate the distribution of polluting activities through issuance of project permits. Industrial projects in China above 30 million yuan used to need the approval from the National Planning Committee. Very large and important projects were also reviewed by the State Council (1984). But the issuance of project permits became more decentralized over the years. By 2001, the approval of all projects was already under provincial control (National Planning Committee, 2001).[6] The continuing decentralization has granted provincial governments the power to strategically allocate polluting activities among their counties.

Compliance with the 2001 mandate

At the end of the five-year period in 2005, China as a whole fell far short of the pollution reduction targets, despite the central government's increasing emphasis on environmental protection. COD was only reduced by 2 percent, and other pollutants increased significantly (e.g., there was a 27 percent increase in SO_2). This was quite astounding, as almost all other national targets, especially growth targets, had been easily met or surpassed.[7] Local governments were usually strongly

[6] The exceptions are very large power plants, inter-province infrastructure projects, "sensitive" industries, and joint ventures with Chinese investments above 200 million U.S. dollars.

[7] See Section One (review of government work in the previous year) of the State Council's "Report on the Work of the Government" from 1999 through 2009 (State Council, 2013).

motivated to achieve their performance targets, and the central government got what it asked for. However, this was not the case for pollution reduction from 2001 through 2005. A common explanation was that local governments were too eager to meet the economic growth targets, and chose to overlook regulations of environmental protection and allowed many heavily water-polluting industries such as pulp and paper manufacturing to expand too fast (Chinese Academy of Environmental Planning, 2006).

After the setback in the Tenth Five-Year Plan, the central government modified its targets of environmental protection in the Eleventh Five-Year Plan for the period of 2006 through 2010. Major pollutants such as COD and SO_2 were to be reduced by 2 percent each year from the 2005 level. These targets were missed again in 2006, when COD increased by 1.2 percent. Then Premier Wen Jiabao criticized the lax enforcement of local governments for this failure in his report to the National People's Congress in 2007 (Wen, 2007). The central government responded by raising the stakes for local officials who did not comply with the mandate (Naughton, 2006). At the end of 2007, the State Council ordered that local officials would be immediately removed if their jurisdiction failed to meet the pollution reduction target (State Council, 2007). As the central government put increasing pressure on the local governments, the overall reduction targets of the Eleventh Five-Year Plan were finally met by the end of 2010. This achievement notwithstanding, the standards were quite modest and it was not clear whether the worsening trend of environmental pollution was indeed reversed (Zhang, 2012).

The implication of the environmental protection policy since 2001 is twofold. It raised the stakes of environmental pollution for the provincial governments, thus incentivizing them to respond by manipulating pollution within their jurisdictions. Moreover, the institutional setting, especially the design of the targets and the choice of monitoring technologies, granted the provincial governments considerable freedom in their responses. Now that the provincial governments have both the incentive and the ability to optimize the allocation of pollution reduction efforts, we expect such efforts to be the lightest in the most downstream counties. This is because, given that there are externalities in reducing river pollution, a province cannot reap the full benefit of its pollution reduction efforts in such counties. As a result, we will see rising levels of pollution in the most downstream county of a province, i.e. the downstream effect. The differential enforcement efforts induced by the pollution control mandates and the negative externalities inherent in water pollution jointly caused this downstream effect through the unexpected strategic responses of the provincial governments.

Empirical strategy

The main objective of our empirical analysis is to identify and measure the downstream effect in China's river pollution. As described before, we expect to see rising levels of pollution in the most downstream county of a province, as a result of the provincial governments' optimal responses to the pollution reduction mandates. The major challenge to our empirical analysis is the selection problem. The most downstream counties differ from other counties in various ways, and firms' responses to these differences depend on their own characteristics as well. Therefore we need to disentangle the many confounding factors of firm location and production choice, and identify, to the best precision possible, the downstream effect that we are interested in. We use the DDD method to achieve this goal.

To illustrate our empirical strategy, let us consider the heuristic map in Fig. 1. Suppose a river flows from the west to the east, crossing the most downstream county *A* in an upstream province *X*, and then the most upstream county *B* in a downstream province *Y*. County *a* is a neighbor of county *A* in province *X*, and county *b* is a neighbor of county *B* in province *Y*. Just like counties *A* and *B*, counties *a* and *b* are also neighbor counties separated by the provincial border. But unlike counties *A* and *B*, counties *a* and *b* are not riverside counties. In addition, counties *A* and *a* share the same provincial characteristics and are likely to share similar geographic features; the same is true of counties *B* and *b*. We call the four counties *A*, *B*, *a*, and *b* a *county group*. Counties will be classified into "types" by their corresponding positions in the county group (e.g., "type-*A* counties" refer to the most downstream riverside counties in an upstream province).

Our identification strategy essentially consists of three steps. In the first step, we compare the difference in water-polluting industries between counties *A* and *B* with that of non-water-polluting industries. This is a standard difference-in-differences

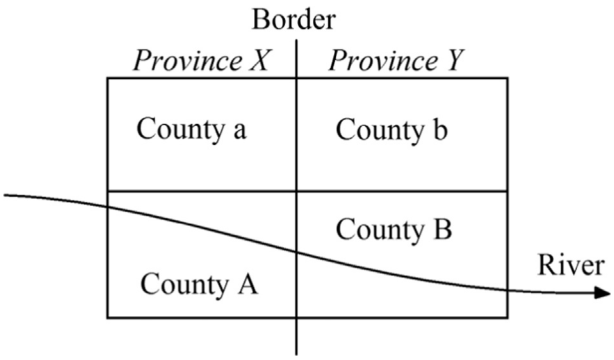

Fig. 1. Heuristic Map of Counties at the Provincial Border.

(DID) exercise, where non-water-polluting industries are the control group for water-polluting industries within the same county. The DID analysis removes the heterogeneities of county *A* and county *B* that are not industry-specific, such as those due to local labor market conditions and geographical features. This first step alone, however, cannot remove all the confounding factors. Water-polluting industries can be inherently different from non-water-polluting ones, making the latter an inappropriate control group for the former. For example, compared with downstream provinces, upstream provinces may be less affluent but have more natural resources, which could be more attractive to water-polluting industries.[8] Then to pursue comparative advantages, upstream provinces (such as *X*) would have more water-polluting industries than downstream provinces (such as *Y*), relative to non-water-polluting industries. In such cases, our DID result may pick up those industry-specific provincial differences in addition to the downstream effect. This is why we need an additional dimension of difference to identify the true downstream effect.

In the second step, we conduct a DID analysis of water-polluting versus non-water-polluting industries between the non-riverside county pairs (such as counties a and b). In the third step, we examine the difference between the DIDs from the previous two steps. This final step removes the confounding factors discussed above. Since county pairs such as *a* and *b* are not located along a major river, the average effect of the DID analysis in the second step is not related to strategic polluting and only captures the aforementioned industry-specific provincial differences. Subtracting this from the first-step DID eliminates the confounding factor and pins down the true downstream effect.

The key identifying assumption of the DDD framework is that there is no selection based on unobservables that correlate with the key explanatory variables. This assumption would be violated if, for example, firms in some industries select into different locations based on their need for inexpensive ground transportation, which is unobservable to the econometrician. If riverside counties have systematically better access to such transportation, then the selection problem will contaminate our DDD analysis. To solve the selection problem, we have controlled for a range of county and industry characteristics and included a full set of industry fixed effects, county group fixed effects, and their interactions with time. We argue that these have captured most of the relevant unobservable variables, and will significantly mitigate potential selection.

We implement the DDD analysis with the following regression:[9]

$$Y_{\mathrm{ijt}}=\beta_0+\beta_1\mathrm{DOWN}_j+\beta_2\mathrm{RIV}_j+\beta_3\mathrm{POL}_i\cdot\mathrm{DOWN}_j+\beta_4\mathrm{POL}_i\cdot\mathrm{RIV}_j+\beta_5\mathrm{DOWN}_j\cdot\mathrm{RIV}_j+\beta_6\mathrm{POL}_i\cdot\mathrm{DOWN}_j \cdot\mathrm{RIV}_j+\beta_7X_{\mathrm{jt}}+\eta_{\mathrm{it}}+\delta_{\mathrm{Jt}}+\varepsilon_{\mathrm{ijt}}$$

where *i*, *j*, and *t* indicate industry, county, and year, respectively. Y_{ijt} is a measure of activity of industry *i* in county *j* in year *t*. DOWN_j is a county dummy: it is set to 1 if county *j* is the most downstream county in its province (such as county *A* and county *a* in our heuristic example, even though county *a* is not a riverside county), and 0 otherwise (such as county *B* and county *b*). RIV_j is another county dummy: it is set to 1 if county *j* is located along a river (such as county *A* and county *B*), and 0 otherwise (such as county *a* and county *b*).[10]

The interaction term $\mathrm{POL}_i\cdot\mathrm{RIV}_j$ captures the industry-specific and riverside-specific effects on *Y*; and the interaction term $\mathrm{POL}_i\cdot\mathrm{DOWN}_j$ captures industry-specific and upstream-province-specific effects. The focus of our DDD analysis is the triple interaction term $\mathrm{POL}_i\cdot\mathrm{DOWN}_j\cdot\mathrm{RIV}_j$, so β_6 is the parameter of primary interest to us. It captures the average effect of being in the most downstream county of a province along a major river on water-polluting activities and firm location choices net of other confounding factors, i.e. the pure downstream effect. We attribute this effect to inter-provincial strategic polluting.

As discussed above, we include a range of county characteristics to try to address the problem of selection based on observable county characteristics. X_{jt} is a set of control variables that represent county *j*'s socioeconomic and demographical characteristics in year *t*. To further mitigate selection based on unobservables, we include η_{it} and δ_{Jt}, the full set of two-digit industry and county group effects, both allowed to vary with time, where δ_{Jt} is the effect of being in county $j\in J$ and year *t*, with *J* being an index of county group. Lastly, ε_{ijt} is the error term. We allow for correlation across two-digit industries within county and cross section by clustering the standard errors.

In the existing literature, many scholars have investigated the effects of environmental regulations on firm pollution and location choices, with recent contributions such as Becker and Henderson (2000), Greenstone (2002) and Kahn and Mansur (2013). Besides exploiting variations in environmental regulations across counties, Kahn and Mansur (2013) use a border pair approach and take into account the variations in electricity prices and labor market regulations as well. Our empirical analysis is similar to these papers in that we examine the factors that cause clustering patterns of industrial activities and firm location choices. However, in our context, there are no variations across provinces or counties in environmental regulations, but we argue that there are variations in the *enforcement* of environmental regulations as a result of the provincial governments' intentional exploitation of the negative externalities in river pollution. To this end, our DDD identification strategy described above controls for geographic, social, economic and institutional characteristics that might

[8] In China, the Eastern coastal region is more developed than the Central region, which is in turn more developed than the Western region. Most of the major rivers are eastbound.

[9] We omit the specifications of the DID analysis in the first two steps since they are standard.

[10] Theoretically the specification should include an additional dummy variable, POL_i, which takes the value of 1 if industry *i* is water-polluting and 0 otherwise. However, since we are already controlling for two-digit industry fixed effects in all regressions, the POL_i dummy is not identified and therefore removed.

affect industrial activities or firm location choices.[11] In section "Mechanism of strategic polluting", we further take advantage of the information on pollution fees and other related empirical findings to analyze the mechanism behind the observed clustering patterns of water-polluting activities and firm location choices in the most downstream counties.

Data

For our main empirical analysis, we construct a sample with data from 8 years (2001 through 2008), each with more than 4000 observations. Each observation represents a two-digit industry in a county in a year, with information on production activity, the number of firms, the number of new firms, whether the industry is water-polluting, county location type (*A*, *B*, *a*, or *b*), and other county characteristics. The data we use come from the following three sources.

Firm data

We use the total value added and the number of firms in an industry as a proxy for its production activities, and we use the number of new firms in an industry in a county as a proxy for firm location choices. The dependent variables are the log of total value added,[12] the number of firms, and the number of new firms per industry per county per year. Such information is generated from the Annual Survey of Above-Scale Industrial Firms data collected by the National Bureau of Statistics, which is available every year from 1998 through 2008.[13] The data contain basic firm information (name, address, industry, age, ownership, etc.) and major financial statement items (total assets, revenues, expenses, etc.). From the firm-level data we calculate the total value added, the number of firms, and the number of new firms in every two-digit industry in a county for each year from 1998 through 2008.[14] We identify seven major water-polluting industries according to the "Report on the First National Census of Polluting Sources" (Ministry of Environmental Protection, 2010). We summarize the total value added and the total number of firms in each of the 30 two-digit industries in the appendix (Table A1). We have excluded mining industries, because the location of mining firms is largely driven by the distribution of natural resources, not by human choices.

County location data

We focus on the 24 longest rivers in China, each of which crosses at least one provincial border. These rivers make up 93.4 percent of annual river runoff in China, and their river basins cover 78.1 percent of the country's land area. We provide summary statistics by river in Table A2.

Along these 24 rivers, we first identify 116 riverside counties (such as counties *A* and *B*) located at provincial borders. Then for each of these counties, we identify a non-riverside neighbor county from the same province which is also at the provincial border (such as counties *a* and *b*). Our final sample has 232 counties in total, or 116 county pairs at provincial borders, of which there are 58 riverside county pairs and 58 non-riverside county pairs. These 232 counties are classified into 58 county groups, each having four neighboring counties (as in Fig. 1). For each county group, we keep an industry if it appears in at least one of the four counties. If an industry is not present in any of the four counties in a county group, we exclude it from our sample to avoid having too many observations with value zero. We end up with a sample with 33,312 observations, covering 232 counties, 30 industries, and 8 years.

To take a first look at the pattern of the raw data, we plot in Fig. 2 the log of average firm value added from 1998 through 2008 for water-polluting industries and non-water-polluting industries in type-*A* counties and in type-*B* counties. It is clear that type-*A* and type-*B* counties produce almost the same level of value added in non-water-polluting industries since 2001, but type-*A* counties have significantly higher value added in water-polluting industries than type-*B* counties. Even though there may be many confounding factors that are uncontrolled for, this clear pattern at least suggests that strategic polluting might be an important factor.

---

[11] There are no institutional variations in electricity prices or labor market regulation across counties in China, and we are not aware of data about them that can be included in the control variable set X_{jt}. Nonetheless, our empirical strategy is still valid as long as any the variation in electricity prices or labor market regulation is not systematically correlated with the differences between water-polluting and non-water-polluting activities across type-A and type-B counties.

[12] Firm-level value added is missing for 2008. We retrieve the missing value added information from the firm's "value-added tax due" in that year, using the then prevailing standard value-added tax rate of 17 percent. Excluding these firms when constructing our sample does not change the regression results.

[13] The above-scale industrial firms are those with annual revenues above 5 million yuan (approximately 600,000 U.S. dollars at the then prevailing exchange rates).

[14] We use the two-digit industry classification for two reasons. First, the seven water-polluting industries are defined by the two-digit classification. Second, with two-digit industries, the number of firms in an industry per county per year has a sample mean of only 2.77 (Table 1). Thus, if we adopt the more refined three-digit or four-digit classifications, there will be too many observations with zero firms.

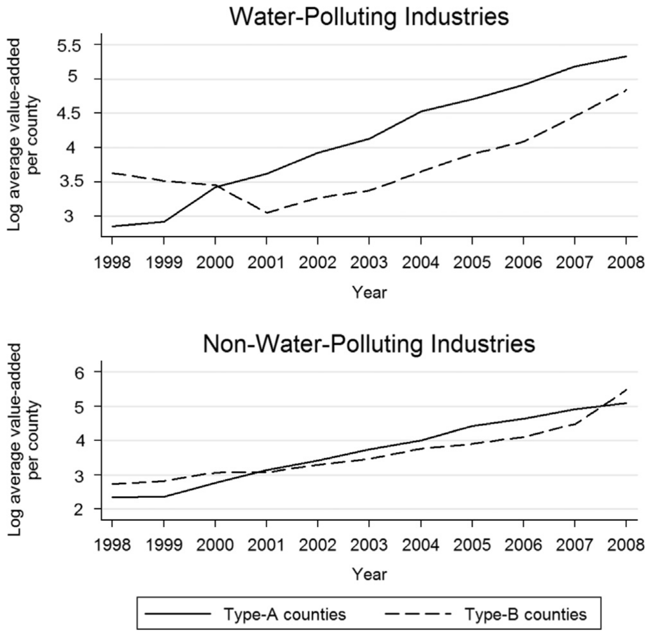

Fig. 2. Log Average Value Added by County Location. *Source*: National Bureau of Statistics Annual Survey of Above-Scale Industrial Firms (1998-2008).

Other county characteristics

We construct county socioeconomic variables (X_{jt} in the regression model) from the China Regional Yearbook, including GDP,[15] population, land area and agricultural share of GDP of each county. To control for nominal price effects, all monetary variables are adjusted to 1998 yuan using province-specific GDP deflators. To measure the proximity of a county to the social and economic center of the province, we calculate the spherical distance from the county's center to the provincial capital. To measure the ease of land transportation, we calculate the distance from a county's center to national highways using annual maps of national highways in China.

Table 1 provides summary statistics for the variables used in our analysis. There are on average 3.20 firms per industry per county per year, generating a total value added of 67.61 million yuan. For the subsample in which new firms enter a county group, on average 1.58 new firms appear per industry per county per year. On average, a county in our sample has 400,000 people, an annual GDP of 3.2 billion yuan, and a land area of about 2600 km^2. The average agricultural share of GDP is about 25 percent, the average distance from a county's center to the capital of its province is 223 km, and the average distance between the county center to the nearest national highway is 54 km.[16] It is evident from Table 1 that the counties in our sample exhibit great variations in almost every aspect.[17]

Estimation results

DID analysis

To investigate whether there is a "polluting thy neighbor" phenomenon in our context, one direct approach is to compare the water-polluting activities and firm location choices among river-side counties *within a province*. If water-polluting activities and new entry into water-polluting industries are more concentrated in the most downstream county of a province, then there is the possibility that the provincial governments strategically reduce the enforcement of environmental protection in the most downstream counties without regard to the negative externalities that impact their downstream neighbor provinces.

To implement this test, we run a DID regression on water-polluting and non-water-polluting industries in the most downstream (type-A) counties, the interior counties, and the most upstream (type-*B*) counties within a province. For each major river, we identify the complete sequence of riverside counties in each province that the river goes through. In total,

[15] In our regressions, we include the GDP of the previous year in the set of control variables to partially address the potential endogeneity problem of current-year GDP.

[16] The national highway system in China has grown at a remarkable speed during our sample period, from only 4800 km in total length in 1998 to 60,300 km in 2008. This results in large variations for the distance from county center to national highways (with a standard deviation of 99 km).

[17] As a robustness check, we drop observations of extreme values from our sample, and our results are not affected at all.

Table 1
Summary statistics.

	Mean	St.Dev.	Min	Max
Dependent variables				
Value added per industry per county (million yuan)	67.61	481.06	0.00	36,689.00
Log (value added per industry per county)	1.63	1.97	0.00	10.51
Number of firms per industry per county	3.20	10.29	0.00	425.00
Number of new firms per industry per county	1.58	4.35	0.00	150.00
Variables of interest				
Type-A/a counties	0.50	–	0.00	1.00
Type-B/b counties	0.50	–	0.00	1.00
Water-polluting industries	0.50	–	–	–
Riverside counties	0.25	0.43	0.00	1.00
Control variables				
Previous-year GDP (million yuan)	3185.57	4492.02	30.31	64,209.42
Log (previous-year GDP)	7.30	1.42	3.41	11.07
Population (10,000 persons)	40.09	35.82	1.00	213.45
Log (population)	2.93	1.63	0.00	5.36
Agricultural share of GDP (%)	24.67	16.57	0.00	89.83
Area (square kilometers)	2646.82	3085.06	83.74	37,146.22
Log (area)	7.53	0.84	4.43	10.52
Distance to provincial capital (kilometers)	223.05	167.81	8.18	1395.48
Log (distance to provincial capital)	5.14	0.79	2.10	7.24
Distance to highway (kilometers)	54.58	99.11	0.01	737.85
Log (distance to highway)	2.81	1.84	−4.68	6.60

Note: $N=34{,}128$. All monetary variables are deflated to 1998 yuan using provincial GDP deflators.

there are 348 riverside interior counties, which, along with the 116 riverside border counties, making up a 464-county sample with 69,226 observations.

The DID results are reported in odd-numbered columns of Table 2. Column (1) shows that, compared with interior counties, both type-*A* and type-*B* counties produce substantially less industrial value added. This reflects the general underdevelopment of all industries in counties at provincial borders. However, compared with interior counties of a province, water-polluting industries in type-*A* counties have 2.6 log points (2.6 percent) more value added than non-water-polluting industries,[18] whereas water-polluting industries in type-*B* counties generate 14.7 log points (13.7 percent) less value added than non-water-polluting industries. This finding provides a preliminary suggestion of the downstream effect that we are trying to identify. We find similar patterns when examining the number of firms in Column (3) and the number of new firms in Column (5).[19]

The within-province DID comparison provides evidence for the concentration of water-polluting production activities and new entry into water-polluting industries in the most downstream county of a province. But one question naturally arises: is this driven by unobservable county characteristics? Even within a province, the most upstream, interior, and downstream counties along a major river can be quite far apart and can differ greatly in many aspects. For example, the concentration of water-polluting production activities in the most downstream county of a province may simply reflect the coast-bound trend in the level of industrialization in China.

To mitigate this potential bias, we use the DID comparison between water-polluting and non-water-polluting industries in pairs of adjacent riverside counties on different sides of provincial borders. Type-*A* and type-*B* counties in such a pair are immediate neighbors of each other, sharing many geographic and other characteristics. The even-numbered columns of Table 2 summarize the cross-border DID results on the paired sample. Note that after controlling for county socioeconomic characteristics, county pair fixed effects and year fixed effects, water-polluting industries in type-*A* counties generate about 16.8 log points (18.3 percent) more value added than non-water-polluting industries when compared with type-*B* counties. Results for the number of firms and the number of new firms show very similar patterns. Thus, the cross-border DID analysis provides another piece of evidence supporting the downstream effect.

Note that in these regressions we control for county characteristics as well as time-variant county group and two-digit industry effects. The control variables for county characteristics have the expected partial effects. Naturally, counties with higher GDP in the previous year and counties less dependent on agricultural production have significantly higher value added per industry. Conditional on other variables, especially previous-year GDP, counties that are more populous (hence

[18] Although the particular result here is not statistically significant, Columns (3) and (5) show very strong and much more significant results that reveal the same pattern.

[19] Poisson models are used when the dependent variable is the number of firms or the number of new firms. We have also run a DID regression on *all* riverside counties to have an overview of the continuous pattern of polluting activities along the rivers. The results, presented in Table A3 in the appendix, show that the value added of water-polluting industries grows by 0.3 percent *more* than non-polluting industries as they move one county down the stream.

Table 2
Within-province and cross-border DID comparison.

	Log (value added)		*Number of firms*		*Number of new firms*	
	(1) Within-province	(2) Cross-border	(3) Within-province	(4) Cross-border	(5) Within-province	(6) Cross-border
(Type-A) · POL	0.026	0.168***	0.155***	0.150**	0.235***	0.211**
	(0.034)	(0.055)	(0.055)	(0.073)	(0.070)	(0.101)
(Type-A)	−0.285***	−0.066	−0.038	0.020	−0.247***	−0.177**
	(0.032)	(0.047)	(0.054)	(0.068)	(0.060)	(0.071)
(Type-B) · POL	−0.147***	0.052	0.067	(0.036)	(0.044)	(0.060)
(Type-B)	−0.466***	−0.558***	−0.484***	(0.052)	(0.031)	(0.057)
Log (previous-year GDP)	0.214***	0.266***	0.192***	0.263***	0.228***	0.275***
	(0.016)	(0.030)	(0.022)	(0.039)	(0.019)	(0.044)
Log (population)	−0.115***	−0.062**	−0.084***	−0.047	0.008	0.057
	(0.011)	(0.027)	(0.015)	(0.040)	(0.013)	(0.039)
Agricultural share of GDP	−0.017***	−0.033***	−0.025***	−0.040***	−0.018***	−0.041***
	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.005)
Log (area)	0.104***	0.436***	0.121***	0.486***	0.128***	0.431***
	(0.015)	(0.046)	(0.021)	(0.047)	(0.023)	(0.055)
Log (distance to provincial capital)	−0.120***	−0.333***	0.001	−0.162***	0.061***	0.038
	(0.019)	(0.062)	(0.023)	(0.060)	(0.022)	(0.070)
Log (distance to highway)	−0.114***	−0.143***	−0.072***	−0.174***	−0.070***	−0.258***
	(0.011)	(0.026)	(0.014)	(0.033)	(0.014)	(0.043)
R^2	0.378	0.393				
No. of Obs.	69,226	16,656	69,226	16,656	35,325	8796

Notes: An observation is an industry–county–year combination. The constant term is included but not reported. A full set of two-digit industry effects and province effects (border county pair effects) interacted with year effects, are also included for Columns (1), (3), (5) ((2), (4), (6)). Standard errors clustered at county-year level are reported in parentheses.

** Significant at 5 percent.

*** Significant at 1 percent.

Table 3
Cross-border DDD regression.

	Log (value added)			*Number of firms*			*Number of new firms*		
	(1) Riverside counties	(2) Non-Riverside counties	(3) All counties	(4) Riverside counties	(5) Non-Riverside counties	(6) All counties	(7) Riverside counties	(8) Non-Riverside counties	(9) All counties
(Type-A/a) · POL · RIV			0.184**			0.177			0.324**
			(0.082)			(0.112)			(0.143)
(Type-A/a) · POL	0.168***	0.029	0.013	0.150**	−0.010	−0.024	0.211**	−0.061	−0.106
	(0.055)	(0.061)	(0.060)	(0.073)	(0.086)	(0.085)	(0.101)	(0.101)	(0.104)
POL · RIV			−0.284***			−0.055			0.008
			(0.061)			(0.060)			(0.079)
(Type-A/a) · RIV			0.266***			0.369***			0.109
			(0.076)			(0.142)			(0.132)
(Type-A/a)	−0.066	−0.374***	−0.348***	0.020	−0.433***	−0.253***	−0.177**	−0.569***	−0.272***
	(0.047)	(0.048)	(0.055)	(0.068)	(0.060)	(0.098)	(0.071)	(0.082)	(0.092)
RIV			−0.200***			−0.354***			−0.174**
			(0.049)			(0.080)			(0.075)
R^2	0.393	0.423	0.371						
No. of Obs.	16,656	16,656	33,312	16,656	16,656	33,312	8796	8796	17,592

Notes: An observation is an industry–county–year combination. Log previous-year GDP, log population, agricultural share of GDP, log county area, log distance to provincial capital, log distance to highway and the constant term are included but not reported. Columns (1), (2), (4), (5), (7) and (8) all include a full set of two-digit industry effects and border county pair effects interacted with year effects. Columns (3), (6) and (9) include the set of county group effects and two-digit industry effects, both interacted with year effects. Standard errors clustered at county-year level are reported in parentheses.

** Significant at 5 percent.

*** Significant at 1 percent.

having fewer per capita resources) produce less industrial value added. On the other hand, more spacious counties produce more value added, probably reflecting the importance of land scarcity and land cost on industrial activities. Counties farther away from the provincial capital, which is usually the most prosperous city in a province, have lower value added. Likewise, counties lacking easy access to national highways also tend to produce less value added per industry, indicating the importance of transportation costs in firm location and industrial activities.

As discussed in section "Empirical strategy", the cross-border DID approach has its limitations too. Hence we use the DDD analysis to further control for confounding factors and better identify the downstream effect.

Baseline model estimations

Table 3 presents our main estimation results for the three dependent variables, the log value added (Columns (1)–(3)), the number of firms (Columns (4)–(6)), and the number of new firms (Columns (7)–(9)). Column (1) reports the cross-border DID estimates, using non-water-polluting industries as the control group for polluting ones. This is the same specification as the one in the cross-border DID analysis. Column (2) then shows the DID estimation results on the non-riverside sample with the same sets of control variables. Contrary to its riverside counterparts, the estimated coefficient on the interaction term is indistinguishable from zero, indicating that there is no excess increase (or decrease) in the value added of water-polluting industries in type-*a* counties relative to type-*b* counties.

Column (3) reports the baseline results using DDD, with the same set of control variables as in previous specifications, except that we now control for time-variant county group effects instead of county pair effects. More importantly, we also include dummy variables and their interaction terms that are of primary interest to us, RIV, DOWN · RIV, POL · RIV, and DOWN · POL · RIV. As the focus of our investigation, we obtain a significantly positive estimate on the triple interaction term. Specifically, we find that type-*A* counties produce 18.4 log points (20.2 percent) more value added per water-polluting industry than type-*B* counties, holding other things constant.

Columns (4) through (6) present the DID and DDD estimates when the dependent variable is the number of firms. The DDD estimate on the triple interaction term is significant both statistically and economically. Water-polluting industries in type-*A* counties have on average 0.18 more firms than in type-*B* counties, a 6 percent increase relative to the sample mean of 3.2 firms.[20] Thus, there are both more production activities and more firms in water-polluting industries in type-*A* counties than in type-*B* counties, after controlling for county characteristics and time-varying county group and industry unobserved effects.

We also examine firm location choices, using the number of newly established firms as the dependent variable. The results are reported in Columns (7) through (9). There are 0.32 more new firms per water-polluting industry in type-*A* counties than in type-*B* counties, or 20 percent of the sample mean of 1.58 firms. This indicates that, all else being equal, type-*A* counties attract more water-polluting firms than type-*B* counties. It is not surprising that the downstream effect as measured by the number of new firms is larger than that measured by the total number of firms, where results are attenuated by the existing firms established before the provincial governments responded to the 2001 policy.

Overall, the DDD results show that water-polluting industries have more production activities (in terms of value added and the number of firms) and more entries (in terms of the number of new firms) in type-*A* counties than in type-*B* counties, after controlling for possible confounding factors, county characteristics and time trends.

Mechanism of strategic polluting

Theoretical analysis

Our empirical analysis so far reveals that, *ceteris paribus*, water-polluting activities and new entry into water-polluting industries are higher in the most downstream county of each province. In this section we try to examine the underlying mechanism behind these findings. We now propose the following mechanism and then test its implications.

The mechanism for the downstream effect relies on two crucial factors, the inherent externalities of river pollution and the differential allocation of enforcement efforts by provincial governments. First off, given the unidirectional externalities of river pollution, the benefit of pollution control within a province is decreasing along the river. In the most downstream county, the gain from keeping the river clean is mainly enjoyed by the downstream neighbor province, making it the least beneficial place in the province to reduce pollution.

At the same time, the provincial government, under the pressure to meet pollution reduction mandates, has both the incentive and the freedom to optimally allocate its enforcement efforts among its counties. The incentive is directly from the mandates. The freedom comes from the institutional setup discussed in section "Institutional background", where the provincial government can affect the stringency of enforcement via the estimation of pollutant discharges, the collection of

[20] The estimate is marginally significant with a *p*-value of 0.114. However, a recent study by Abadie et al. shows that when the "sample" is the entire population, conventional methods usually overestimate standard errors because there is essentially no sampling error (Abadie et al., 2014). In our case, the population is counties along major rivers in China (as the mandates only apply to the major rivers), which is exactly the "sample" we use. Hence what we report are in fact upper bounds of standard errors, which underestimate the significance of estimated coefficients.

pollution fees, the inspection of violations, etc., all of which affect firms' production cost, and therefore the distribution of polluting activities.

The above two factors jointly lead to the provincial government allocating less effort in more downstream counties, and the least effort in the most downstream county. Then, other things being equal, water-polluting firms are more inclined to locate or expand production in downstream counties where enforcements are more lenient.[21] As a result, the most downstream county in a province will attract more polluting activities than an otherwise identical county.

Note that both the externalities and the differential enforcement are crucial in generating the downstream effect. Without the unidirectional externalities of river pollution, the optimal allocation of provincial governments' enforcement effort may not imply the least effort in the most downstream county. Equally importantly, without the variation in the stringency of enforcement, externalities per se cannot induce profit-maximizing firms to locate in the most downstream counties.

In this simple model, we should expect the following hypotheses to hold.

(H1) Enforcement will be the most lenient in the most downstream county of a province.
(H2) Water-polluting production activities and new entry into water-polluting industries will be the greatest in the most downstream county of a province.
(H3) Firms that are less cost-sensitive or more environmentally conscious will be less affected by regional differences in environmental protection enforcement, thus will contribute less to the downstream effect.
(H4) Without the pressure from the pollution reduction mandates, water-polluting production activities and new entry into water-polluting industries will vary less among riverside counties.

(H1), (H2) and (H3) are straightforward from the above discussion. (H4) results from the necessity of differential enforcement effort in generating the downstream effect, which we will discuss in detail in section “Before and after the 2001 policy”. Our empirical results in the preceding section provide strong evidence supporting (H2). Now we turn to the other hypotheses.

Pollution fees as a proxy for enforcement efforts

We use information on pollution fees to test (H1) about the lenient enforcement in the most downstream county of a province. The National Bureau of Statistics in China conducted the Census of Industrial Firms in 2005, which covered all industrial firms (including small firms with revenues below 5 million yuan) and had more information than the Annual Surveys of Above-Scale Industrial Firms. In particular, it included information about pollution fees firms paid in 2004. Firms in China that discharge water, air, and solid waste pollutants are required by law to pay pollution fees (State Council, 2003). For each pollutant type, pollution fees are usually proportional to the quantity of discharges, which is monitored by local BEPs. The collection of pollution fees clearly depends on the stringency of environmental regulation enforcement by local governments, thus pollution fees are a good proxy for the enforcement efforts.[22]

To test (H1), we use the within-province DID model to investigate how the collection of pollution fees varies in counties along a major river within a province. Since we have only cross sectional data, we run regressions of the natural log of pollution fees paid by 25,642 individual firms located in riverside counties instead of aggregating them to the industry level. The explanatory variables of primary interest are the location dummies, the water-polluting industry dummy, and their interaction terms, especially POL · DOWN. Moreover, we include in the right-hand side measures of firm production activities, namely the log of firm value added, of firm assets, and of the number of employees. This is because low pollution fees per se can be the result of lax enforcement or less pollution, or both. By controlling for firm production levels, which are proxies for the level of pollution, we could better identify the effect of lax enforcement in the most downstream counties, if there is any. As in previous regressions of the paper, we also include the full set of two-digit industry fixed effects and province fixed effects.

Table 4 presents the results. Most importantly, the estimated coefficient for the interaction term, POL · DOWN, is significantly negative. This shows that, *ceteris paribus*, water-polluting firms located in the most downstream county of a province on average pay 11.5 percent less pollution fees than those located in the interior counties. Note that water-polluting firms in the most upstream county do not pay significantly less for their pollution than those located in the interior counties. Note also that we are controlling for firm size (total assets and the number of employees) and firm production activities (firm value added). Thus the above results are already conditional on a given level of firm polluting activities. This finding provides evidence of lax enforcement efforts in the most downstream county of a province, which is consistent with our hypothesis.

[21] See Carlton (1983) for a model of firm location and size choices that spells out the effects of factor prices.

[22] Holding constant firms' polluting activities, pollution fees are clearly monotonic in enforcement efforts. However, firms' polluting activities are likely to decrease as enforcement efforts increase. But even so it would be paradoxical if overall pollution fees decreased as enforcement efforts increase.

Table 4
Pollution fee in 2004: within-province comparison.

	Log (pollution fee)
(Type-A)·POL	−0.115**
	(0.054)
(Type-A)	−0.081**
	(0.038)
(Type-B)·POL	−0.019
	(0.064)
(Type-B)	−0.124***
	(0.039)
Log(firm value added)	0.177***
	(0.013)
Log(firm asset)	0.060***
	(0.009)
Log(firm employee)	0.260***
	(0.012)
R^2	0.218
No. of Obs.	25,642

Notes: An observation is a firm in 2004. Log previous-year GDP, log population, agricultural share of GDP, log county area, log distance to provincial capital, log distance to highway and the constant term are included but not reported. A full set of two-digit industry fixed effects and province fixed effects is included. Heteroskedasticity robust standard errors are reported in parentheses.

** Significant at 5 percent.

*** Significant at 1 percent.

Table 5
Falsification test: interior county groups.

	Log (value added) (1)	*Number of firms* (2)	*Number of new firms* (3)
(Type-A/a)·POL·RIV	0.075	−0.159*	−0.140
	(0.077)	(0.093)	(0.116)
(Type-A/a)·POL	−0.014	−0.049	−0.141**
	(0.056)	(0.055)	(0.071)
POL·RIV	−0.072*	0.040	−0.015
	(0.043)	(0.066)	(0.080)
(Type-A/a)·RIV	−0.054	0.140**	0.256***
	(0.044)	(0.059)	(0.074)
(Type-A/a)	0.011	−0.040	−0.081
	(0.034)	(0.038)	(0.051)
RIV	0.082***	−0.004	−0.045
	(0.022)	(0.029)	(0.039)
Log (previous-year GDP)	0.667***	0.559***	0.415***
	(0.032)	(0.030)	(0.037)
Log (population)	−0.050***	−0.023**	−0.017
	(0.011)	(0.012)	(0.015)
Agricultural share of GDP	−0.009***	−0.009***	−0.008***
	(0.001)	(0.002)	(0.002)
Log (area)	0.114***	0.119***	0.229***
	(0.030)	(0.026)	(0.039)
Log (distance to provincial capital)	−0.213***	−0.011	−0.035
	(0.035)	(0.034)	(0.047)
Log (distance to highway)	−0.120***	−0.043***	−0.081***
	(0.009)	(0.009)	(0.015)
R^2	0.415		
No. of Obs.	50,163	50,163	28,830

Notes: An observation is an industry–county–year combination. The constant term is included but not reported. A full set of county group effects and two-digit industry effects, both interacted with year effects, is also included. Standard errors clustered at county-year level are reported in parentheses.

* Significant at 10 percent.

** Significant at 5 percent.

*** Significant at 1 percent.

Table 6
Cross-border DDD regression by types of firm ownership.

	Log (value added)			*Number of firms*			*Number of new firms*		
	(1) SOE	(2) POE	(3) Foreign	(4) SOE	(5) POE	(6) Foreign	(7) SOE	(8) POE	(9) Foreign
(Type-A/a) · POL · RIV	0.054	0.207**	−0.059	−0.031	0.209*	0.009	0.926	0.278*	0.442*
	(0.045)	(0.081)	(0.047)	(0.120)	(0.113)	(0.162)	(0.664)	(0.158)	(0.226)
(Type-A/a) · POL	0.037	−0.076	0.018	0.156*	0.057	−0.171**	−0.502	0.021	−0.368***
	(0.031)	(0.056)	(0.033)	(0.085)	(0.067)	(0.071)	(0.498)	(0.090)	(0.098)
POL · RIV	−0.015	−0.359***	0.059*	0.032	−0.153**	0.053	−0.581	−0.085	−0.041
	(0.033)	(0.061)	(0.035)	(0.087)	(0.070)	(0.093)	(0.492)	(0.090)	(0.171)
(Type-A/a) · RIV	0.026	0.123**	0.121**	0.005	0.322***	−0.122	0.316	0.392**	−0.039
	(0.029)	(0.057)	(0.061)	(0.096)	(0.097)	(0.163)	(0.370)	(0.163)	(0.311)
(Type-A/a)	−0.026	−0.246***	−0.021	0.088	−0.440***	0.362**	−0.119	−0.519***	0.436**
	(0.020)	(0.042)	(0.042)	(0.066)	(0.062)	(0.152)	(0.276)	(0.102)	(0.193)
RIV	0.020	−0.048	−0.063	0.078	−0.061	0.050	−0.032	−0.185*	−0.495**
	(0.020)	(0.040)	(0.044)	(0.065)	(0.057)	(0.124)	(0.262)	(0.101)	(0.242)
R^2	0.255	0.320	0.396						
No. of Obs.	26,487	26,575	26,533	26,487	26,575	26,533	14,475	14,589	14,515

Notes: An observation is an industry–county–year combination. Log previous-year GDP, log population, agricultural share of GDP, log county area, log distance to provincial capital, log distance to highway and the constant term are included but not reported. A full set of county group effects and two-digit industry effects, both interacted with year effects, is also included. Standard errors clustered at county-year level are reported in parentheses.

* Significant at 10 percent.

** Significant at 5 percent.

*** Significant at 1 percent.

Interior county groups

The DDD analysis on county groups at provincial borders in section "Estimation results" provides direct evidence in support of (H2). To further examine whether the downstream effect is the result of being at provincial borders, we perform the same DDD analysis on interior counties as a falsification test.

Specifically, we define an interior county group by identifying four adjacent interior counties located at the center of each province, two of which are located along a major river and two of which are not.[23] The sample consists of 102 riverside counties and 102 non-riverside ones, or 51 county groups, where all counties are interior.

We then estimate exactly the same DDD model on this new sample. This is in effect adding a fourth dimension of difference, using interior county groups as the control for county groups on the provincial border.

The estimation results are reported in Table 5. The three columns examine the log value added, the number of firms, and the number of new firms with a full set of control variables and fixed effects as those in Table 3. Contrary to the results for county groups at provincial borders, we do not see the same downstream effect for any of the three regressions. That is, there is no increase in water-polluting production activities or entry into water-polluting industries in "type-A" counties of an *interior* county group. This is in sharp contrast with what we find for counties on provincial borders, and provides further support to (H2).

Firm heterogeneity in ownership

To test (H3), we take advantage of the heterogeneity in firm ownership in our sample. We divide the whole sample into three subsamples by ownership type: state-owned enterprises (SOE), private firms and foreign firms. Compared with private firms, SOEs are less cost-sensitive, either because of their inefficiency in operation or because they have social goals other than maximizing profits (e.g., improving local employment). Foreign firms are cost-sensitive, but are often bounded by more stringent environmental regulations in their home countries, so are more environmentally conscious than private firms. At the same time, they have less access to the Chinese administrative network, and may benefit less than their local competitors from the lax enforcement of environmental policies. Therefore by (H3), SOEs and foreign firms are less attracted to the most downstream county of a province than private firms.

For each of the SOE, private, and foreign firm samples, we implement the same DDD regression as that in our baseline model using the natural log of value added per industry per county per year as the dependent variable. The results are shown in Columns (1)–(3) of Table 6. The estimate on the triple interaction term, $\text{POL}_i \cdot \text{DOWN}_j \cdot \text{RIV}_j$, is not distinguishable from zero for SOEs and is slightly negative for foreign firms. On the other hand, the private firm sample produces a significantly positive estimate of 20.7 log points (23.0 percent). Similarly, Columns (4)–(9) show similar results for the number of firms and the number of new firms. The only irregularity arises in Column (9), where foreign firms also display some downstream effects. This, however, is mainly driven by highly skewed distribution of the number of new foreign firms. More than 90 percent of our observations have zero entry of foreign firms, and the skewness is over 16. The limited variation in the dependent variable may have been the cause of this unusually large coefficient on the number of new firms. Overall, Table 6 suggests that SOEs and foreign firms contribute less to the downstream effect than private firms. Private firms are the most sensitive to the provincial governments' differential allocation of enforcement efforts, thus they are the main contributor to the downstream effect.

Before and after the 2001 policy

As (H4) predicts, we expect to see more downstream effects since 2001 but not so much before 2001. The 2001 pollution reduction mandates triggered differential allocation of enforcement efforts by the provincial governments among their counties. Prior to 2001, however, there was no pressure from the central government to reduce pollution. Hence the provincial governments could exert any amount of effort they wanted, which was likely to be little, as they were primarily interested in spurring economic growth to gain the favor of the central government and to be promoted (Zhou, 2008).[24] With limited *total* enforcement effort, there would be limited *variation* in the allocation of enforcement efforts within a province as a result. This small variation implies that we should not expect to see as strong a downstream effect before 2001.

To test (H4), we pool our samples before and after 2001 and run a quadruple difference regression (DDDD), where the additional dimension of difference is between pre- and post-2001. The following regression equation summarizes this test:

$$\begin{aligned} Y_{ijt} = {} & \gamma_0 + \gamma_1 \text{AFTER}_t \cdot \text{POL}_i \cdot \text{DOWN}_j \cdot \text{RIV}_j + \gamma_2 \text{POL}_i \cdot \text{DOWN}_j \cdot \text{RIV}_j \\ & + \gamma_3 \text{AFTER}_t \cdot D_{ijt} + \gamma_4 D_{ijt} + \gamma_5 X_{jt} + \delta_{Jt} + \eta_{it} + \varepsilon_{ijt} \end{aligned}$$

[23] Along a major river within a province, we construct interior county groups in three different ways: (i) the most upstream *interior* county group; (ii) the most downstream *interior* county group; and (iii) the interior county group at the center. We present the result using definition (iii), but the other two definitions yield the same result.

[24] Citizens in the province, on the other hand, had virtually no influence on the promotion of provincial government officials (Zhou, 2008).

Table 7
Before and after the Tenth Five-Year Plan.

	Log (value added) (1)	*Number of firms* (2)	*Number of new firms* (3)
(Type-A/a) · POL · RIV · AFTER	0.137	0.108	0.300***
	(0.130)	(0.170)	(0.023)
(Type-A/a) · POL · RIV	0.050	0.070	0.037***
	(0.101)	(0.128)	(0.007)
(Type-A/a) · POL · AFTER	0.013	−0.109	−0.217***
	(0.093)	(0.111)	(0.006)
(Type-A/a) · POL	0.005	0.085	0.100***
	(0.071)	(0.071)	(0.022)
POL · RIV · AFTER	−0.196**	0.122	−0.135***
	(0.098)	(0.095)	(0.013)
POL · RIV	−0.093	−0.177**	0.139***
	(0.076)	(0.074)	(0.008)
(Type-A/a) · RIV · AFTER	0.559***	0.919***	−0.486***
	(0.140)	(0.228)	(0.046)
(Type-A/a) · RIV	−0.260**	−0.535***	0.576***
	(0.117)	(0.177)	(0.010)
R^2	0.364		
No. of Obs.	45,449	45,449	22,998

Notes: An observation is an industry–county–year combination. Dummy variables (Type-A/a) and RIV, and their interaction terms with AFTER, as well as log previous-year GDP, log population, agricultural share of GDP, log county area, log distance to provincial capital, log distance to highway and the constant term are included but not reported. A full set of county group effects and two-digit industry effects, both interacted with year effects, is also included. Standard errors clustered at county-year level are reported in parentheses.

** Significant at 5 percent.

*** Significant at 1 percent.

where AFTER_t takes the value of 1 if and only if $t \geq 2001$. D_{ijt} is the vector of county and industry dummy variables and their interactions in the baseline DDD. The other components of the model is the same as those in section "Empirical strategy".[25]

Table 7 summarizes the estimation results for the log of value added, the number of firms, and the number of new firms. The estimated coefficients for the quadruple interaction term are indeed positive and larger than those on $\text{POL}_i \cdot \text{DOWN}_j \cdot \text{RIV}_j$.[26] We do find, however, that the two coefficients in Column (2) are comparable in sizes and have overlapping confidence intervals. We conjecture that it might be because the total number of firms captures both the strategic response of the provincial governments to the 2001 policy and the preexisting conditions of the industries. Thus the before–after difference is attenuated. Nonetheless, we still find some *increase* in the downstream effects after 2001. Moreover, our estimation is likely to be a lower bound of the difference between pre- and post-mandate years. This is because there might have been anticipation of the 2001 environmental policy, which would induce differential allocation of enforcement efforts even before the policy was implemented. If such anticipation effects do exist, then the pre-2001 results may also display some weak evidence for the downstream effect.

In sum, the empirical evidence in this section support the four hypotheses. Evidence from testing (H1) suggests that the downstream effect we have identified is indeed due to strategic polluting by the provincial governments via the allocation of enforcement efforts. The falsification test on interior counties finds that the downstream effect is uniquely observed at the provincial borders, providing further evidence in support of (H2). Tests of (H3) extend the model prediction to heterogeneous firms, and suggest interesting patterns of the contribution to the downstream effect across different ownership types. Lastly, tests of (H4) find some evidence that the 2001 mandates increased the downstream effects by giving the provincial governments more pressure.

Concluding remarks

Using the DDD identification strategy, we identify the downstream effect and provide strong evidence of strategic polluting along major rivers in China. In particular, both water-polluting production activities and new entry into water-polluting industries are significantly higher in the most downstream county of a province, relative to otherwise similar counties. In investigating the mechanism behind the downstream effect, we find evidence that under the pressure from the central government, the provincial governments allocate their enforcement efforts so that the most downstream county has

[25] Alternative specifications where the X_{jt}'s are separately estimated over time yield similar patterns. In addition, replacing AFTER_t with individual year dummies, i.e. estimating the downstream effects year by year, shows that the downstream effects are roughly increasing over the years.

[26] As discussed earlier, the standard errors are overestimated. Yet still, the DDDD coefficients in Column (3) is strongly significant and sizable. The ones in Columns (1) and (2), although not statistically significant, still shed some light on the difference between downstream effects in pre- and post-mandate years.

the most lenient enforcement of environmental regulation. As a result, we observe a sharp increase in water-polluting activities in the most downstream counties.

Correction of the incentives for local governments poses a challenging policy problem. In a democratic society, local governments concerned with the welfare of voters in their own jurisdictions but not those in other jurisdictions also have strategic polluting incentives. Researchers have examined the efficacy of decentralization, transfer payments, and international trade in mitigating the free riding incentives for local governments, such as Sigman (2005), Bernauer and Kuhn (2010), and Lipscomb and Mobarak (2013). In China, local governments are not necessarily concerned with the welfare of their residents, but are required by the central government to maintain a certain level of environmental quality in their jurisdictions. Without taking into account local governments' strategic responses, the increasing pressure from the central government to protect the environment may increase local governments' incentive to environmentally free ride and distort the distribution of polluting activities. Direct monitoring by the central authority, especially in counties at provincial borders, may keep the free-riding incentives in check. Incorporating the feedback from downstream provinces when evaluating local government officials may also reduce the incentives to engage in strategic polluting.[27]

Acknowledgments

We thank the editor, two anonymous reviewers, Xu Cheng, Hanming Fang, and participants at the Thirteenth NBER-CCER Annual Conference and Peking University for helpful discussions and comments.

Appendix A. Supplementary data

Supplementary data associated with this paper can be found in the online version at http://dx.doi.org/10.1016/j.jeem.2015.01.002.

References

Abadie, A., Athey, S., Imbens, G.W., Wooldridge, J.M., 2014. Finite Population Causal Standard Errors. Mimeo.
Becker, R., Henderson, V., 2000. Effects of air quality regulations on polluting industries. J. Polit. Econ. 108 (2), 379–421.
Bernauer, T., Kuhn, P.M., 2010. Is there an environmental version of the Kantian Peace? Insights from water pollution in Europe. Eur. J. Int. Relations XX (X), 1–26.
Bureau of Environmental Protection, 2002. Implementing the Tenth Five-Year Plan for Environmental Protection. Bureau of Environmental Protection, The People's Republic of China.
Carlton, D.W., 1983. The location and employment choices of new firms: an econometric model with discrete and continuous endogenous variables. Rev. Econ. Stat. 65 (3), 440–449.
Chan, C.K., Yao, X., 2008. Air pollution in mega cities in China. Atmos. Environ. 42 (1), 1–42.
Chinese Academy of Environmental Planning, 2006. Report on the Completion of the Tenth Five-Year Plan for Environmental Protection. Ministry of Environmental Protection of the People's Republic of China Document 2006-00172.
Deng, X., 1992. Talks in Wuchang, Shenzhen, Zhuhai, and Shanghai. News of the Communist Party of China Online.
Gray, W.B., Shadbegian, R.J., 2004. 'Optimal' pollution abatement: whose benefits matter, and how much?. J. Environ. Econ. Manag. 47, 510–534.
Greenstone, M., 2002. The impacts of environmental regulations on industrial activity: evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures. J. Polit. Econ. 110 (6), 1175–1219.
Helland, E., Whitford, A.B., 2003. Pollution incidence and political jurisdiction: evidence from the TRI. J. Environ. Econ. Manag. 46, 403–424.
Kahn, M.E., 2004. Domestic pollution havens: evidence from cancer deaths in border counties. J. Urban Econ. 56, 51–69.
Kahn, M.E., Mansur, E.T., 2013. Do local energy prices and regulation affect the geographic concentration of employment? J. Public Econ. 101, 105–114.
Konisky, D.M., Woods, N.D., 2012. Environmental free riding in state water pollution enforcement. State Polit. Policy Q. 12 (3), 227–251.
Lipscomb, M., Mobarak, A.M., 2013. Decentralization and Pollution Spillovers: Evidence from the Re-drawing of County Borders in Brazil. Unpublished Manuscript, University of Virginia and Yale University.
Liu, S., 2006. Who is Dragging Environmental Protection Down? Behind Unmet Environmental Targets of the Tenth Five-Year Plan. China Youth Daily, April 13, 2006. Retrieved from ⟨http://env.people.com.cn/GB/1072/4295135.html⟩.
Ministry of Environmental Protection, 2010. Report on the First National Census of Polluting Sources. Ministry of Environmental Protection, the People's Republic of China.
National Development and Reform Committee, January 2002. The Tenth Five-Year Plan for Environmental Protection. China Environmental Sciences Press, Beijing.
National Planning Committee, 2001. Notice on the cancellation of some administrative examination and approval items. National Planning Committee Document [2001]2440.
Naughton, B., 2005. The new common economic program: China's Eleventh Five Year Plan and What it Means. China Leadership Monitor, no. 16.
Sigman, H., 2002. International spillovers and water quality in rivers: do countries free ride?. Am. Econ. Rev. 92 (4), 1152–1159.
Sigman, H., 2005. Transboundary spillovers and decentralization of environmental policies. J. Environ. Econ. Manag. 50, 82–101.
State Council, 1984. Reply to National Planning Committee's Interim Provisions on Improving the Central Planning System. The People's Republic of China State Council Document [1984]138.
State Council, 2001. The Tenth Five-Year Plan for National Economic and Social Development of the People's Republic of China. The National People's Congress of the People's Republic of China Online.

[27] Inter-governmental collaboration across provinces is often difficult to implement. For example, a serious river pollution dispute between two counties at the Jiangsu–Zhejiang provincial border went on for 5 years (Zhai and Tong, 2006). By the time the central government finally intervened, the downstream county had already suffered great losses in fishery and built a dam to block the waterway in a tit-for-tat manner. To the best of our knowledge, there is no formal mechanism or anecdotal evidence that there is coordination among China's local governments in environmental protection.

State Council, 2003. Regulation of the Collection and Usage of Pollutant Discharge Fees. The People's Republic of China State Council Document [2003]369.
State Council, 2007. Approval of Provisions on the Evaluation, Inspection, and Assessment of Pollution Reduction and Energy Saving. The People's Republic of China State Council Document [2007] 36.
State Council, 2013. Report on the Work of the Government 1954–2013. The Central People's Government of the People's Republic of China Online.
vanRooij, B., 2010. The people vs. pollution: understanding citizen action against pollution in China. J. Contemp. China 19 (63), 55–77.
Wang, X., 2002. Explaining the Tenth Five-Year Plan for Environmental Protection. Bureau of Environmental Supervision, Ministry of Environmental Protection, The People's Republic of China.
Wen, J., May 2007. Report on the Work of Government. Report on the Fifth Session of the Tenth National People's Congress.
World Bank, 2006. China Water Quality Management: Policy and Institutional Considerations, Washington, DC.
World Bank, 2007. Cost of Pollution in China: Economic Estimates of Physical Damages, Washington, DC.
Wu, Z., Cheng, S., Fu, G., 2000. Human impacts on self-cleaning effects in freshwater ecosystems of the Yangtze River area. In: Ecosystem Service and Sustainable Watershed Management in North China International Conference Paper.
Zhai, M., Tong, J., 2006. Observation on a Five-Year Inter-Provincial River Pollution Dispute. Minjian Magazine, April 14, 2006. Retrieved from ⟨http://news.21cn.com/today/legend/2006/04/14/2544649.shtml.
Zhang, Z., 2012. Effective environmental protection in the context of government decentralization. Int. Econ. Econ. Policy 9 (1), 53–82.
Zhou, L.-A., 2008. Local Governments in Transition: Incentive and Governance of Officials in China. Shanghai People's Press, Shanghai.