Just Water? Environmental Jurisprudence, Water Quality and Infant Mortality in India*

Sandeep Bhupatiraju
World Bank

Daniel L. Chen
World Bank, Toulouse School of Economics

Shareen Joshi
Georgetown University

Peter Neis
Toulouse School of Economics

Shashank Singh
Indian Institute of Technology

PRELIMINARY DRAFT – PLEASE DO NOT CITE WITHOUT PERMISSION
December 21, 2021

Abstract

India's surface water displays high levels of toxicity. Policies to improve it have been inconsistently effective. We assemble a novel dataset that includes environmental court cases from the Indian Supreme Court, High Courts and Green Tribunal, assessments of the impacts of these cases on the environment, a comprehensive history of the rulings of judges who serve on these cases, surface-water pollution, and also mortality estimates from two recent demographic surveys. We use an instrumental variables framework to exploit the process of random judge assignment within the Indian justice system to estimate the effect of "green" rulings on pollution levels throughout India. We find that the rulings precipitated immediate reductions in surface water pollution. These effects however, are confined to the year of the ruling and do not show persistence. We find no statistically significant impact of the rulings on neonatal mortality or infant mortality. This is suggestive evidence that judicial policies can succeed in lowering short-term pollution, but this is insufficient to lower infant mortality in the districts of India.

JEL Codes: Q53, Q56, C36

Keywords: environmental justice, development, pollution, infant mortality.

*We are grateful to the World Bank and Georgetown University's Laudato Si grant opportunity for financial support. We are also grateful to Lechuan Qiu, Ritesh Das, Viknesh Nagarathinam, Shevani Tiwari, Shilpa Rao, Sunny Anand, Kasturi Kadalam, Praharsha Enagandula, Akhilesh Rajagopalan, Connor Brennan, Aanika Veedon and an outstanding team of law students from National University of Juridical Sciences in India for valuable research assistance in coding cases. We are also grateful to Quy Toan Do, Jishnu Das, Shibani Ghosh, Vikram Raghavan, Vasujith Ram, Vijayendra Rao, Martin Ravallion, Nicholas Robinson, Mrinal Satish, and Nayantara Vohra for very helpful advice, critical insights and very stimulating discussions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of the World Bank, its Board of Executive Directors, or the governments they represent.

1 Introduction

India is facing a water crisis (Government of India, 2018). It is estimated that 600 million Indians face high to extreme water stress and about 200,000 people die every year due to exposure to contaminated water (Global Alliance on Health and Pollution, 2019). Though popular narratives have often emphasized the roles of population growth, the rapid pace of urbanization and strong economic growth, weak environmental governance is also understood to play a role (Woodhouse and Muller, 2017).

India's water governance framework has some significant structural weaknesses. On the legal side, it lacks a formal system of rights over water and environmental resources more generally, and so no single entity bears the responsibility of protecting water resources, or shoulders the blame for failing to protect them (Ghosh, 2019). Though there are numerous formal laws and regulatory bodies on firms who pollute, these are poorly suited to deal with the growing challenge of water toxicity (Greenstone and Hanna, 2014).

Enforcement of environmental laws is complicated by a complex institutional structure. Though State Pollution Control Boards (SPCBs) are tasked with ensuring that firms comply with the relevant laws, they face chronic problems of under-staffing, inadequate budgets and coordination challenges with the Central Pollution Control Board (CPCB) in New Delhi (Mehta, 2019). Recent experiments illustrate that even outsourcing some of the most time-consuming tasks fail to improve performance, largely because of the deeper structural issues in the system (Duflo, Greenstone, Pande, and Ryan; 2013).

India's system of democratic politics is not conducive to reform. The millions of firms that are located in highly polluting industrial clusters provide valuable employment and valuable products for export (Joshi and Shambaugh, 2018). As in the case in many contexts across the world, there is very little political will to prioritize environmental quality at the cost of economic growth (Stiglitz, Sen and Fitoussi, 2010).

In this landscape of near-total gridlock in the legislative and executive systems, India's judiciary has taken an activist stance towards environmental conservation (Malleson, 2016; Ghosh, 2019). Over the past 30 years, the judiciary has passed some landmark verdicts on issues of water toxicity. The first of these – MC Mehta versus Union of India – successfully curtailed the levels of pollution in the Ganga river in 1987, and the drop in pollution has been shown to have lowered child mortality downstream (Do, Joshi and Stolper, 2018). This case as well as other well-known cases at the Supreme Court include *Vellore Citizens Welfare Forum v. Union of India*, *Subhash Kumar v. State of Bihar & Ors.* and *Samit Mehta v. Union of India & Ors.* More recently, with the establishment of the Green Tribunal in 2011, numerous cases related to water pollution have emerged from the judiciary.

What is the broad impact of judicial rulings on environmental outcomes? There is currently no empirical study that examines the overall impact of litigation on environmental toxicity in India. This paper fits into this gap. We curate cases related to water toxicity that have been heard at the high courts, supreme court and Green Tribunal of India between 1987 and the current time. We examine these cases and code assessments of "pro-environment" verdicts. We then link the profiles of judges who hear these cases. For each judge we construct a corpus of rulings that were written *prior* to the green verdict. Next, we analyze the causal relationship between a pro-environmental ruling and actual environmental outcomes. Since rulings may be endogenous to outcomes, we use an IV framework, with the textual features of judges' cases to predict the likelihood of a green verdict. Since judges are randomly assigned within the courts of India, these serve as an instrument capturing judges' writing style to predict the outcome of cases. Finally, we deploy the same instrumental framework to look at the effects of green rulings on mortality rates at the district-level.

We find that the rulings precipitated reductions in two specific measures of water quality – biological oxygen demand (BOD) and chemical oxygen demand (COD) – the most common measures of industrial pollution in surface water. These effects however, are confined to the year of the ruling and do not show persistence. We find no significant effect of green pro-environmental rulings on domestic river pollution (Fecal Coliforms and Total Coliforms.) We also find no significant impact of the rulings on neonatal mortality or infant mortality at the district-year level. We interpret this as suggestive evidence that judicial policies can succeed in lowering short-term pollution, but the many issues with enforcement and oversight that have already been documented in the literature inhibit their ability to bring real improvements in health at the grassroots of society, particularly for newborns who are most vulnerable to contaminated water.

The remainder of this paper is organized as follows. Section 2 presents some background information on environmental jurisprudence in India. Section 3 provides an overview of the many sources of data that were curated for this project. Section 4 presents empirical models, section 5 presents results. We discuss the implications of our results in section 6. The final section concludes.

2 Background

India has an extensive environmental management system that includes laws, statutory mandates, regulatory system, and institutional frameworks to implement and enforce environmental protections. There are currently more than 200 laws on the books (World Bank, 2013; UNDP, 2009). Several institutions are involved in the design and implementation of environmental policies. These include the Ministry of Environment and Forests (MoEF), the Central Pollution Control Board (CPCB), State Departments of Environment, State Pollution Control Boards (SPCBs) and Municipal Corporations. Below we summarize the main legal instruments, and the systems of enforcement for these laws with a focus on water management in the past four decades.

2.1 Water Laws

Water has been formally regulated in India since colonial times (Singhal, 2010).[1] In independent India, the responsibility for protecting most environmental resources is shared between the central and state governments, with the central government having responsibility for policy and regulatory formulations and the state governments for ensuring implementation and enforcement of national policies and laws.

One of the earliest set of regulations came from the Factories Act of 1948, which required every factory to make effective arrangements for the treatment of waste and effluent that emerged from the manufacturing process. The formalization of the rules and the implementation of this however, was left to the states. The River Boards Act of 1956 mandated the establishment of River Boards for the regulation and development of inter-state rivers and river valleys.[2] Section 133 of the Code of Criminal Procedure allow district magistrates to pass orders to address issues of public nuisance or unlawful obstructions of rivers or waterways.[3] The most significant post-colonial piece of legislation that addresses the issue of water pollution is the Water (Prevention and Control of Pollution) Act of 1974. The Act has 64 sections which establish the Central and State Pollution Control Boards, describe the powers of the Boards, specify the requirements for testing water at state laboratories, outline the measures that the Boards must take to prevent and control water pollution; and outline penalties and punishment procedures for the flouting of these rules.

[1] Many acts were imposed on India by the British colonial regime. Some examples, we found in our cursory research are as follows: (1) The Fair Ways Act of 1823 expanded the authority of the Central Government to regulate or prohibit the throwing of garbage into India's canals and waterways; (2) The Shore Nuisance (Bombay and Kolaba) Act of 1853 authorized the collector of land revenue to remove "nuisances and encroachments below high- water mark in the Islands of Bombay and Kolaba; (3) Section 277 of the Indian Penal Code, formalized in 1860, made it illegal and punishable to voluntarily corrupt or foul the water of "any public spring or reservoir, so as to render it less fit for the purpose for which it is ordinarily used; (4) The Sarais Act of 1867 imposed water standards on the keeper of a Sarai (a place that provides free drinking water) and imposed fines on violations of these standards; (5) The Drainage Act of 1873 prohibited any interference or alteration of the flow of water in rivers and streams; (6) The Indian Fisheries Act of 1897 made it a punishable offence to intentionally insert any substance into waterways that have an adverse effect on fish populations.

[2] https://indiankanoon.org/doc/1608688/

[3] https://indiankanoon.org/doc/983382/

While this act went a long way towards recognizing the issue of surface water pollution in India, it had some limitations. First, it left out some key sources of pollution. Groundwater, and non-point sources of water pollution such as agricultural runoff and water discharged from municipal sources are outside the act's purview. The act also limits the accountability of the pollution control boards by limiting the opportunities for citizens or companies to contest their actions in civil courts.[4]

The Water Act of 1974 was ultimately subsumed into broader environmental legislation. In 1986, the Government of India passed the Environment Act in the aftermath of the Bhopal Gas Disaster.[5]. The act intends to establish a framework for the central government to coordinate with various central and state authorities. The central government was given the power to appoint officers, give directives to firms to close, or impose penalties for non-compliance. It provides specific details on the handling of offences by companies, citizens or government agencies.

More recently, additional acts have been passed to address water pollution. These include the Municipal Solid Wastes (Management and Handling) Rules in the year 2000. This however, was replaced by Solid Waste Management Rules (SWM), 2016.

2.2 Enforcement Mechanisms

Under the Water Act, the Air Act and the Environmental Protection Act, the pollution control boards have a variety of policies at their disposal to ensure compliance and enforcement of environmental laws. They can issue and revoke consents to operate, require self-monitoring and reporting, conduct sampling, inspect facilities, require corrective action and prescribe compliance schedules. They cannot however, alter the guidelines or standards specified by the central government.

The principle tool of ensuring compliance with water laws in India is the process of inspections (Duflo, Greenstone, Pande, and Ryan, 2018). Section 21 of the Water Act and Section 26 of the Air Act empowers SPCB officials to enter the premises of polluters and take samples of effluents or emissions. The purpose of this system is largely to deter citizens from engaging in polluting activities (Abbot, 2009; Epple and Visscher, 1984).

In the context of India however, shortfalls in staffing and budges has curtailed the effectiveness of this system. In an experiment, Duflo et al. (2018) doubled the rate of inspection for treatment plants and required that the extra inspections be assigned randomly. They find that treatment plants only slightly increased compliance. This is largely due to poor targeting: random inspections in treated firms produced fewer violations than the regulator's own discretionary inspections. The authors demonstrate in a formal model that it is efficient for the regulator to aggressively target discretionary inspections to the dirtiest and most polluting firms, and provide only minimal inspections to the vast majority of firms. The lack of information on actual levels of polluting behavior stymies their efforts.

The systems of monitoring vary significantly across states. Gupta (1996) argues that states may deliberately pursue a policy of lax enforcement to attract investment. A recent World Bank report points out that the frequency of on-site visits to verify compliance is determined by the pollution potential (red/orange/green) and size (based on the value of capital investment) of the industry. Although CPCB has set its guidelines

[4]The act clearly says that "*No civil court shall have jurisdiction to entertain any suit or proceeding in respect of any matter which an appellate authority constituted under this Act is empowered by or under this Act to determine, and no injunction shall be granted by any court or other authority in respect of any action taken or to be taken in pursuance of any power conferred by or under this Act.*" This effectively means that actions taken by those who work for the pollution control board cannot be charged with offenses under the Act.

[5]As the introduction says, however, the act was "Where as the decisions were taken at the United Nations Conference on the Human Environment held at Stockholm in June, 1972, in which India participated, to take appropriate steps for the protection and improvement of human environment. Where as it is considered necessary further to implement the decisions aforesaid in so far as they relate to the protection and improvement of environment and the prevention of hazards to human beings, other living creatures, plants and property" (http://www.moef.nic.in/sites/default/files/eprotect_act_1986.pdf)

regarding the frequency of visits, individual states differ in their interpretation of this guidance (World Bank, 2013). For example, red category facilities are supposed to be inspected once a month in Gujarat, once per quarter in Orissa, and once every two years in West Bengal although the guidelines set by CPCB is once in three months for large and medium scale industries.

Besides inspections, the SPCBs also carry out training workshops for firms and other polluting entities. They are also authorized to issue notifications for hazardous wastes, bio-medical wastes, municipal solid wastes and electronic wastes in their respective States (World Bank, 2013).

2.3 Water Jurisprudence

Legislative and executive failures, together with the pressures of economic growth and population growth, necessitated involvement of the judiciary in issues related to the environment (Rajamani, 2007; Bhuwania, 2017; Mehta, 2019). In the aftermath of the political emergency of 1976, the judiciary took on this challenge through a renewed commitment to protecting citizen's fundamental rights (Dias, 1994; Bhuwania, 2017). The commitment to environmental jurisprudence intensified after the massive Bhopal gas disaster (Abraham and Abraham, 1991; Dias, 1994). The landmark cases pertinent to water pollution, particularly *MC Mehta versus Union of India*, emerged at this time.

Faced with environmental laws that were deficient in the coverage, compliance mechanisms and liability provisions, the courts had to improvise. Early landmark cases show that the courts reach deep into the constitution to identify the explicit as well as implicit legal sources. One frequently cited source is Article 21 of the Constitution, which guarantees Indian citizens the fundamental right to life.[6] Articles 47 and 48A, which fall under the non-binding "Directive Principles of State Policy" and require the government to improve public health and protect and improve the environment. Finally, Article 51 A(g) have also defined one of the fundamental duties of citizenship to "maintain a hygienic environment".

These constitutional provisions, interpreted broadly, provide Indian courts with considerable scope to weigh in on environmental jurisprudence. Over time, this legal framework has also absorbed additional legal principles that are drawn from international and foreign legal systems. Terms like "sustainable development", the "polluter pays" principle, and the "public trust" doctrine have also entered Indian environment discourse over the past three decades (Ghosh, 2019). Though these principles were not articulated in Indian statutory law, they have become regarded as an intrinsic part of Indian environmental law with some adjustments and adaptations that are needed in the Indian context.

3 Data

Estimating the impact of environmental litigation on environmental as well as human capital outcomes requires data with comprehensive information on all three sets of variables. We compile a unique database of all cases that pertain to water pollution that have been heard in the courts of India for the past 30 years and combine this with existing data on both water pollution and mortality from population surveys. We aggregate and then link these data together at the district-year level.

Analysis of demographic data in India is increasingly being conducted at the district level (Dreze and Murthi, 2001; Government of India, 2017; Mohanty, Fink, Chauhan and Canning, 2016; Singh, Kumar, Pathak, Chauhan and Banerjee, 2017; Spears, Ghosh and Cumming, 2013). Since the average Indian district contains a population of about 2-3 million people, and many critical decisions about policy are made at this level, district-level aggregates are meaningful and show considerable variation across the country (Government of India, 2017).

[6] This act states that "No person shall be deprived of his life or personal liberty except according to procedures established by law".

The different components of the working sample we construct for our analysis are summarized below, and greater detail the processes of data compilation are provided in the Appendix.

Legal cases and Judges There is no publicly available database of environmental litigation in India that is suitable for statistical analysis. To address this gap, we extracted all judgements passed by the National Green Tribunal of India, the state High Courts, and the Supreme Court of India that include a mention of the Water (Prevention and Control of Pollution) Act of 1974 , Air (Prevention and Control of Pollution) Act of 1981, and the Environment Protection Act of 1986. This unique data set consists of approximately 4000 observations (978 observations for the Water Act and 3021 for the Air Act). We focus here on the water cases only. By scraping publicly available websites, we were able to obtain texts of judgment as well as meta information on all pending and disposed cases, such as the dates of filing, registration and disposals, transfers between courts, judge, litigants and advocate names, acts involved and case types.

To determine whether a particular judgment is likely to have a positive impact on the environment, we rely on manual reading, interpretation and categorization by a team of law students.[7][8] We consider a case in our sample to be a "Green Verdict" on the basis of these recommendations. Specifically, we take the median of the scores assigned to the cases across the coders who coded the case and define it as "Green Verdict" if the median assigned environmental impact is positive.[9]

In addition to the environmental impact of cases, our coders also identified the precise location of the ruling, the geographic scope of the ruling (within the district, across all districts in a state or across the entire country), the names of the judges who ruled on the case, the basic attributes of the case and the month and year of the ruling. Summary statistics of all 978 cases are presented in the appendix. Summary statistics of the 516 cases that were successfully matched to the pollution data, and the 777 cases that were successfully matched to the mortality data are present in Table 1. We observe no discernible difference in the subsamples that defined the common support of the pollution and mortality data.

Note that the average case in our sample has a green score of 0.25 (the range is -2 to 2). 21 percent of cases are constitutional cases and 26 percent are appeals. 82 percent feature the government as the respondent and only 21 percent feature the government as the petitioner. There are on average 2 judges per case.[10]

The cases touch on a variety of themes. Most of them deal with pollution and environmental contamination, but judges deliberate on these issues in a variety of ways. In a very simple analysis of the incidence of keywords, we found the word "Pollution" in all the cases. India's "Constitution" is referenced in nearly half of these.[11] A third of the cases appear to be filed as Public Interest Litigation (PILS). The words "Public Interest" and "Public Trust" are cited 289 and 44 times respectively. Some

[7] These students, located in India, were trained by a lawyer with expertise of Indian law to read the judgments and label them based on their likely impact on the environment.

[8] We drafted a detailed training manual which provides information on how to use the portal, how to read and extract information from the judgement and FAQs. To ensure consistency in how cases were read and evaluated, we created a case coding portal using oTree, which is an open source framework for interactive tasks and games. To avoid errors and double-check the labels assigned by students, each judgment was assigned to at least two students for labelling independently. Discrepancies in labelling will be reconciled by assigning the judgment to a third student.

[9] Coders were asked to form an opinion on whether a case was likely to have "a positive effect on the environment" on a scale of -2 to 2 (-2: strongly anti-environment; -1: mildly anti-environment; 0: no impact on the environment; 1: mild positive effect on the environment and 2: strong positive effect on the environment)

[10] Of the 978 cases, 12 cases do not have the names of the judges who heard the case, 489 cases were heard by a single judge, 431 have two judges and 37 have three.

[11] Among these 477 references, Article 21, 47, 48 and 51 are cited 145, 11, 60 and 44 times respectively. The "Right to Life" is specifically mentioned 84 times. Other sections of the constitution are also routinely mentioned.

terms from international law are also cited. ”Polluter Pays” and ”Sustainable Development” are 115 and 75 times respectively. These variations will guide our analysis later in this paper.

Judge Biographies and Case Histories Our analysis also incorporates the biographic characteristics of judges. There is no publicly accessible database of judges for the courts of India. We have thus sourced and combined the lists of judges who have served at the high-courts and supreme-court of India since the date of the establishment of the courts.

Given that we are examining cases that are based on legislation from 1974, we are able to focus our attention on the post-1974 period. We draw these data from two sources: (a) the Judges Handbooks that have been released by the Supreme Court of India in 2014 and 2018; (b) the websites of the various High Court websites that list the names, biographies and career trajectories of the judges who have ever served at these courts. Details of the full sample are presented in the Appendix.

Summary statistics of the sample of judges who matched with the environmental cases however, are presented in Table 1. Note that 97 percent of the judges are male, and we observe about 1.87 environmental cases per judge. For each of these judges, we are able to extract a complete case history from our judicial database.[12]

Water Pollution To measure water quality, we use two sources of data. The first is the river pollution data that were compiled from the annual reports of India's Central Pollution Control Board (CPCB). These data were originally curated and digitized by Greenstone and Hanna (2014) and then further refined by Do, Joshi and Stopler (2018). For this analysis, we further extended the dataset's time-coverage to the year 2019, the last year available from the CPCB. The dataset now includes 2865 monitors over the time period 1986-2019. Our second source of data on water pollution is India's Water Resources Information System (WRIS). This is a a repository of national water resources data that receives input from many central and state agencies and provides a ”Single Window” source of updated data on water resources and related themes. The data covers 153 districts from 1984 to 2020.

The two sources of water data differ in the number of observations, districts covered and the specific locations within districts. They also differ in the types of pollution indicators that are reported. To address these issues, we combine both types of data and then aggregate the combined sample at the district-level. Since the CPCB does not report mean values of pollution after 2014, we rely on the maximum observed values in any given district and month for the entire period. Given that concerns over water quality can be triggered by irregularities in recorded pollution in most settings, we believe the maximum values are appropriate for study in our research design. Details of this process are described in the appendix.

Our main indicators of river quality are biological-oxygen-demand (BOD) and chemical-oxygen-demand (COD). These are common indicators of industrial water pollution (Brown and Caldwell 2001). BOD captures the amount of dissolved oxygen needed by water-borne, aerobic organisms to break down organic material present at a certain temperature (usually 20 degrees Celsius) and over a specific time period (usually five days). COD captures the amount of oxygen that can be consumed by reactions in a measured solution. The units for both measures of pollution are milligrams of oxygen consumed per liter (mg/l). We consider the logarithm of the raw value of these two pollutants as primary pollutants of interest.

We also consider a few other indicators of water quality: total coliforms (TOTCOLI), conductivity and temperature. Total coliforms are oft-used measures of domestic (as opposed to industrial) pollu-

[12]To do this, we scraped data from the public website Indian Kanoon. This yielded 7.2 million text cases in total. We were able to successfully identify judge names for 2.6 million of these cases. We then use fuzzy string matching to match the judges from the judge bios dataset to these cases. We have on an average 202 cases per judge (from these 2.6 million cases).

tion, which was a major focus of water policy in India. It is measured as the "most probable number" of coliform organisms per 100 milliliters of water (MPN/100 ml, reported in thousands). Conductivity is a measure of the ability of water to pass an electrical current. Dissolved salts can increase salinity and conductivity while inorganic chemicals (such as oil) reduce conductivity. According to the Environmental Protection Agency, conductivity is only useful as a general measure of water quality. Each water body tends to have a relatively constant range of conductivity that, once established, can be used as a baseline for comparison with regular conductivity measurements. Significant changes in conductivity could then be an indicator that a discharge or some other source of pollution has entered the aquatic resource.[13]. Our last measure of water quality, temperature, is not intended to be a measure of water pollution (though it can increase conductivity). We consider this indicator mainly as a falsification check. We expect to find no significant impacts of environmental rulings on this measure of pollution.

This list of pollution measures is admittedly limited to basic indicators. Other pollutants that are known to affect human health are not recorded consistently in our time period. We note that while these data are quite detailed, India's data systems for water in the time period being considered here are limited in their coverage, robustness, and efficiency (Government of India, 2018). Detailed data on a wide range of pollutants, particularly the presence of toxic heavy metals, is unavailable for the past 30 years.

Mortality To construct district-level estimates of child mortality in India, we draw on two national population-based household surveys that have been used to measure national and sub-national health outcomes in India that are representative at the district level and cover the time-period of the pollution data and legal case-data. These are the second round of the District Level Household Survey (DLHS-2:2002-04) and the fourth round of the National Family Health Survey (NFHS-4: 2015-16). The DLHS2 has been previously used to analyze the impacts of pollution on mortality (Do, Joshi and Stopler, 2018). The NFHS4, conducted 13 years after the DLHS 2, is also representative at the district level and has been used to examine demographic trends (Joshi et al., 2020). Details about the construction of mortality estimates is presented in the Appendix.

Combining data on pollution, court cases and judge case histories at the district-year level produces a sample of 8,856 observations that covers 153 districts for the time period 1984 to 2020. 516 court cases occurred in these districts during the specified time-frame. We are able to successfully identify an average of two judges per case. The key summary statistics of this working sample are summarized in Panel (a) of Table 2.

Combining data on mortality, court cases and judge case histories at the district-year level produces a sample of 24,169 observations that covers 678 districts for the time period 1974 to 2020 and is matched to 777 court cases.

4 Econometric Models

Our first goal is to estimate the impact of litigation on pollution levels. If green verdicts from the courts of India were to emerge randomly and are local in scope and impact, we would expect the following regression to identify the relationship between environmental rulings and outcomes:

$$Y_{dt} = \beta_1 + \beta_2 GreenVerdicts_{dt} + \beta_3 \mathbb{1}\{\text{At least one case in } d, t\} + X_{dt}\theta + \epsilon_{dt}. \tag{1}$$

Here Y_{dt} can be either pollution ($Pollution_{dt}$) or mortality ($Mortality_{dt}$). $Pollution_{dt}$ is a measure of pollution in district d at time t, $GreenVerdicts_{dt}$ measures the fraction of rulings in district d in year t

[13] https://www.epa.gov/national-aquatic-resource-surveys/indicators-conductivity

which are pro-environment (i.e. the median score assigned in the manual coding process described above is greater than 0) and X_{dt} is a vector of district and location-by-time characteristics, which includes year and district fixed effects. $Mortality_dt$ is the percentage of children born in a district d in year t who lost their lives within 1 month and 1 year of their date of birth. We also examine the incidence of mortality in the first year conditional on one month survival.

The variable $GreenVerdicts_{dt}$ denotes the number of green verdicts that occurred in district d in year t. This variable is determined through the manual coding of cases (as described earlier). Of the 978 cases in our sample, We found that 401 cases clearly pertained to a specific location – this was clear to at least two coders who read each case. A further 115 in the sample lacked information on the district of origin but it was apparent to the coders that these cases were applicable to the entire state. For these cases, we assumed that on the date of the judgement, the verdict applied to all the districts in the state. An additional 2 cases in our sample were pertinent to the entire country. Here we again assumed that on the date of the judgement, the verdict applied to the entire country. This approach assumes that a ruling that has been coded as applicable to district d applies to that specific district.

The assumption that a verdict in district d at time t will impact pollution in that very district in that time period as well as subsequent time-periods is justified in light of how India's common law system works. Judges establish common law through written opinions that are binding on future decisions of lower courts in the same jurisdiction. Moreover, given that many of these rulings pertain to specific environmental disputes that pertain to local firms and local institutions, rulings are quite specific and require actions such as the closure of a firm, the installation of special equipment or the imposition of fines to ensure greater compliance with environmental laws.

The main challenge in estimating this equation is that green verdicts from the courts are likely to be endogenous to environmental as well as mortality outcomes: pollution is affected by economic growth, the proliferation of particular types of pollutants in the environment, as well as investments in education, the growth of awareness in a population and the pressures of democratic politics. These are likely to result in a biased estimation of β_2, the effect of environmental rulings on observed outcomes. The direction of the bias will depend on which of the unobservables affect the emergence of litigation. Take the case of pollution as the dependent variable. If litigation is more likely to emerge in significantly polluted locations, we can expect the coefficient to be upward biased. If however, litigation is more likely to emerge in wealthy locations where there are also efforts to mitigate pollution, then we would expect the coefficient to be downward biased.

To address this issue, we instrument for green verdicts using judge-level characteristics under the assumption that the judicial system of India assigns cases to judges randomly, but judge characteristics do influence the likelihood of green verdicts in cases (see for example, Dobbie, Goldin and Yang, 2018 and Bhuller, Dahl, Løken and Mogstad, 2020 for recent examples).

Our first stage regression is as follows:

$$GreenVerdicts_{dt} = \gamma_1 + \gamma_{2t} JudgeChars_{dt} + X_{dt}\theta + u_{dt} \tag{2}$$

$GreenVerdicts_{dt}$ is the fraction of environmental cases in location i in year t which are coded as pro-environmental. The instrumental variables are district-year averages over variables capturing judges' writing styles for all judges sitting on environmental cases decided in year t and related to district d. Given that only 12 percent of district-year observations have any case at all, and there are an average of 0.18 cases per district year, with a maximum of 13 in the entirety of the corpus, the averages of judge attributes in no case approximate the averages of these variables for judges of a court.[14]

The main instruments are a set of 25 indicators that summarize the writing styles of the judges in our sample on cases heard *prior* to the pollution case in question. We construct these indicators by applying the

[14]The average High Court of India has at least 19 judges at any given time and this group changes each year, with approximately one-third of the judges transferred.

"doc2vec" algorithm on the corpus of cases of the 302 judges in our pollution sample and 398 judges in the mortality sample (Le and Mikolov, 2014). [15] These vectors can be interpreted as numeric representations of the semantic structure of a body of text.

Figure 4 presents some visualizations of this approach. Since it is not possible to produce a direct visualization of high dimensional vectors that analyze a corpus of rulings, we use a technique called t-distributed Stochastic Neighbour Embedding (t-SNE) to produce two-dimensional representations of the original 25 vectors. The top panel presents the two dimensional visualization of the case vectors (colored by the hand-labelled impact score), the middle panel presents the judge level embedding (colored by the mean impact score of the cases the judge has adjudicated) and the lowest panel presents the judge embedding along with the vector representation of key phrases which were jointly trained along with the case vectors by Doc2Vec. We note that there is considerable variation in the writing style across judges, and also considerable variation across cases. This variation loosely corresponds to the incidence of key words from Indian environmental jurisprudence.

X_{dt} includes a set of control variables pertinent to district d and time t. These include district-year averages of the number of green cases in our sample where the government is a petitioner, the number of cases in our sample that are appeals, and the cases that are regarded by our team of coders as constitutional cases. X_{dt} also consists of control variables that captures economic development at the district-level, as measured by satellite: the total forest cover and the calibrated levels of night-lights.[16]

4.1 First stage at the case-level

Given that the first-stage regression pertains to cases in the courts but the overall regression pertains to districts, we take a small detour to understand the first-stage at the case-level. In the top-left panel of Figure 4 we estimate equation 2 using the leave-one-out cross-validation approach to generate a prediction of the likelihood of a green verdict with the set of judge writing characteristics. We see that these judge characteristics are indeed highly predictive for the likelihood of a green verdict. Separately, in the top-right panel, we also run a regression of the predicted likelihood of a case being green, taking into account only control variables, on the the predicted likelihood of a case being green using the full set of judge characteristics and controls. Here we can see that the predictive power from controls only is very low and seems to uncorrelated with the prediction from the full set of judge characteristics and controls.

In the middle panel of Figure 4, we aggregate our variables on a district-year level and run the same regressions as in the top row for all district-years which have at least one case. We see, that the results are very similar to the case-level ones. Finally, in the bottom row, we rerun these regressions on a district-year level, including also district-years with no cases. This level-analysis corresponds to our main econometric specification.

Next, we move to a formal 2SLS framework, where the second stage-regression is as follows:

$$Pollution_{dt} = \delta_1 + \delta_2 Gree\widehat{nV}erdict_{dt} + X_{dt}\phi + \eta_{dt}. \tag{3}$$

Coefficient δ_2 in equation (3) now provides the estimated impact of green verdicts on pollution in the

[15] doc2vec is a package that provides an efficient framework for text analysis and natural language processing (NLP). The algorithm takes as a corpus of texts (here, judge rulings) as an input, applies a neural network algorithm that analyzes the co-occurrence of specific words in relation to other words, and creates a 25-dimensional vector representation of the entire body of text. Stop words such as "is", "are", "the", "and", "we", "our", "ours", "ourselves", "you", "your", "yours," etc. are removed from the list of tokens. It is assumed that the closer tokens are to each other, the greater is their semantic relationship. The 25 dimensions produced with doc2vec are ultimately a numeric representation of the semantic meaning of each token within a wider body of language.

[16] Forest cover data comes from Vegetation Continuous Fields (VCF), a MODIS product that measures tree cover at 250m resolution from 2000 to 2019. VCF is predicted from a machine learning algorithm based on broad spectrum satellite images and trained with human-categorized data, which can distinguish between crops, plantations and primary forest cover. These data were linked to the SHRUG as part of work conducted by Asher, Garg, and Novosad (2018).

average district. This point estimate provides a reference point for the magnitude of the court's impact on environmental outcomes.

Here $Pollution_{dt}$ is a measure of pollution in location d at time t. *Green Verdicts*, the endogenous variable, measures the predicted number of pro-environment rulings related to water in location i in year t, and X_{it} is a vector of district and location-by-time characteristics, which includes district and year-month fixed effects.

It is plausible that the potential effect of a judgement occurs over time rather than all at once. To take this into account, we must interpret each judgement as a policy and use (together with the IV approach) a distributed lag model. To address this possibility, we estimate the dynamic model with leads and lags for the policy.

5 Results

5.1 Impacts on Pollution

To obtain a point estimate of the impact of green verdicts on pollution levels, we estimate equation 1 (OLS estimation) and equation 3 (IV). In Table 2, we see results from three different specifications: omission of the districts and years that have no environmental verdicts at all (columns 1 and 2), inclusion of dummies for those districts and years (columns 3 and 4), inclusion of dummies and fixed effects for districts and years (columns 5 and 6) and the inclusion of dummies, fixed effects and covariates related to the cases (constitutional case, appeal case, and the involvement of the government as a respondent in the case) in the full specification (columns 7 and 8). Note that for our preferred specification (columns 7 and 8), both the OLS and IV estimates for the variable "Fraction of Green Cases" are negative and statistically significant at the 1 percent level. The IV coefficient, -0.267, suggests that a 1% increase in the fraction of green cases in a district in a given year reduces the highest observed BOD value in the district by approximately 26.7% in that very year.

Our empirical strategy hinges on the assumption that green verdicts in specific districts and years are affected by the random assignment of judges to cases. Judge attributes such as gender, education and ideology expressed in previous rulings may affect the outcomes of cases.

We explore the strength of these assumptions by a closer look at the first-stage regressions. Table 3 presents the results of the first-stage across a range of specifications. Since the regression model is estimated at the district-year level, the instruments are also averages of the attributes of cases at the district-year level. These include the fraction of judges who were assigned green cases in a district-year who have a post-graduate degree and a set of 25 textual variables that summarize the corpus of cases in the record of the judges, to create these textual variables we removed all the water pollution cases from the corpus to mitigate concerns of endogeneity. In the full specification (Table 3, column 5), we note that the addition of these instruments, a full set of covariates, year and district fixed-effects and clustered standard errors provides a robust first-stage.

In column (1) of Table 3, we see just the textual IVs included in the specification. The first-stage F-statistic is 107.9. In our preferred specification (Column 4), which features a full set of district-year fixed effects and a dummy for having a green case, we see an F-statistic of 65.20. This is well above the threshold that is typically employed to evaluate the power of the first-stage regressions and increases our confidence in our first-stage (Andrews, Stock and Sun, 2019).

We also explore the relationship between green verdicts and judge attributes at the case-level. Figure 3 depicts (a) the mean of the actual coded values (on the scale of -2 to 2) versus the predicted values of the cases on the basis of only the judge characteristics; (b) the predicted values of the verdict based on ex-ante case characteristics regressed on predicted values of the verdict based on judge characteristics. We infer from this graph that judge characteristics are a powerful predictor of case outcomes. Green verdicts are clearly associated with judge characteristics. However, the judge's tendency to issue green verdicts is not correlated with the prediction of a green verdict based on case characteristics. This further strengthens our confidence in our instruments: under the assumptions of random judge assignment, we find that judge attributes are indeed an appropriate instrumental variable for case-verdicts, exogenous to case characteristics, as has been noted in many other studies in law and economics (Belloni, Chen, Chernuzhukov, Hansen 2012).

Table 4 presents the results of our preferred specification for additional pollution outcomes: ln(COD), ln(TOTCOLI), ln(conductivity) and ln(temperature). The estimates from our preferred specification (in Table 2) are also included here, along with the corresponding F-statistics from the first-stage regressions. We note that across all specifications, F-statistics exceed 60. This strengthens our confidence in these instruments and our identification strategy more generally.

Table 5 presents the results of three key pollutants – BOD, COD, and TOTCOLI – with data that is not aggregated at the yearly level. Rather, it contains monthly observations at the district-year level. All key variables are defined exactly as the year-district sample whose results we have seen thus far. We note however, that this sample is about 7-10 times larger than the yearly sample, depending on the pollutant. Many district-month observations are missing in this data, creating an unbalanced sample with considerable noise.

Next we estimate equation (3) with dynamic effects: we consider effects three years in advance of the ruling, and five years after. The coefficients for these three leads and five lags are presented for three pollutants – BOD, COD and TOTCOLI – in Figure 4. We note that the coefficients are negative and statistically significant at the 1% level for both BOD and COD. We believe this is largely driven by the curbing of pollutant behavior during the time that the case is being deliberated at the courts. The filing of an environmental case against a polluter generates considerable publicity in India (Baxi, 1985; Sathe, 2002). The courts are highly salient in the state-society interface (Kapur, Mehta, Vaishnav, 2018). Most cases are resolved within about 3 years. Indeed, the average tenure of a judge at a court in our sample is also about 3 years. We can expect a firm that is subject to an environmental case to be scrutinized by pollution inspectors, particularly if it is a large polluter to begin with (Duflo, Greenstone, Pande and Ryan, 2018). As discussed earlier, the Water Act is largely enforced through the central and state-run pollution control boards that lack the capacity to ensure compliance with environmental laws or court mandates.

We also note that even though all three pollutants decline in the immediate aftermath of the verdict, pollution rises thereafter. This is consistent with many past analyses of court-led environmental activism in India. As discussed earlier in this paper, the courts lack an enforcement mechanism. As a result, the period of adherence that follows strong verdicts is often not sustained (Ghosh, 2019). Water clean-up projects that are designed in the aftermath of such rulings for example, are often built in the format of PPPs and the scale of the projects becomes unsustainable almost immediately after the project is constructed (Shambaugh and Joshi, 2019, 2021). Many water clean-up projects along the Ganga river became rent-seeking opportunities almost immediately after they were created, leading to persistent pollution despite considerable judicial activism to protect the river (Alley, 2002).

5.2 Impacts on Mortality

Table 7 estimates the second-stage impacts of green verdicts on mortality. We consider three measures of mortality in our estimation of equation (3) . We consider three dependent variables: death in the first year of life (column 1), death in the first month of life (column 2), and death in the first year conditional on surviving the first month (column 3).

The coefficient of interest to us is β_2 in equation (3), which measures the impact of predicted green verdicts on mortality outcomes in a district-year. We note that this coefficient is negative but not statistically significant in any of the regressions.

Next we estimate the mortality equation (3) with dynamic effects: we consider effects three years before, and five years after the ruling. The coefficients for these three leads and five lags are estimated for all three measures of mortality considered in Table 6. Results are in Figure 6. We note no pre-trends prior to the verdict for any of the mortality indicators. Again, we observe no

We emphasize that these results must be interpreted cautiously. As is seen in Figure 1, our sample of districts with green verdicts is small. In previous work, Do, Stolper and Joshi (2018) found localized effects along a single river. The considerable ecological, demographic and institutional diversity of India together with the rarity of mortality in recent years may make it difficult to find strong effects. Examining localized effects of green verdicts in this sample is an important next step for our research.

5.3 Robustness Checks: Neighboring Districts

Our empirical strategy hinges on the assumption that judges are randomly assigned once we condition on case characteristics and judge characteristics (which include histories of their previous judgements) and also district and year-month fixed effects. Implicit in this assumption is that these variables fully explain the emergence of green verdicts in polluted locations. To bolster this argument, we employ a placebo test that provides additional evidence supporting our exclusion restriction. Specifically, we regress green verdicts on judge characteristics in a geographically neighboring district and then examine whether these green districts in neighboring districts affect pollution and mortality in the districts in our sample.[17]

[TO BE ADDED LATER]

6 Discussion

Our estimates of the impact of green verdicts on water pollution levels are the first documented empirical evidence of the judiciary's success in India's regulation of water quality over the past three decades.

We find it striking that even though green rulings clearly reduce pollution, the effects are not big enough to improve environmental quality over time. Water quality along the Ganga and Yamuna rivers, for example, remains poor. Recent evidence from the stringent Indian Covid-19 lockdown (March 2020-June 2020) found a reduction in irrigation and power demands, increased water storage, increased flow and a significant improvement in the concentrations of pollutants such as dissolved oxygen, BOD and nitrates (Dutta, Dubey and Kumar, 2020).[18] Prior to the pandemic however, the number of classified 'polluted river stretches' in India doubled from 150 to 302, and the gap between sewage load and sewage treatment capacity expanded (*Daily Mail* 2015).

This raises two questions: First, why does the judiciary succeed where other Indian government action – such as the ambitious National River Conservation Plan (NRCP) – fails? And second, why does the effectiveness of the judiciary dissipate in the year after the ruling?

On the first question, it is plausible that decisions from the judiciary differ from those made by the executive in that they mandate agents to take specific and verifiable actions aimed at pollution-mitigation, and they are also held immediately accountable for these actions. Furthermore, the set of stakeholders empowered to monitor the execution of a judicial decision, which includes citizens, might reduce the scope for non-compliance (Duflo et al. 2018).

On the second question, we believe that the introduction discussion in the background section of this paper is very relevant. Environmental regulation in India has some structural weaknesses. There is a large corpus of laws on the books, but the enforcement systems are complex and ultimately, no single entity is ultimately responsible for protecting water resources (Ghosh, 2019). Unlike air quality, which is

[17] We use geospatial maps with district boundaries to construct lists of neighboring districts for each district in our sample and average judge characteristics across these neighboring districts, if they have cases present for a specific district and year.

[18] Dutta, Dubey and Kumar (2020) even found that the river became fit for drinking for the first time in years.

more observable and traceable to a source, water toxicity can be invisible to the naked eye and transports undiscerned in flowing waterwarys to locations far away from the original source (Greenstone and Hanna, 2014; Do, Stolper and Joshi, 2018).

India's system of democratic politics is not conducive to reform. The millions of firms that are located in highly polluting industrial clusters provide valuable employment and valuable products for export (Joshi and Shambaugh, 2018). As in the case in many contexts across the world, there is very little political will to prioritize environmental quality at the cost of economic growth (Stiglitz, Sen and Fitoussi, 2010).

Here is is also notable that we did not find a mortality impact of the rulings. While the lack of significance may be driven by noise in the data, it is also plausible that citizens are aware of water pollution and adopt pollution-mitigation systems close to the location of final water use, thereby insulating human health from the challenge of water toxicity. We believe this is an important area for future research.

7 Conclusion

This paper provided an empirical study of the broad impact of judicial rulings on environmental outcomes in India, a developing country with some of the highest levels of water toxicity in the world. Our analysis is based on a novel dataset that combines legal, environmental and demographic variables at the level of districts. Our empirical model seeks to identity the causal relationship between a pro-environmental ruling and actual environmental outcomes. Since rulings may be endogenous to outcomes, we use an IV framework, with the textual features of the judges who preside over these cases to predict the likelihood of a green verdict. In the second stage of analysis we consider both pollution and mortality as key outcomes.

We find that the rulings precipitated reductions in two specific measures of water quality – biological oxygen demand (BOD) and chemical oxygen demand (COD) – the most common measures of industrial pollution in surface water. These effects however, are confined to the year of the ruling. We find neither any long-term impact on pollution nor any immediate or long-term impact of the rulings on neonatal mortality or infant mortality. This suggests that judicial policies do succeed in lowering pollution, but the challenge of enforcement limits their impact on both longer-term pollution and human health. Sustained improvements in water quality and child health in India require more than green verdicts.

8 References/Useful Links

Abbot, C. (2009). *Enforcing pollution control regulation: strengthening sanctions and improving deterrence.* Bloomsbury Publishing.

Abraham, C. M., and Abraham, S. (1991). The Bhopal case and the development of environmental law in India. *International & Comparative Law Quarterly*, 40(2), 334-365.

Alley, K. D. (2002). *On the banks of the Ganga: When wastewater meets a sacred river.* University of Michigan Press.

Andrews, I., Stock, J. H., and Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11, 727-753.

Asher, S., Garg, T., & Novosad, P. (2020). The ecological impact of transportation infrastructure. *The Economic Journal*, 130(629), 1173-1199.

Baxi, U. (1985). "Taking suffering seriously: Social action litigation in the Supreme Court of India". Third World Legal Studies, 107, 4(6): 1-6.

Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6): 2369-2429.

Bhuller, M., Dahl, G. B., Løken, K. V., and Mogstad, M. (2020). Incarceration, recidivism, and employment. *Journal of Political Economy*, 128(4), 1269-1324.

Bhuwania, A. (2017). *Courting the people: Public interest litigation in post-emergency India.* Cambridge University Press.

Dias, A. (1994). Judicial activism in the development and enforcement of environmental law: Some comparative insights from the Indian experience. Journal of Environmental. Law, 6, 243.

Dobbie, W., Goldin, J., and Yang, C. S. (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, 108(2), 201-40.

Duflo, E., Greenstone, M., Pande, R., & Ryan, N. (2018). The value of regulatory discretion: Estimates from environmental inspections in India. Econometrica, 86(6), 2123-2160.

Dutta, V., Dubey, D., and Kumar, S. (2020). Cleaning the River Ganga: Impact of lockdown on water quality and future implications on river rejuvenation strategies. *Science of the Total Environment*, 743, 140756.

Epple, D., and Visscher, M. (1984). Environmental pollution: Modeling occurrence, detection, and deterrence. *The Journal of Law and Economics*, 27(1), 29-60.

Fuller, R., Sandilya, K., and Hanrahan, D. (2019). Pollution and Health Metrics: Global, Regional, and Country Analysis. Global Alliance on Health and Pollution.

Ghosh, S. (Ed.). (2019). *Indian Environmental Law: Key Concepts and Principles.* Orient BlackSwan.

Government of India (2018). Composite water management index. Niti Aayog, National Institution for Transforming India, Government of India.

Joshi, S., Borkotoky, K., Gautam, A., Datta, N., Achyut, P., Nanda, P., and Verma, R. (2020). Poised for a dividend? Changes in the life trajectories of India's young women over the past two decades. *PLoS One*, 15(12), e0242876.

Kapur, D., Mehta, P. B., and Vaishnav, M. (Eds.). (2018). *Rethinking Public Institutions in India.* Oxford University Press.

Le, Q., and Mikolov, T. (2014). Distributed representations of sentences and documents. In *International conference on machine learning*, Proceedings of the 31st International Conference on Machine Learning, pp. 1188-1196.

Malleson, K. (2016). *The new judiciary: The effects of expansion and activism.* Routledge.

Rajamani, L. (2007). "Public Interest Environmental Litigation in India: Exploring Issues of Access, Participation, Equity, Effectiveness and Sustainability", *Journal of Environmental Law*, Volume 19, Issue 3, 2007, Pages 293–321.

Sathe, S. P. (2002). *Judicial activism in India.* Oxford University Press, USA.

Shambaugh, G., and Joshi, S. (2021). Bridges over Troubled Waters? The Political Economy of Public-Private Partnerships in the Water Sector. *Sustainability*, 13(18), 10127.

Joshi, S., and Shambaugh, G. (2018). Oversized solutions to big problems: The political economy of partnerships and environmental cleanup in India. *Environmental Development*, 28, 3-18.

United Nations Development Program (2009). *Analysis of existing environmental instruments in India.*

Woodhouse, P., and Muller, M. (2017). Water governance—An historical perspective on current debates. *World Development*, 92, 225-241.

World Bank (2013) *India : Diagnostic Assessment of Select Environmental Challenges*, Volume 2. *Economic Growth and Environmental Sustainability, What Are the Tradeoffs?.* World Bank, Washington, DC.

Figures and Tables

A: Max log(BOD mg/l) per District

B: River Pollution Cases per District, 1982 - 2020

C: Overlap between Cases and Pollution

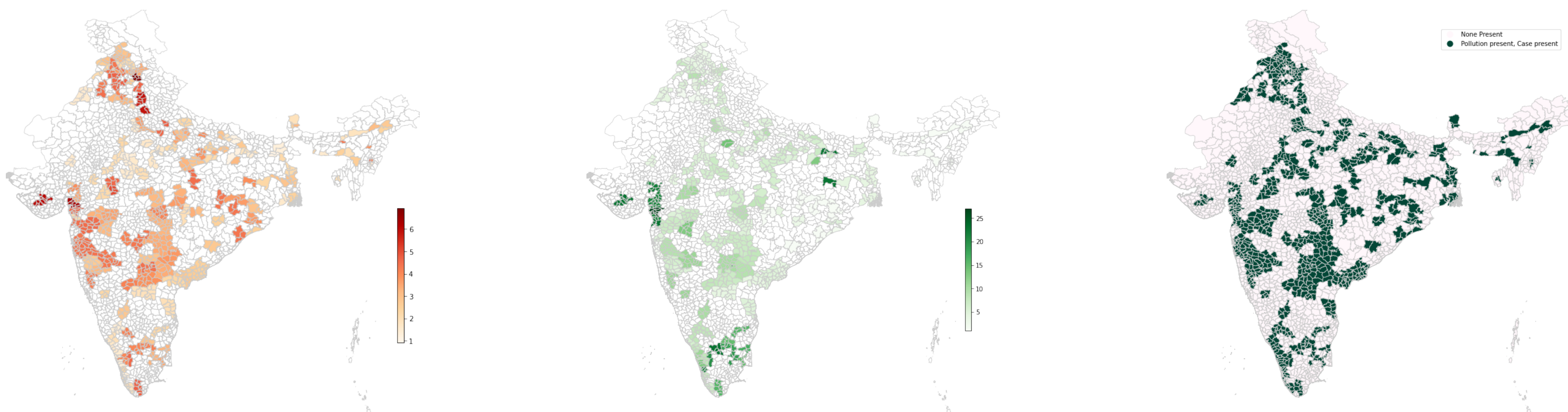

Figure 1: Maps of Available Data

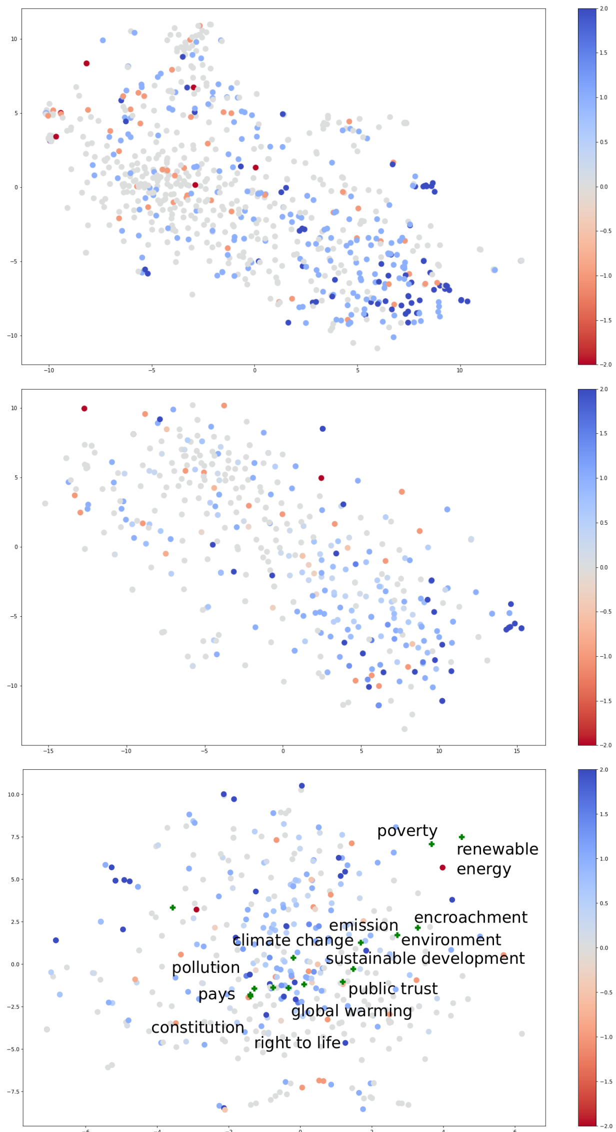

Figure 2: A visual illustration of judge-writing styles. Notes: Each case in our corpus is represented as a 25 dimensional vector using Doc2Vec. The top panel presents the two dimensional visualization of the case vectors (colored by the hand-labelled impact score), the middle panel presents the judge level embedding (colored by the mean impact score of the cases the judge has adjudicated) and the lowest panel presents the judge embedding along with the vector representation of key phrases which were jointly trained along with the case vectors by Doc2Vec.

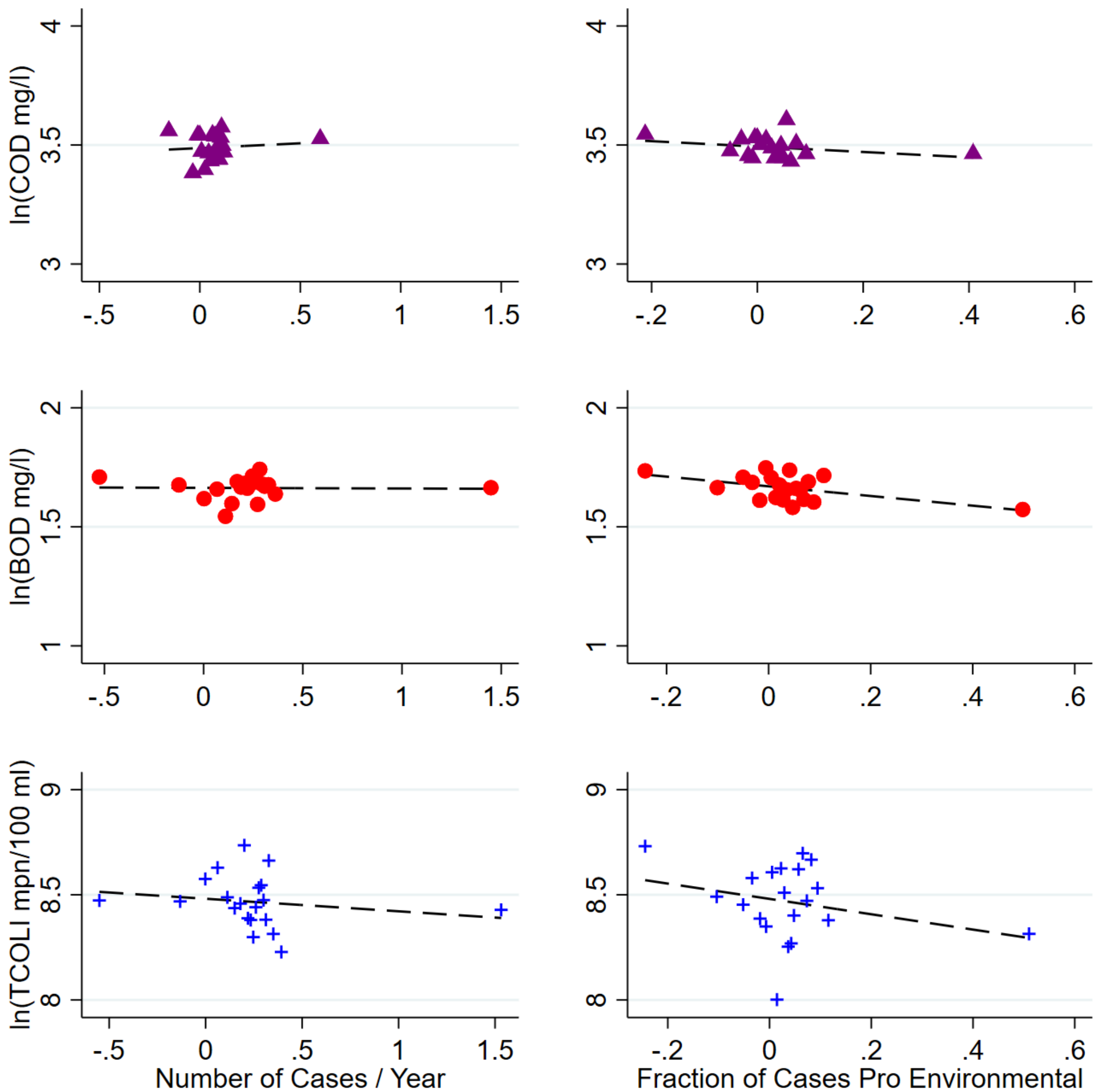

Figure 3: Binscatters of the relationship between water pollution and the number of environmental cases per district-year (left) and between water pollution and the fraction of pro green cases per district-year (right). Variables are residualized by the district-year means of case characteristics (Whether the government is a respondent, whether it is an appeal and or a constitutional case) and by district and year fixed effects.

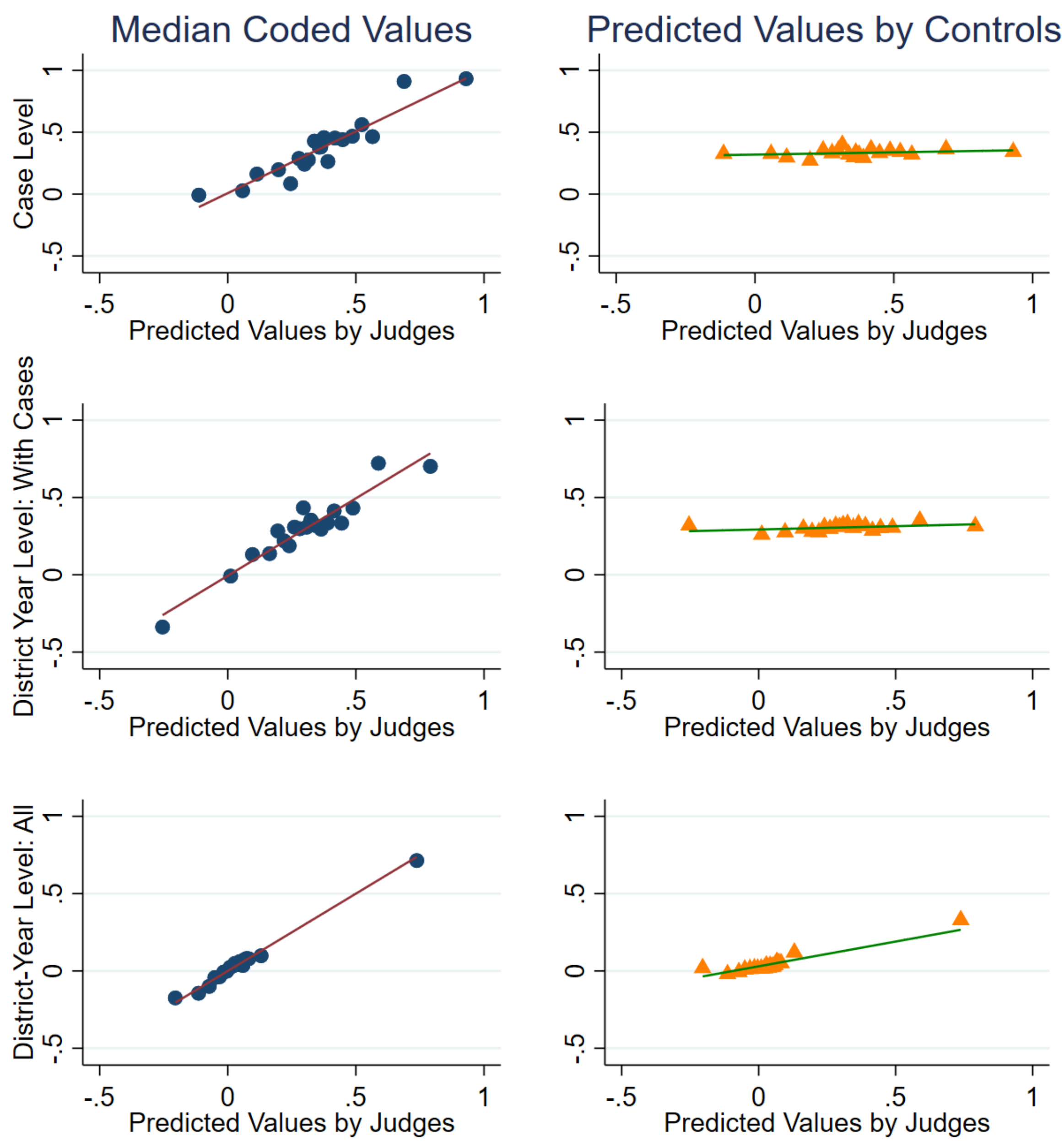

Figure 4: Graphical First Stage. The graphs on the left are binscatters of (residualized) median codings of cases on the (residualized) predicted codings by case characteristics. The graphs on the left are binscatters of the predicted (residualized) median case codings by case characteristics on predicted (residualized) median case codings by judge characteristics. The top row is on a case-level, the middle row on a district-year level including only district-years with at least one year and the bottom row is on the district-year level including all observations.

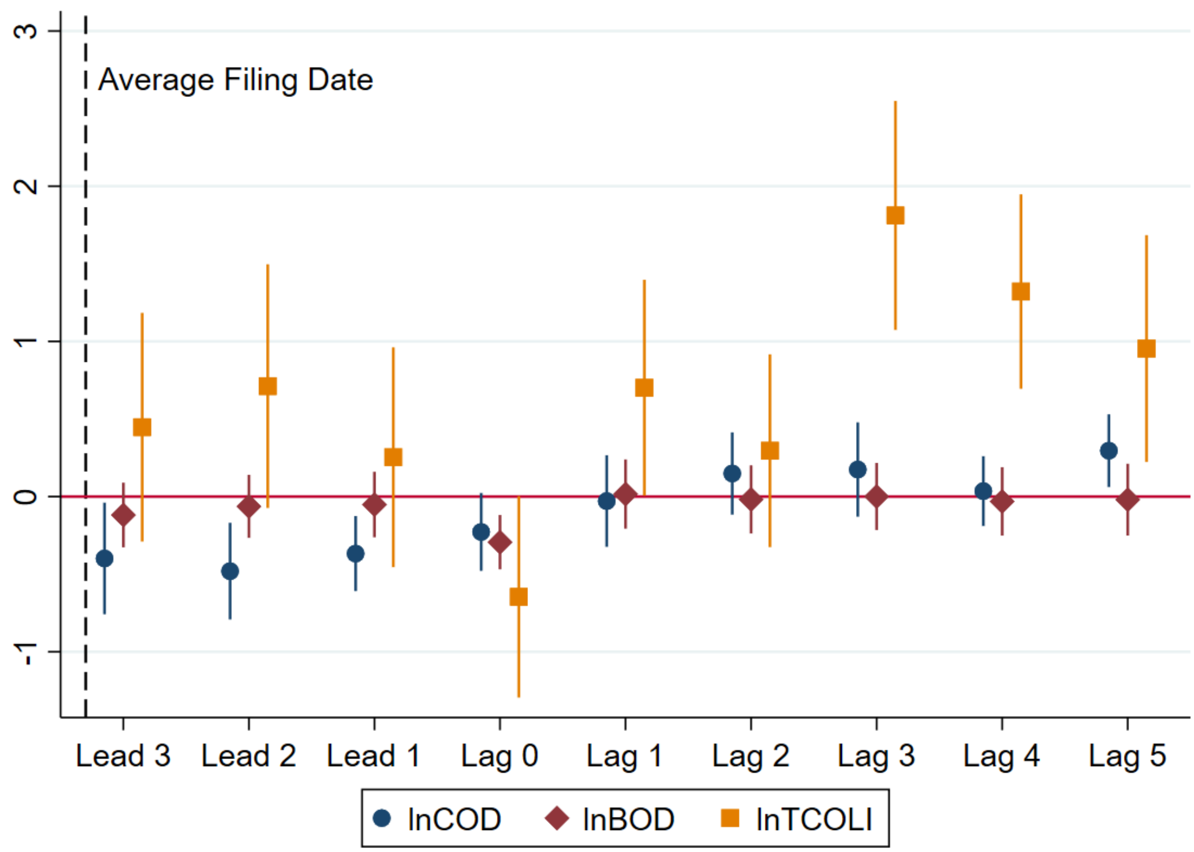

Figure 5: Lags and Leads of Pollution Regressions

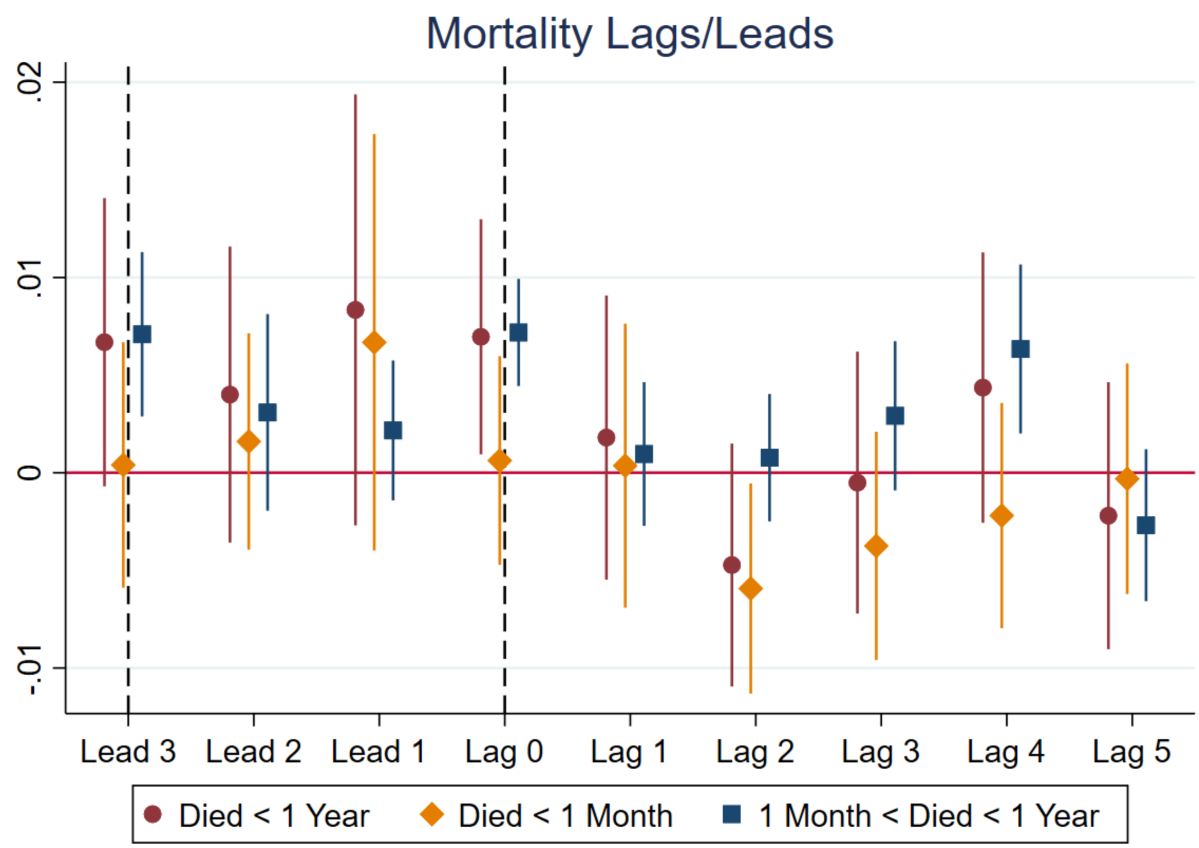

Figure 6: Lags and Leads of Mortality Regressions

Table 1: Summary Statistics for each source of data

	N	Mean	SD	Min	Max
Pollution (Monitor-Year)					
Max BOD (mg/l)	23,413	9.57	38	0	1820
Max COD (mg/l)	6,089	39.95	63	0	1750
Max Total Coliform (mpn/100 ml)/106	19,628	7	322	0	23,000
Max Temperature (°C)	24,623	29	6	0	269
Max Conductivity (µmhos/cm)	22,843	2,281	9440	0	513,000
Case Level Data - Pollution					
Appeal	516	0.25	0	0	1
Constitutional	516	0.21	0	0	1
Government is Respondent	516	0.82	0	0	1
Government is Petitioner	516	0.14	0	0	1
Number of Judges	516	2	1	0	3
Environmental Impact (Median Coding)	516	0.34	1	-2	2
Maximum Forest Cover	286	24.04	15	4	66
Total Forest Cover	286	70,997.99	354796	161	2198364
Maximum Nightlights	176	16.16	17	1	63
Total Caliberated Nightlights	176	4,048.10	16031	3	88983
Case Level Data - Mortality					
Appeal	777	0.25	0	0	1
Constitutional	777	0.22	0	0	1
Government is Respondent	777	0.86	0	0	1
Government is Petitioner	777	0.11	0	0	1
Number of Judges	777	2	1	0	3
Environmental Impact (Median Coding)	777	0.35	1	-2	2
Maximum Forest Cover	557	25.42	15	1	72
Total Forest Cover	557	65,954.68	295902	119	2737216
Maximum Nightlights	331	23.07	23	0	63
Total Caliberated Nightlights	331	12,542.39	32648	1	261839
Judge Level Data (Pollution Sample)					
Male	302	0.97	0	0	1
Graduate Level Education	302	0.39	0	0	1
Post-Graduate Level Education	302	0.13	0	0	1
Judge Level Data (Mortality Sample)					
Male	398	0.96	0	0	1
Graduate Level Education	398	0.38	0	0	1
Post-Graduate Level Education	398	0.12	0	0	1

Table 2: Summary Statistics of the two working samples

District-Year Level Data (Pollution Sample)					
Case Present	8856	0.12	0	0	1
Number of Pro-Environemental Cases	8856	0.18	1	0	13
Share of Cases Pro-Environmental	8856	0.03	0	0	1
Average Number of Judges / Case	8856	0.22	1	0	3
Share of Appeal Cases	8856	0.02	0	0	1
Share of Constitutional Cases	8856	0.04	0	0	1
Share of Cases w/ Government as Petitioner	8856	0.01	0	0	1
Share of Cases w/ Government as Respondent	8856	0.10	0	0	1
Max BOD (mg/l)	5650	12.53	34	0	1025
Max COD (mg/l)	3053	55.65	80	1	1750
Max Total Coliform (mpn/100 ml)/10^6	5057	15.09	514	0	23000
Max Temperature (°C)	5614	29.69	6	0	269
Max Conductivity (µmhos/cm)	5476	1,936.84	7327	0	81800
log Max BOD (mg/l)	5649	1.66	1	-2	7
log Max COD (mg/l)	3053	3.49	1	0	7
log Max Total Coliform (mpn/100 ml)	5057	8.47	3	1	24
log Max Temperature (°C)	5541	3.39	0	2	6
log Max Conductivity (µmhos/cm)	5475	5.99	2	-1	11
log Max BOD (mg/l) (MA)	6254	1.67	1	-2	7
log Max COD (mg/l) (MA)	5742	3.41	1	0	7
log Max Total Coliform (mpn/100 ml) (MA)	5888	8.52	3	1	24
log Max Temperature (°C) (MA)	6185	3.38	0	0	6
log Max Conductivity (µmhos/cm) (MA)	6237	6.02	2	-1	11
District-Year Level Data (Mortality Sample)					
Case Present	24169	0.07	0	0	1
Share of Cases Pro-Environmental	24169	0.02	0	0	1
Average Number of Judges / Case	24169	0.11	0	0	3
Share of Appeal Cases	24169	0.01	0	0	1
Share of Constitutional Cases	24169	0.02	0	0	1
Share of Cases w/ Government as Petitioner	24169	0.01	0	0	1
Share of Cases w/ Government as Respondent	24169	0.06	0	0	1
Infants dying aged <1 Year (%)	24169	0.08	0	0	1
Infants dying aged <1 Month (%)	24169	0.06	0	0	1
Infants dying, conditional on surviving first month (%)	24169	0.03	0	0	1

Table 3: Comparison of Yearly log(BOD) specifications

	Log of Yearly Maximum log(BOD) per District (mg/l)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Fraction of Green Cases	0.177^{*}	0.196	0.177^{*}	0.196	-0.183^{***}	-0.302^{***}	-0.162^{**}	-0.267^{***}
	(0.0973)	(0.153)	(0.0973)	(0.153)	(0.0702)	(0.102)	(0.0684)	(0.0986)
Dummy for Presence of a Case			0.202^{***}	0.197^{***}	0.0814^{*}	0.117^{**}	0.0366	0.0702
			(0.0652)	(0.0595)	(0.0441)	(0.0509)	(0.106)	(0.108)
District-years with no cases	Dropped	Dropped	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied
Year and District FEs					Yes	Yes	Yes	Yes
Covariates							Yes	Yes
Clustering	District	District	District	District	District	District	District	District
K-P First Stage F		107.9		110.7		74.34		65.20
Adj. R2	0.000972	0.000931	0.00588	0.00587	-0.0450	-0.0456	-0.0444	-0.0449
N	859	859	5649	5649	5649	5649	5649	5649

Note: Cases are defined as green case if the median reader classified them as either having a "mild positive impact" or a "strong positive impact" (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table 4: Comparison of Yearly log(BOD) First Stage Specifications

	Fraction of Green Cases per Year and District			
	(1)	(2)	(3)	(4)
Dummy for Presence of a Case		0.133^{***}	0.143^{***}	0.0800^{**}
		(0.0240)	(0.0253)	(0.0373)
Government is Respondent				0.0791^{**}
				(0.0358)
Appeal				-0.0167
				(0.0395)
Constitutional				0.113^{**}
				(0.0477)
District-years with no cases	Dropped	Dummied	Dummied	Dummied
25 Textual IVs	Yes	Yes	Yes	Yes
Year and District FEs			Yes	Yes
Covariates				Yes
Clustering	District	District	District	District
F	107.9	110.7	74.34	65.20
N	859	5649	5649	5649

Note: Cases are defined as green case if the median reader classified them as either having a "mild positive impact" or a "strong positive impact" (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table 5: Yearly Pollution Regressions

	(1) ln(COD)	(2) ln(BOD)	(3) ln(TCOLI)	(4) ln(Conductivity)	(5) ln(Temperature)
Fraction of Green Cases	-0.240*	-0.267***	0.0486	-0.0836	-0.0272
	(0.130)	(0.0986)	(0.355)	(0.111)	(0.0185)
Dummy for Presence of a Case	0.288**	0.0702	0.130	-0.0673	0.00199
	(0.114)	(0.108)	(0.214)	(0.117)	(0.0280)
District-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes
Clustering	District	District	District	District	District
K-P First Stage F	92.17	65.20	71.12	60.03	71.54
Adj. R2	-0.0800	-0.0449	-0.0516	-0.0474	-0.0466
N	3053	5649	5057	5475	5541

Note: Cases are defined as green case if the median reader classified them as either having a ”mild positive impact” or a ”strong positive impact” (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table 6: Monthly Pollution Regressions

	(1) ln(COD)	(2) ln(BOD)	(3) ln(TCOLI)
Fraction of Green Cases	-0.0401	0.0502	0.893**
	(0.120)	(0.0989)	(0.370)
Case Dummy	0.186**	0.0505	-0.0877
	(0.0918)	(0.110)	(0.266)
District-years with no cases	Dummied	Dummied	Dummied
Year, Month and District FEs	Yes	Yes	Yes
Covariates	Yes	Yes	Yes
Clustering	District	District	District
K-P First Stage F	103.7	87.79	184.5
Adj. R2	-0.00971	-0.00980	-0.0111
N	30955	34677	30871

Note: Cases are defined as green case if the median reader classified them as either having a ”mild positive impact” or a ”strong positive impact” (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table 7: Yearly Mortality Regressions

	(1) Died1Y	(2) Died1M	(3) Died1YC
Fraction of Green Cases	0.000128	-0.000879	0.00101
	(0.00228)	(0.00190)	(0.00124)
Case Dummy	-0.00232***	-0.00157**	-0.000783*
	(0.000892)	(0.000737)	(0.000411)
District-years with no cases	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes
Covariates	Yes	Yes	Yes
Clustering	District	District	District
K-P First Stage F	162.8	162.8	162.8
Adj. R2	0.922	0.948	0.981
N	15024	15024	15024

Note: The time-period of the mortality sample spans 1997–2017. Cases are defined as green case if the median reader classified them as either having a ”mild positive impact” or a ”strong positive impact” (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Fraction of green cases is instrumented for by the district-year means of 25 textual features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table 8: Neighbouring Districts

	(1) ln(COD)	(2) ln(BOD)	(3) ln(TCOLI)	(4) ln(Conductivity)	(5) ln(Temperature)
Neighbouring Fraction of GreenCases	-0.341***	-0.0608	-0.00778	-0.0523	-0.0111
	(0.101)	(0.0830)	(0.373)	(0.107)	(0.0151)
Case Dummy	0.263***	0.0154	0.155	-0.130	-0.0280**
	(0.0987)	(0.0803)	(0.202)	(0.128)	(0.0118)
District-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes
Clustering	District	District	District	District	District
K-P First Stage F	144.0	106.3	95.72	97.66	108.9
Adj. R2	0.576	0.606	0.602	0.688	0.461
N	3053	5649	5057	5475	5541

Notes: Pollution indicators for each district are ”matched” with pollution cases in upto five neighboring districts (averaged). Cases in a neighboring district are defined as green case if the median reader classified them as either having a ”mild positive impact” or a ”strong positive impact” (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates for the neighboring district are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table 9: Neighbouring Districts Impact on Mortality

	(1) Died $<$ 1 Year	(2) Died $<$ 1 Month	(3) Died $<$ 1 Year, $>$ 1 Month
Neighbouring Fraction of GreenCases	0.00217	-0.00402	0.00728**
	(0.00529)	(0.00386)	(0.00334)
Case Dummy	0.00453	0.00596	-0.00113
	(0.00491)	(0.00424)	(0.00306)
District-years with no cases	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes
Covariates	Yes	Yes	Yes
Clustering	District	District	District
K-P First Stage F	227.6	227.6	227.7
Adj. R2	0.223	0.161	0.102
N	21143	21143	21001

Note: Mortality indicators for each district are ”matched” with pollution cases in upto five neighboring districts (averaged). Cases in a neighboring district are defined as green case if the median reader classified them as either having a ”mild positive impact” or a ”strong positive impact” (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates for the neighboring district are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table 10: Do Cases have an Impact on the State Level?

	(1) ln(COD)	(2) ln(BOD)	(3) ln(TCOLI)	(4) ln(Conductivity)	(5) ln(Temperature)
Fraction of GreenCases per State	-0.190	-0.187*	0.330	-0.0742	-0.00847
	(0.119)	(0.103)	(0.360)	(0.102)	(0.0146)
Case in State	-0.0252	0.0218	-0.0768	-0.0379	0.00225
	(0.0681)	(0.0515)	(0.122)	(0.0465)	(0.00765)
Case in District	0.217***	0.0888*	0.236*	0.0633	-0.000291
	(0.0737)	(0.0532)	(0.131)	(0.0738)	(0.0114)
District-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes
Clustering	District	District	District	District	District
K-P First Stage F	145.7	137.8	133.2	178.5	179.7
Adj. R2	-0.0798	-0.0470	-0.0491	-0.0481	-0.0469
N	3049	5619	5055	5446	5510

Note: Rewrite

A Appendix Tables

Table A1: Summary Statistics

Cases and District Controls– Pollution Working Sample					
Appeal	516	0.25	0	0	1
Constitutional	516	0.21	0	0	1
Government is Respondent	516	0.82	0	0	1
Number of Judges	516	2	1	0	3
Environmental Impact (Median Coding)	516	0.34	1	-2	2
Maximum Forest Cover	288	24.14	15	4	69
Total Forest Cover	288	62,604.25	328,253	161	2,097,769
Maximum Nightlights	178	16.12	17	1	63
Total Caliberated Nightlights	178	4,446.92	16,898	3	88,983
Cases and District Controls–Mortality Working Sample					
Appeal	777	0.25	0	0	1
Constitutional	777	0.22	0	0	1
Government is Respondent	777	0.86	0	0	1
Number of Judges	777	2	1	0	3
Environmental Impact (Median Coding)	777	0.35	1	-2	2
Maximum Forest Cover	552	25.61	16	1	73
Total Forest Cover	552	63,453.08	290,488	131	2,737,216
Maximum Nightlights	329	22.78	23	0	63
Total Caliberated Nightlights	329	11,899.85	32,118	1	261,839

Note: Summary statistics of cases and control variables that were featured in the two working samples.

Table A2: Pollution Regressions with District-Level Control Variables

	(1) ln(COD)	(2) ln(BOD)	(3) ln(TCOLI)	(4) ln(Conductivity)	(5) ln(Temperature)
Fraction of Green Cases	-0.513**	-0.270*	-0.262	-0.210	-0.0470
	(0.232)	(0.153)	(0.421)	(0.173)	(0.0309)
Dummy for Presence of a Case	0.149	0.115	-0.341	-0.172	0.0255
	(0.172)	(0.174)	(0.216)	(0.123)	(0.0653)
District-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes
Clustering	District	District	District	District	District
K-P First Stage F	183281465.2	240.0	159.4	251.6	189.0
Adj. R2	-0.230	-0.0910	-0.102	-0.0967	-0.0977
N	961	2126	1852	2077	2073

Note: Cases are defined as green case if the median reader classified them as either having a ”mild positive impact” or a ”strong positive impact” (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual (LSA) features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table A3: Yearly Pollution Regressions with LSA Text features

	(1) ln(COD)	(2) ln(BOD)	(3) ln(TCOLI)	(4) ln(Conductivity)	(5) ln(Temperature)
Fraction of Green Cases	-0.117	-0.215*	0.576	-0.0575	-0.0254
	(0.121)	(0.121)	(0.386)	(0.139)	(0.0216)
Dummy for Presence of a Case	0.235**	0.0536	-0.0350	-0.0742	0.00141
	(0.113)	(0.109)	(0.212)	(0.123)	(0.0281)
District-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes
Clustering	District	District	District	District	District
K-P First Stage F	143.9	65.54	67.33	34.43	51.62
Adj. R2	-0.0800	-0.0446	-0.0562	-0.0474	-0.0466
N	3053	5649	5057	5475	5541

Note: Cases are defined as green case if the median reader classified them as either having a ”mild positive impact” or a ”strong positive impact” (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual (LSA) features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table A4: Yearly Pollution Regressions with 3 year Moving Averages

	(1) ln(COD)	(2) ln(BOD)	(3) ln(TCOLI)	(4) ln(Conductivity)	(5) ln(Temperature)
Fraction of Green Cases	-0.193***	-0.211**	0.00642	0.0302	-0.0405**
	(0.0725)	(0.0948)	(0.363)	(0.113)	(0.0189)
Dummy for Presence of a Case	0.179**	0.0749	0.274	-0.0416	0.00524
	(0.0717)	(0.106)	(0.216)	(0.112)	(0.0273)
District-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes
Clustering	District	District	District	District	District
K-P First Stage F	58.54	62.57	58.93	62.43	62.08
Adj. R2	0.663	0.623	0.629	0.690	0.519
N	5742	6254	5888	6237	6185

Note: DESCRIBE POLLUTION Cases are defined as green case if the median reader classified them as either having a ”mild positive impact” or a ”strong positive impact” (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table A5: Yearly Pollution Regressions based on Aggregated Monthly Pollution

	(1) ln(COD)	(2) ln(BOD)	(3) ln(TCOLI)
Fraction of Green Cases	-0.0121	0.0583	0.860**
	(0.118)	(0.0999)	(0.367)
Case Dummy	0.171*	0.0420	-0.128
	(0.0931)	(0.113)	(0.266)
District-years with no cases	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes
Covariates	Yes	Yes	Yes
Clustering	District	District	District
K-P First Stage F	103.9	88.53	187.1
Adj. R2	0.000919	-0.00938	-0.0108
N	30955	34677	30871

Note: DESCRIBE POLLUTION Cases are defined as green case if the median reader classified them as either having a ”mild positive impact” or a ”strong positive impact” (see text for more details.) Fraction of Green Cases is equal to 0 if there a no cases in a district-year. Included covariates are the district-year means of case characteristics such as whether the government is a respondent and if it is an appeal and or a constitutional case. Fraction of green cases is instrumented for by the district-year means of 25 textual features representing the writing style of judges. Robust standard errors are clustered at the district level.

Table A6: Neighbouring Districts Without Big Cities

	(1) ln(COD)	(2) ln(BOD)	(3) ln(TCOLI)	(4) ln(Conductivity)	(5) ln(Temperature)
Neighbouring Fraction of GreenCases	-0.273**	-0.0159	-0.121	-0.0684	-0.0159
	(0.106)	(0.0861)	(0.342)	(0.0963)	(0.0153)
Case Dummy	0.227**	0.00264	0.0459	-0.192	-0.0291**
	(0.101)	(0.0847)	(0.190)	(0.126)	(0.0123)
District-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes
Clustering	District	District	District	District	District
K-P First Stage F	227.0	122.5	111.4	117.6	125.8
Adj. R2	0.562	0.585	0.559	0.693	0.462
N	2908	5383	4810	5219	5282

Note: Rewrite

Table A7: Mortality Regressions Weighted

	All India Post 1989			Ganga Basin Post 1989			Ganga Basin Post 1997		
	(1) died1y	(2) died1m	(3) died1yc	(4) died1y	(5) died1m	(6) died1yc	(7) died1y	(8) died1m	(9) died1yc
Fraction of Green Cases	-0.00241	-0.00241	-0.0000133	-0.00314	-0.00163	-0.00161	-0.00314	-0.00163	-0.00161
	(0.00277)	(0.00225)	(0.00161)	(0.00335)	(0.00275)	(0.00197)	(0.00335)	(0.00275)	(0.00197)
Case Dummy	0.00365	0.00318	0.000496	0.000437	-0.00244	0.00298	0.000437	-0.00244	0.00298
	(0.00315)	(0.00270)	(0.00138)	(0.00921)	(0.00622)	(0.00588)	(0.00921)	(0.00622)	(0.00588)
District-years with no cases	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied	Dummied
Year and District FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Clustering	District	District	District	District	District	District	District	District	District
K-P First Stage F	229.4	229.4	229.4	530.2	530.2	530.2	530.2	530.2	530.2
Adj. R2	0.366	0.285	0.231	0.296	0.214	0.209	0.296	0.214	0.209
N	8322	8322	8322	3993	3993	3993	3993	3993	3993

Note: Rewrite