

Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India*

Esther Duflo, Michael Greenstone, Rohini Pande and Nicholas Ryan

September 21, 2012

Abstract

In a wide range of markets, from environmental and corporate audits to credit ratings, third-party entities report to the firms on which they report. We use a two-year field experiment on environmental audits of industrial plants to show that the resultant conflict of interest can corrupt information provision and undermine regulation. In treatment plants, auditors were randomly assigned to plants, paid from a central pool of funds and their reporting was itself audited (backchecked) on a random basis. In contrast, control plants directly chose and paid auditors. Backcheck data and an endline pollution survey show that monitoring and incentives strongly influence auditor and firm behavior. Relative to auditors assigned to treatment plants, auditors chosen by control plants are much more likely to falsely report plants as just meeting regulatory limits. And after two years of changed auditing practices, we observe fewer large polluters and, therefore, reduced pollution levels among treatment plants relative to the control group.

*We thank Sanjiv Tyagi, R. G. Shah, and Hardik Shah for advice and support over the course of this project. We thank Pankaj Verma, Eric Dodge, Vipin Awatramani, Logan Clark, Yuanjian Li and Nick Hagerty for excellent research assistance. We thank the Sustainability Science Program (SSP), the Harvard Environmental Economics Program, the Center for Energy and Environmental Policy Research (CEEPR), the International Initiative for Impact Evaluation (3ie), the International Growth Centre (IGC) and the National Science Foundation (NSF Award #1066006) for financial support. Ryan thanks an AMID Early Stage Researcher Fellowship for financial support. All views and errors are solely ours.

I Introduction

In a wide range of settings, third-party auditing is used to monitor the compliance of firms with regulations. Third-party audits are the norm in financial regulation, as public companies typically have to file independently audited annual financial statements and, in many countries, credit ratings agencies serve an important regulatory role (White, 2010). Independent audits play a large role in several consumer and commodity markets, including those for food safety and healthcare, flowers, timber and many durable goods (Hatanaka et al., 2005; Reynolds et al., 2007; Dranove and Jin, 2010). Third-party auditors also verify firm compliance with domestic and international environmental regulations.¹

The structure of third-party audit markets may create a conflict of interest for auditors between providing credible reports and maintaining business with their clients, corrupting information provision and undermining regulatory goals. Events brought to light by the recent financial crisis, suggest that this is a real possibility. Observational evidence attributes the poor quality of credit ratings during the subprime boom to financial incentives faced by rating agencies (Griffin and Tang, 2011; Strobl and Xia, 2011).² Yet, despite calls for reforming the third-party audit system to increase the independence of financial auditors, regulatory reforms to date have not addressed the core problem that firms, when they directly hire their auditor, can shop for a favorable report.³

In this paper, we provide causal evidence on how the conflict of interest can undermine information provision by third-party auditors and enforcement of regulations against the firms

¹Third party auditing is also used to enforce international environmental standards, including ISO 14001 certification and verification of carbon abatement in the carbon offset market (Potoski and Prakash, 2005; Bhattacharyya, 2011).

²80% of collateralized debt obligations originally rated AAA by Standard & Poor's in 2005-2007 were downgraded to below investment grade by mid-2009 (International Monetary Fund, 2009). Griffin and Tang (2011) use data for a single credit agency and find that the agency's internal surveillance team's judgements on CDO ratings were more accurate than the business-oriented ratings team's and that the difference between these measures predicts future downgrades. Strobl and Xia (2011) compared ratings for the same companies provided by two credit rating agencies, where one agency uses a issuer-pay model and the other an investor-pay model. The difference in ratings across agencies is more pronounced when the issuer-pay rating agency plausibly has more business at stake. For broad overviews of the corporate audit and credit ratings markets see Ronen (2010) and White (2010), respectively.

³In 2002, the Sarbanes-Oxley Act overhauled the regulatory structure for corporate audits, and made auditors of public companies subject to oversight by a private-sector, nonprofit corporation. This corporation determines who can perform audits, conducts investigations and sets fines. Three former SEC Chairmen testified in favor mandatory auditor rotation, which was not adopted. In 2003 the Securities and Exchange Commission adopted rules on auditor independence that focused on restrictions on and disclosure of non-audit activities. In 2008 New York State Attorney General Cuomo reached an agreement with credit rating agencies which required upfront payment for their ratings. The Dodd-Frank financial reform bill and the Sarbanes-Oxley Act restrict the services that auditors or credit rating agencies can offer firms that they audit.

they audit. We also show how a simple package of reforms to promote auditor independence can restore accuracy and improve regulatory compliance. Our evidence comes from a two-year field experiment conducted in collaboration with the environmental regulatory body in Gujarat, India. Since 1996, Gujarat has had a third-party audit system for plants with a high pollution potential. Every year auditors submit plant pollution readings to the regulator (Gujarat High Court, 1996).⁴ The basic financial arrangement underlying these audits is the typical practice the world over—plants hire and pay auditors directly.

Our experiment covered all 473 audit-eligible plants in two populous and heavily polluted industrial regions of Gujarat. In each region, half the plants were randomized into treatment. A treatment plant was randomly assigned an auditor who was paid centrally, and was, therefore, financially independent of the plant audited. For each auditor we independently verified a random sample of pollution readings (henceforth we call these super-audits backchecks). While the probability of being back-checked was public knowledge, actual backcheck visits by an independent technical agency were unannounced. During the backcheck, the agency collected pollution readings for the same pollutants at the same sampling locations as were reported in audits. In the first year, backcheck data was used to measure auditor accuracy but no consequence of poor performance was explicitly specified. At the start of the second year, treatment auditors were informed that their pay would be linked to their reporting accuracy, as measured by the backchecks. For data analysis purposes we collected similar “backcheck” data for a random sample of control plants during the final reporting season. Finally, at the end of the second year we conducted an independent endline survey of pollution outcomes in all treatment and control plants.

We have three main findings. First, auditors chosen by control plants systematically underreport pollution readings, relative to the truth as measured by backchecks. The average difference between audit and backcheck pollution readings across reported pollutants is -0.30 standard deviations in the control group, indicating a significant negative bias. In particular, auditors for control plants incorrectly report many pollution readings as compliant. In the control group, an audit report is 29 percentage points more likely to show a plant as compliant than a corresponding backcheck of the same pollutant.

Second, the treatment essentially eliminates this bias. Treatment auditors submit much

⁴In addition to pollution readings, the report also describes the production process and physical state of the plant, including measures taken for pollution control.

higher reports. Auditors working in the treatment group improve the accuracy of their reports to the extent that, by the end of the experiment, the mean audit report is statistically no lower than the mean corresponding backcheck. This large improvement in auditor reporting is evident in specifications with auditor fixed effects that compare the behavior of the same auditors working in the treatment and control plants. Auditors working in treatment plants reduce the share of plants falsely reported as compliant by 23 percentage points against the base of 29.

Third, plants respond to greater audit scrutiny under the treatment by reducing pollution, as measured by the endline survey. The decline in actual pollutant concentrations due to the treatment is roughly the same size as the change in auditor reporting of pollutant concentrations, which is to say rather large. Pollution declines most for the water pollutants subject to greatest regulatory scrutiny.

These results support a key insight offered by the literature on market behavior of information intermediaries, that competition can lead to inaccurate information provision (Dranove and Jin, 2010; Bolton et al., 2011).⁵ More broadly, our findings suggest an economic environment in which plants shop for auditors willing to report compliance, and auditors balance payments by these plants and the continuation value of leniency in reporting against the risk of being disbarred by the regulator. By directly assigning auditors to plants and paying them from a central pool our treatment ended the ability of plants to shop for auditors willing to provide false reports. This, in turn, reduced the value for auditors of leniency in reporting for repeat business. The introduction of back-checks increased the probability of false reports being detected. Both changes increased auditors' incentives to report truthfully. We use non-experimental variation in financial incentives, from the introduction of performance pay in the second year, to show that this component of the treatment package independently contributed to increased truth-telling. The treatment increased the cost for plants of filing a falsely low report which improves the relative cost of complying with regulation by abating pollution.

Further evidence on the importance of incentives is provided by evidence that auditor and plant behavior is sensitive to anticipated penalties. Specifically, we use administrative records

⁵Bolton et al. (2011), for instance, show that competition among Credit Rating Agencies can lead to rating shopping and that rating inflation is particularly likely when part of the payment occurs after rating is completed, as in our setting.

of actions taken by GPCB to measure pollutant-specific anticipated penalties. Auditors falsely report the most for pollutants which are most likely to trigger a sanction for exceeding the regulatory standard. While truth-telling is higher for treatment plants, we continue to observe more false compliance for pollutants with higher anticipated penalties of non-compliance. Moreover, plant reductions in pollution occur only in the right tail of the water pollutant distribution, where the regulator is most likely to impose costly sanctions.

Our findings contribute to several literatures. Corruption in developing countries and the growth and welfare costs of corrupt behavior have been widely documented (Olken and Pande, 2012). A growing body of evidence has shown that, even in settings with weak enforcement institutions, reforms that strengthen financial incentives and increase monitoring can improve compliance with regulations (Olken, 2007; Duflo et al., 2012). Enforcement of environmental regulations is an important area for social welfare as a growing literature documents high levels of pollution in emerging economies and large costs of this pollution, including lower labor productivity, higher infant mortality and lower life expectancy (Hanna and Oliva, 2011; Hanna and Greenstone, 2011; Chen et al., 2010). Of course, our findings are valid in the specific context of the reform evaluated and may not apply to environmental regulation in other countries or to other sectors, such as finance. That caution notwithstanding, this paper presents remarkably clear evidence on one appropriate structure of third-party auditing as a regulatory tool. Altering economic incentives caused auditors to switch from biased reporting to truth-telling in a very short span of time, and this change led to an improvement in real outcomes, here pollution. To our knowledge, there is no comparable evidence in the prior literature for any sector.

The remainder of the paper is organized as follows. In Section II, we describe the background and experimental design. In Section III we model the behavior of polluting plants and environmental auditors and lay out expected responses to the experiment in terms of model parameters. In Section IV we discuss our data collection and in Section V present the results. We conclude and discuss some implications for policy in Section VI.

II Background and Experimental Design

A Study Context

Our study occurred in Gujarat, one of India's fastest growing industrial states, having attracted the largest share of investment after licensing reform of any state in India (Chakravorty, 2003). Since 1991-1992, the peak of industrial licensing reform, net state domestic product has grown at an average of 8% per year. Gujarat has about 5 percent of the Indian population, but accounts for 9 percent of India's registered manufacturing employment and 19 percent of output (Authors' calculation, Annual Survey of Industries, 2004-05). This industrial growth has, however, been accompanied by a degradation of air and water quality. Eight industrial clusters are categorized as critically polluted, tied for the most of any state, and including the two most polluted in the country (Central Pollution Control Board, 2009b). Gujarat contains three of India's five most polluted rivers (Central Pollution Control Board, 2007). Essentially all large cities in the state, as well as some industrial areas, are classified as non-attainment with respect to the National Ambient Air Quality standards for Respirable Suspended Particulate Matter (RSPM) (Central Pollution Control Board, 2009a).

Such high levels of industrial pollution persist despite a stringent regulatory framework for pollution control, as set by national and state laws and court orders (Hanna and Greenstone, 2011).⁶ Most decisions on environmental regulations and all enforcements occur at the state-level. State Pollution Control Boards are responsible for enforcing the provisions of the Water Act and the subsequent Air (1981) and Environmental Protection (1986) Acts and their attendant regulations. In Gujarat, the Gujarat Pollution Control Board (GPCB) enforces these traditional command-and-control pollution regulations. GPCB is responsible for monitoring and regulation of approximately 20,000 plants.

B Environmental Audit Regulation

The main instruments that GPCB uses to limit industrial pollution are plant-level inspections and environmental audits. This paper focus on the environmental audit system.

In 1996, in order to remedy the perceived failure of inspections in enforcing pollution standards, the High Court of Gujarat introduced the first third-party environmental audit

⁶The Water (Prevention and Control of Pollution) Act of 1974 created the Central Pollution Control Board as a coordinating body to set pollution standards and the state boards as enforcement agencies.

system in India (Gujarat High Court, 1996). Under the scheme, plants with high pollution potential must submit a yearly environmental audit, conducted by an audit plant hired and paid for by the plant. Audit-eligible plants are categorized as Schedule I (most polluting) or Schedule II (less polluting). The classification of pollution potential is done on three dimensions: what product the plant produces, where it sends its waste effluent and the volume of that effluent.⁷

Schedule I plants must be audited by Schedule I auditors, usually an engineering college or similar institution. Schedule II plants must be audited by a private audit plant, called a Schedule II auditor. This study concerns only the reporting of Schedule II auditors and henceforth we refer to plants in Schedule II as “audit eligible.” Auditors visit each plant three times per year for about one day, observe the plant’s environmental management practices and measure pollution outputs. Auditors compile the results of these visits and submit their audit report to GPCB by February 15th of the following year. The format and scope of the audit report are fixed by the environmental audit regulation. The final audit report, which is sent to the plant and GPCB, describes the production process and physical state of the plant, including the measures the plant has taken for pollution control and the results of pollution sampling during each of the visits. Finally, auditors provide recommendations to the plant.

The basic structure of the environmental audit system includes various safeguards and strong penalties. Most of this structure was explicitly laid down by the Court in its original 1996 ruling, though some features were developed by GPCB in response to court order. Each four-member audit team is certified by the regulator and must be recertified every two years.⁸ Each audit team is limited to auditing at most 15 plants per year and an audit firm, which may contain several teams, can audit a plant at most three years in a row. In principle, non-submission of an audit is punishable by closure and disconnection of water and electricity connections. In practice, as we will see, some plants do not submit reports, usually arguing that they are not actually eligible. A report showing non-compliance with the terms of a plant’s environmental consent can also be punished by closure or payment of compensation

⁷For example, plants that produce certain types of dyes and dye intermediates are classified in Schedule II, roughly, if their effluent is between 25 and 100 thousand liters per day, with variations around this classification based on whether the effluent discharged by the plant goes on to further treatment in a common effluent treatment plant (CETP). A plant with effluent below 25,000 liters would be exempt from the audit requirement.

⁸The team members are required to have degrees in environmental engineering, chemical engineering, chemistry and biology. and at least two members must have at least one year’s experience in environmental management.

(Gujarat High Court, 1996). Moreover, auditors with reports found to be inaccurate are liable to be de-certified and their reports on behalf of other plants declared void, an incentive to build a reputation for quality.

C Environmental Audit Market

The audit system, as originally implemented, is widely perceived to function poorly. Industry recently litigated against the scheme, somewhat ironically and without success, to get the High Court of Gujarat to throw out the audit requirement on account of GPCB not following up on audit reports (Gujarat High Court, 2010). The regulator, for its part, believes that industry's use of auditors who are willing to file false reports causes submitted reports to be unreliable. Review of the auditors by the GPCB is mostly *pro forma*.

Consistent with the idea of auditor shopping, the environmental audit market is characterized by strong price competition. Informally, during interviews conducted prior to the experiment both auditors and plants claimed that an audit report could be purchased for as low as INR 10,000-15,000. We collected information on audit prices after the experiment from both auditors and plants. Conditional on reporting any payment, plants in the treatment and control report very similar average payments of around INR 24,000. Auditors report a mean minimum plant willingness-to-pay of INR 25,000 and a mean maximum of INR 52,000.⁹ Using GPCB pollution sampling and analysis charges, we estimated that the cost of performing a thorough audit should be between INR 20,000-40,000 depending on the plant's characteristics. As about 80% of our audited sample are textile plants, for which the cost basis of an audit is around INR 40,000, we take the sum of evidence on audit prices as showing that a significant portion of audit reports are being purchased at prices below the cost of an audit duly performed. Many of the best-reputed auditors (as identified by GPCB) reported that they had scaled back or closed their audit businesses to focus on environmental consulting.

D Experimental Sample

Our sample is the population of all audit-eligible plants in the GPCB regions of Ahmedabad and Surat, the two largest cities of Gujarat. We obtained from GPCB a list of all 2,771 red

⁹The fact that the mean reported by auditors is higher than the mean reported by plants is consistent with the fact that low-cost auditors audit more plants. The minimum reporting here may be more reliable as auditors may have expected answers given at this conference to be accounted for in future audit reforms.

category (i.e., high pollution potential) plants with reported capital investment less than INR 100m (about USD 2.2m), which were designated small or medium scale, and selected the 633 audit-eligible plants as a provisional sample. We randomly assigned half of the plants within this provisional sample, stratified on region, to the audit treatment group. After the randomization, we collected the required detailed sectoral information and eliminated plants found to be ineligible from both treatment and control, reaching the study sample of 473 plants, 49.2% of them belonging to the treatment group.

Table 1 presents summary statistics for plants in the study sample that submitted an audit report in either year of the experiment using administrative baseline data from GPCB records. We condition the sample on audit submission as most analysis of auditor reporting uses data from submitted audit reports. Plants submitting audits in the audit treatment and control look very similar. Panel A reports plant characteristics. Most sample plants are textile factories eligible for environmental audit due to having high volumes of waste effluent. Textiles is the largest registered manufacturing sector by employment in India and second largest in Gujarat (Authors' calculation, ASI 2005). Both treatment and the control plants have similar pollution potential as measured by effluent quantity and type of fuel used. Treatment plants are 10 percentage points less likely to have a bag filter, a type of air pollution control equipment, installed, but are similar to control plants with respect to other equipment such as cyclones and scrubbers. In the same comparison of covariate balance for the full study sample, unconditional on submission (not shown), bag filter installation is also the only significant difference between plants assigned to the treatment and control groups.

Table 1, Panel B reports various interactions of sample plants with GPCB in the year prior to study by treatment status. First, a little over 80% of this group, which was selected for being audit-eligible and submitting at least once in year 1 or year 2 of the study, submitted an audit report in the prior year. Roughly the same fraction was inspected and over forty percent of sample plants were mandated to install equipment. Second, based on the inspection reports a significant number of sample plants were subject to costly regulatory actions: Around a quarter were cited for any type of violation and fully ten percent of plants, in both treatment and control, had their utilities disconnected at least once. About three percent posted a bank guarantee (i.e., bond) against future environmental performance. These variables are balanced across treatment and control plants. Consistent with being less likely to have a

bag filter, which abates air pollution, treatment plants were more likely to have received a citation for an air pollutant violation than control but equally likely for water pollutants and all citations together. In summary, our experiment occurred in a setting where information reported by environmental auditors should matter to plants as the regulator has a meaningful track record of action and could employ audit reports in the same way.

E Experimental Design

We evaluated a modified audit system, in collaboration with the GPCB, which sought to increase auditor’s financial independence and the accuracy of auditor reporting.¹⁰ Treatment plants were assigned to the audit treatment once, in 2009, for the audit years 2009 (hereafter year 1) and 2010 (year 2). We made three changes to the existing audit system: random assignment of auditors to plants, payments to auditors from a central pool, and back-checks of auditor performance. In year 2, direct incentive pay for auditor accuracy was added to the basic set of three reforms. We now discuss the components of the audit treatment in turn.

Random Assignment and Financial Independence In the treatment group, auditors were randomly assigned to plants by the research team rather than being selected by the audited plants. Auditors working in the treatment plants were paid from a central pool of funds raised for the study. The payment was fixed (INR 45,000 in the first year, and depending on accuracy in year 2). This rate was estimated by applying GPCB’s own sampling charges to the average plant characteristics in the audit sample and adding a small margin: it should thus represent the average cost of completing an audit. However it was in the high range of the market price.¹¹

In terms of auditor participation, all GPCB-certified Schedule II auditors were solicited for their interest to participate in the treatment separately in each year. In both years interest was oversubscribed, relative to the number of treatment plants, and so at the beginning of each year auditors were randomly assigned to a number of plants allocated proportional to

¹⁰This experiment was designed and undertaken concurrently with the evaluation of another intervention, an increase in inspection frequency for some plants, which was conducted stratified on the audit treatment and which we study in a separate paper

¹¹Paying a higher than market price for the audits could be considered to be a fourth component of the intervention in its own right. However, arguably the level of payment is part of a plant’s strategy and is endogenous to market structure. Moreover, a high payment was not in itself sufficient to ensure the audit actually performed the audit: they could easily have pocketed the difference. Finally, our analysis shows responsiveness of treatment to expected penalties – if a higher wage simply reduced shirking then we would not expect auditor actions to remain sensitive to expected penalties.

their capacity, measured in number of certified audit teams.

The strong auditor interest in the program may reflect both that treatment audit payments, were in the high range of the market and that the audit treatment offered auditors better working terms. Auditors remained free to use auditing capacity not allocated in the treatment to conduct audits in the control group. This aspect of the design allows measurement of the effect of the audit treatment on auditor reporting within the same set of audit plants. In the first year, out of 42 auditors, 24 worked in control only, 9 in treatment only, and 9 in both. In 2010, out of 34 auditors, 7 worked in control only, 12 in treatment only, and 15 in both.¹²

Backchecks A randomly selected 20% of auditor plant readings in the treatment were backchecked in the field by technical staff of independent engineering colleges (which were certified as Schedule I auditors). Auditors were aware of the possibility of being backchecked and aware that backcheck results would be used for quality control, although in the first year the terms of any sanction for poor performance were left somewhat vague and auditors were paid the same amount regardless of accuracy. The results of backchecks where not were only sent to GPCB as part of aggregated reports on the accuracy of auditors, and therefore carried no consequence for audited plants. In season 3 of year 2, 20% of control group audits were also backchecked in similar conditions. That data was not sent to GPCB, but used only for the evaluation.

Incentive Pay In year 2 incentive pay for auditor accuracy, as measured by backchecks in the treatment, was added to the basic set of reforms. Incentive payments used a formula that was first applied to year 1 to demonstrate to each auditor how accuracy was measured. The pay formula first calculated the difference between audit report pollution concentration readings and backcheck readings, standardized by the variability of backcheck readings, for each pollutant. It then averaged the scores for six water and three air pollutants into indices for each media and created the overall measure of auditor quality as the average of these two. An index of zero thus means that an auditor always reported readings that matched backchecks exactly and of one that the average auditor and backcheck readings were one standard deviation apart. Auditors were grouped into three bins of the least accurate quartile,

¹²The increase of auditors working in treatment only or both in the second year comes from the fact that, in the first year, some auditors were not able to participate because they had already reached their capacity when the program was announced.

paid Rs. 35,000 per audit, the next least accurate quartile, paid Rs. 40,000 per audit, and the most accurate half, paid Rs. 52,500 per audit, maintaining the average pay of the prior year.

III Model

This section presents a simple model to clarify the incentives faced by auditors and plants. We discuss the setting, the solution and the relation of the model to our experimental treatments.

A Set-up

We consider a three-stage game between a plant, an auditor and the regulator. In the first stage, plants observe their abatement cost and set a payment schedule for their auditor. Plants produce output with pollution by maximizing the objective

$$S - \mathbf{1}\{\text{Abate}\} \cdot c - w - \mathbf{1}\{\text{Fine Imposed}\} \cdot f.$$

where S is the value of production, c is the cost of abating pollution, w is the amount paid to environmental auditors and f is any fine imposed by the regulator for violations of pollution standards. All plants pollute at a level p_h , above the regulatory limit, with no abatement. By paying an abatement cost c plants can reduce their emissions to p_l , below the limit. Plants draw c from a cumulative distribution function G .

Plants make a joint decision about whether to abate pollution and what schedule of payments to offer auditors in order to minimize their overall cost of production. We assume that S is sufficiently large that environmental regulation does not cause plants to shut down. The payment $w(\hat{p}|p)$ offered auditors can be a function of both the plant's true pollution level p and the auditor's report \hat{p} on the pollution level. Let the notation $w_{lh} = w(\hat{p} = p_l|p = p_h)$ and similarly for other conditional payments. The goal of the plant when offering payments is to minimize the sum of the cost of an audit and any potential fine.¹³

Auditors observe the pollution level of plants and report on plant pollution to the regulator. The payoff to an auditor is given by

$$w - a + \mathbb{E}V(\hat{p}|p)$$

¹³We assume not filing an audit is very costly and so is not considered as an action. In practice not filing is similar to reporting $\hat{p} = p_h$.

where w is the payment for performing an audit this period, a is the cost of preparing an audit and the final term is the expected value of future audit business. The continuation value of future audit business depends on audit reports \hat{p} through an exogenous function V .¹⁴ We assume leniency increases auditors' continuation value, so that $V_{lh} = V(p_l|p_h) > V_{hh} = V_{hl} = V_{ll} = V_0$ where the base continuation value V_0 is the same for all other conditional reports. Define $\Delta V = V_{lh} - V_0$ as the incremental continuation value of falsely-low reporting, i.e. the benefit of notoriety.¹⁵ The expected value depends on the likelihood that an auditor will be disbarred if found to be inaccurate.

We model the audit market as competitive but assume that auditors nonetheless receive some surplus from participating in the market. The marginal cost of an audit may be interpreted as including some rent to being certified as an auditor or the opportunity cost to auditors of doing audits, which can employ their staff in environmental consultancy roles if auditing is not profitable.

Finally, the regulator reviews audit reports \hat{p} and may impose fines on plants and disbar inaccurate auditors. With probability q the regulator observes the true state of pollution at a plant, either through the reports of its own staff or backchecks of auditor performance.¹⁶ The regulator decides whether to impose a fixed fine f based on the audit report and whether to disbar auditors that report falsely. The continuation value for a disbarred auditor is zero. As the focus of the paper is on auditor behavior we take a simple view of the regulator as seeking only to minimize pollution.

B Equilibrium, and response to Experimental Treatments in the Model

We restrict attention to equilibria with regulation, defined as equilibria in which the regulator imposes fines when p_h is observed and always disbars auditors reporting falsely.¹⁷ Moreover,

¹⁴We think of V as measuring the effect of auditor reputation *with plants*, though strictly auditors are homogeneous and have no type for client or other plants to infer.

¹⁵Bolton et al. (2011) and many other models of information intermediaries also take continuation values as exogenous and, moreover, as wholly independent of current actions. In our set up the continuation values may have been affected by the treatment, as leniency is less valuable in a system with random assignment of auditors. Since the experiment was conducted for only two years, changes in continuation values were likely to be small and we do not emphasize them in the model.

¹⁶We assume the regulator does not use this direct observation to fine plants, as backchecks are not legally valid grounds to impose fines. Alternatively, plant abatement costs may incorporate the baseline likelihood of being penalized by the regulator directly.

¹⁷There are equilibria without regulation but we believe that equilibria with regulation are more likely to occur given that the regulator is relatively long-lived. If the game were dynamic the regulator would be considered a long-lived player against short-lived auditors and plants. In the manner of Kreps and Wilson (1982), it would then select its most preferred equilibrium, in which some plants abate, despite moving last.

we observe frequent sanctions for high pollution in practice, as will be discussed in Section V.

Proposition 1. *There is a unique subgame-perfect equilibrium with regulation in which one of three outcomes obtains. Firms clean up if $c \leq \min\{f, qV_{lh} - \Delta V\}$, underreport if $qV_{lh} - \Delta V < \min\{c, f\}$ and defy the regulation if $f < \min\{c, qV_{lh} - \Delta V\}$.*

The Model Appendix, Section XI, derives the equilibrium, and Figure 1 provides a graphical illustration. The plant minimizes costs across one of three strategies. Most simply, if abatement cost is low the plant can abate to p_l and pay auditors at cost for giving a truthful report. If abatement cost is high, the plant can choose between buying a truthful, high report and incurring a fine and buying a falsely low report. In any equilibrium with regulation, inducing low reporting requires $w_{lh} \geq w_{hh} - \Delta V + qV_{lh}$, as the plant has to compensate the auditor for the risk of disbarment qV_{lh} that comes with a false report. The plant gets a discount on false reporting due to the reputational benefit ΔV of notoriety to auditors.

Baseline case. In the *status quo* audit system (as seen in the control group), the probability q of audit reports being verifiable was very low. As $q \rightarrow 0$ the cost of buying falsely low reports $qV_{lh} - \Delta V$ will be negative and thus less than strictly positive abatement costs and fines. We anticipate pervasive auditor underreporting and, as $G(qV_{lh} - \Delta V) = 0$ for small q , no plants will abate in response to audits.

Backchecks. We model the backcheck treatment as raising q , the probability an audit report is verifiable.

Proposition 2. *A raise in q from the baseline sufficient for $qV_{lh} - \Delta V > f$ induces a share $G(f)$ of plants to abate.*

The stated rise in q is enough that polluting plants would prefer incurring a fine to paying the auditor a premium for false reporting. Given such a rise, the relevant margin of decision for a plant is whether to abate or to incur a fine. The $G(f)$ share of plants with costs beneath the fine abate. The necessary rise in q is smaller the larger is the continuation value of auditing business for auditors, and larger is the benefit of notoriety.

Incentive Pay. The incentive pay treatment augments the auditor continuation value for accurate reports so that inducing low reporting is more costly than a fine if $q(V_{lh} + B) - \Delta V > f$, where B is the incentive pay bonus.

Assignments. We model the assignment of auditors to plants as plants required to offer payments unconditional on the contents of the audit report, so $w(\hat{p}|p) = w(p)$. Thus, plants cannot refuse to pay an auditor to which they were assigned if the assigned auditor submits a high report.¹⁸ Assignments may also reduce the incremental continuation value of notoriety, ΔV , to the extent that auditors expected that the modified audit system might persist beyond the experiment.

Proposition 3. *Under assignments, a rise in q from the baseline sufficient for $qV_{lh} \geq \Delta V$ induces a share $G(f)$ of plants to abate.*

With a flat fee auditors now consider only their reputations when deciding whether to report falsely. This makes the required raise in q smaller than in the case with conditional payments. If assignments reduce V_{lh} and thus $\Delta V = V_{lh} - V_0$, then the q required for abatement is smaller than if ΔV does not change. Both channels through which assignments may work push in the same direction. The consequence for abatement still depends on the distribution of plant costs.

To summarize, if the combined treatment of random assignments, central payment and backchecks is weak, auditors will falsely report and plants will not abate. If it is strong enough, in that $q(V_{lh} + B) \geq \Delta V$, then the response will be a combination of abatement, from a $G(f)$ share of plants, and higher reporting from the rest. With false reporting more costly, plants trade-off abatement and the risk of fines, and abate if abatement costs are low. In this case, the composition of the response to the experiment reveals the extent to which plants are able to abate at costs less than the fine for a detected violation.

IV Data and Empirical Strategy

A Data

The key outcomes of interest are accuracy of auditor reporting and the pollution response of client plants. Two data sources are used to measure accuracy. First, audit reports filed by auditors to GPCB in 2009 and 2010. These reports cover a mandated set of water and

¹⁸In the treatment, auditors were paid with funds raised for the study, not by plants. As a flat fee above cost will not enter into the choice between reports, the salient aspect of this change is the fact that auditors were paid for reports regardless of whether they reported violations. Firms could however still have offered to pay more below the table for a favorable report

air pollutants, described in Table 7. For most of the analysis, we focus on final-outlet concentrations for water samples, as these are the readings that have a direct impact on the environment and are therefore most closely attended by both auditors and GPCB. For air, we focus on boiler-stack samples for the widest comparability across the entire sample, as most plants have boilers.¹⁹ Air samples are somewhat more costly and time-consuming to collect than water samples.²⁰ Water pollution, in principle, may be more stable over time if a plant is treating its waste as effluent mixes together in the treatment process, in contrast to stack gas, which is instantaneously released. The variability of water pollutant readings with the plant process type and process stage, however, is greater than that for air pollution that all stem from combustion.

Not all plants in the sample submit audits, despite being *prima facie* audit-eligible, for several reasons.²¹ As shown in Table 2, treatment and control plants were about as likely to submit audit reports. Control plants were slightly more likely to report in the first year and treatment in the second year, though neither of these differences is statistically significant. A steady 70% of treatment plants submitted audit reports in each year, whereas 74% and 64% of control plants submitted audit reports in years 1 and 2, respectively. The rates of submission during the experiment are comparable to what GPCB records indicate for submission in 2008, when 72% and 69% of plants submitted in treatment and control, respectively, before the experiment began. The treatment, therefore, does not appear to have induced more plants to submit audit reports.

The second source of data for accuracy is the backchecks, which were conducted in a sample of treatment plants throughout 2009 and 2010 and in a single wave, which we call the “midline” survey, after the third season of audit visits in 2010 in both the treatment and control. Midline data allows a comparison of backcheck and actual readings, and therefore direct measurement of the comparative accuracy of auditors working in treatment and control

¹⁹Depending on the number of stages of effluent treatment, water pollutant samples may be taken during audits prior to any treatment, after primary treatment, after secondary treatment, or at the final outlet from the plant. Similarly air samples may come from stacks (chimneys) attached to industrial boilers, thermic-fluid heaters or various industrial processes, such as chemical reaction vessels.

²⁰The cost of collecting these samples can be benchmarked with the charges set by GPCB, which run from INR 85 for Total Dissolved Solids and Total Suspended Solids to INR 250 and 325 for Chemical Oxygen Demand and Biochemical Oxygen Demand. All air pollutants are charged at INR 400 and manual sampling of air pollutants in the field takes approximately half an hour per pollutant.

²¹GPCB had judged them not-audit-eligible in the past, they had changed products since the data used to determine eligibility was compiled, they protested their audit-eligible status, they closed or they simply chose not to submit and incur the risk of a penalty rather than hire an environmental auditor.

plants. Backchecks were conducted on the same pollutants and locations as the audits, and were scheduled to occur close to the initial sampling.²²

The midline sample was drawn from treatment and control groups in order to maximize the number of plants covered by auditors working simultaneously in both the treatment and control groups and to use information on the dates of audit visits to conduct backchecks that were as close as possible to the date of the initial visit. In the treatment group, the sample plants were randomly selected stratified by auditor. In the control group, the sample plants were drawn non-randomly in order to ensure coverage of auditors working in the treatment simultaneously. Priority for the survey was first given to plants that previously submitted an audit report by an auditor working in the treatment group. The control sample was completed by adding those plants for which auditors submitted a date for the audit visit and finally by adding randomly-selected plants for which auditors had not submitted a date.

The overall composition of the midline sample remains well-balanced along observables that were shown in Table 1. The 130 plants in the midline treatment and control significantly differ at the 5% level in that midline treatment plants are less likely to have a bag filter installed, as are treatment plants in the overall sample, and consume more water and generate more waste water than midline control plants. The primary outcome of interest is the accuracy of auditor reporting, which is the difference between audit and backcheck readings, and thus controls for a plant's true pollution levels. These few differences on observables at baseline would therefore need to differentially effect audit and backcheck readings to influence midline results.

The source of data on plant pollution response is an endline survey conducted from April through July of 2011.²³ The survey was mostly conducted by the same agencies that did backchecks and included the same pollution samples discussed above, in addition to economic and abatement cost measures that we do not use in this paper. The endline survey also covered plants outside the audit sample that were subject to an additional experimental treatment, discussed in another paper. We use survey data only plants in the audit sample (treatment and control group) that were not subject to the other intervention.

²²The median backcheck lag in the midline sample was 36 days. It was 6 days in the rest of the study, when backchecks were used for monitoring and incentives.

²³The mean survey date, at the end of May, was approximately six months after the surveyed plant would have had their final audit visit under the audit treatment and three and a half months after the audit reports in year two were due to the regulator.

In all cases pollutant measures are standardized by subtracting the mean and dividing by the standard deviation of the same pollutant reading amongst backcheck samples at the same sampling location.

B Empirical Strategy

Given the randomized design, the empirical strategy is straightforward. Our specifications, however, vary with outcome variable (and sample). First, to measure the impact of audit treatment on reported compliance we pool samples of pollutant readings from both audits and backchecks for plant j . We then estimate ordinary least-squares regressions of a difference-in-difference form:

$$\mathbf{1}\{Compliant\} = \beta_1 \mathbf{1}\{AuditReport\} \times T_j + \beta_2 \mathbf{1}\{AuditReport\} + \beta_3 T_j + \alpha_r + \epsilon_j, \quad (1)$$

α_r are fixed effects for the region $r \in \{Ahmedabad, Surat\}$ and treatment coefficient β_3 measures the difference in compliance, as measured by backchecks, across treatment and control. The coefficients of interest are β_2 , which measures how much more likely an audit report is than a backcheck to be compliant, and β_1 , which measures how treatment changes this difference between audit reports and backchecks. In some specifications we replace the compliance dummy with a dummy for being between 75% and 100% of the regulatory standard, as a measure of the tendency to just barely underreport.

Next, we consider reported pollutant i concentration in plant j as reported in audit reports. The advantage of these data is their availability for both years of the experiment. We estimate similar ordinary least-squares regressions:

$$y_{ij} = \beta T_j + \alpha_r + \alpha_y + \epsilon_{ij}, \quad (2)$$

α_y are fixed effects for the year $y \in \{2009, 2010\}$. We pool standardized measures for different pollutants in most specifications and cluster standard errors are at the plant level.²⁴ We report regressions without and with fixed effects α_a for the auditor a , to focus on the within-auditor difference in behavior across the two incentive systems.

One concern with using plant-level concentration of pollutants as an outcome is that they

²⁴On average, we have 7 pollutants per plant.

are a composition of the “true” level of a pollutant, measurement error and an auditor’s manipulation of the results. If the treatment affects plant behavior and therefore true level of pollution, then changes in audit reports of concentrations reflect changes in pollution and changes in reporting, and will underestimate the reporting effect if plants respond to more accurate auditing by reducing pollution levels.

We, therefore, also consider the difference between back check and audit reading as an outcome variable. This controls at the plant level for any possible effect of the intervention on actual pollution levels. These differences are available on a comparable basis only in the midline sample.

$$y_{ij}^D = y_{ij}^{Audit} - y_{ij}^{Backcheck},$$

where y_{ij}^D is the audit reading less the backcheck reading. Samples are matched on pollutant i , plant j , sampling location (boiler stack or final outlet) and date. Negative values indicate underreporting and positive overreporting. Our third outcome, from the endline survey, is actual plant-level pollution. Our regression framework is as in Equation2, omitting the year fixed effect.

Finally, to measure changes in pollution in the endline survey, we report OLS regressions (of the form given by Equation2) considering both the continuous pollution outcome and a compliance dummy as outcome variables. We also estimate quantile regressions of the form

$$Q_{y_{ij}|X_j}(\tau) = \beta T_j + \alpha_r + \epsilon_{ij}$$

where $Q(\tau)$ is the τ quantile of the pollutant concentration conditional on treatment status and regional indicator variables.

V Results

We structure our findings in three parts. First, we use the audit reports and backcheck data to examine whether the treatment altered pollution reporting by auditors and increased truth-telling. Next, we examine how existing regulatory incentives for plants and treatment-induced financial incentives for auditors influence the extent of truth-telling by auditors. We conclude by examining the impact of two years of altered auditor incentives on plant polluting behavior.

A Auditor Behavior

A.1 Reported Compliance with Regulatory Standards

The traditional third-party audit system is likely to differentially alter auditor incentives for false reporting across the pollution distribution: the conflict of interest, if realized, should cause auditors to increase reports of false compliance. To examine this, we start with a graphical illustration of the impact of the audit treatment on the full distribution of auditor readings. Figure 2 shows the distributions of Suspended Particulate Matter (SPM) concentrations, an important air pollutant. Panels A and B show distributions from audit reports and backchecks for the control and treatment, respectively. In each distribution a vertical line marks the pollutant standard, and we report the share of the probability mass that falls between 75% of the standard and the standard (this area is shaded in grey).

The figure shows a striking pattern: Reports from control auditors are clustered immediately below the standard, while reports from treatment auditors show much more dispersed distributions similar to those for backcheck readings. Specifically, Panel A shows a large spike of readings in the audit control immediately beneath the standard. A full 73% of control audit reports are in the area from 75% to 100% of the standard, as against 19% of the readings in the backcheck distribution. The backcheck readings are an important counterfactual, as such a spike may occur without misreporting if plants abate pollution just enough to be in compliance with the standards. In the audit treatment group, shown in Panel B, 39% of readings are in the area beneath the standard. This is far lower than in the control group audits but still higher than in the backcheck distributions for either treatment or control plants, at 19% and 14%, respectively. The reduction in clustering below the standard in the treatment audit reports implies that far more plants are reported as out of compliance, at levels up to twice the standard of 150 mg/Nm^3 or more. More plants are also reported with emission levels well below the standard. We attribute this increase in the left tail in the treatment relative to control to the cost of sampling air pollutants: it is cheaper to report narrow compliance by default than to properly sample and document a very low reading.

This pattern of control readings being just compliant holds across the full range of pollutants. Table 3 shows regressions for two outcomes: whether a reading is narrowly compliant (between 75% and 100%) (Panel A), and whether it is compliant (Panel B). Our estimation

equation is as given by Equation (1). Panel A, column (1) shows that audit reports are 27 percentage points more likely to show narrow compliance than backchecks in the control, on a base of 23 percentage points. In the treatment, audit reports are 19 percentage points less likely than in the control to falsely report a reading in this narrow range, a reduction of 69% relative to the control. Similarly, in panel B, column (1), treatment auditors are 81% ($-0.234/0.288$) less likely to falsely report for compliance as a whole. These effects are seen for both water pollutants, in column (2), and air pollutants, in column (3).

Figure 3 summarizes the changes in the density of the reported pollutant distribution due to the audit treatment. For this figure, all pollutant readings are standardized by subtracting the standard and dividing by the pollutant standard deviation in backchecks, so that the horizontal axis indicates the number of standard deviations above or below the limit. We regress indicator dummies for a pollutant reading belonging to a particular 0.05-standard-deviation-width bin on an indicator for audit treatment. Weakly negative values indicate compliance and positive values non-compliance. As expected the treatment dramatically reduces the amount of mass just beneath the standard. Density in the bins above the standard rises.²⁵

Though the compliance threshold is discrete the regulator and the public care about the degree of pollution emitted and therefore reported. Table ??, Panel A, shows ordinary least-squares regressions of standardized audit pollutant concentrations, pooled by pollutant type, on audit treatment status. We consider all pollutant readings reported across the two years of the experiment by auditors. As shown in column (1), the mean audit report reading for all pollutants in treatment plants is a significant 0.103 standard deviations higher than the mean report in the control. The coefficients on the audit treatment are similar for both water and air pollutants considered separately, as shown in columns (3) through (6). Adding auditor fixed effects, as shown in column (2), increases the point estimate of the audit treatment to 0.131 standard deviations, but the effects with and without auditor fixed effects remain statistically indistinguishable.

The magnitude of these effects is substantial: Consider the estimate in column (2) of 0.131 standard deviations for all pollutants, which is roughly of the same size as the effect for Biochemical Oxygen Demand (BOD) estimated alone (not shown). The standard deviation of

²⁵Using binned rather than quantile regressions allows us to normalize each pollutant relative to its own regulatory standard.

BOD in backchecks in final-outlet samples was 203 mg/l and the mean 191 mg/l, as against a concentration standard of 30 mg/l.²⁶ An effect of size 0.131 standard deviations would thus be 26.7 mg/l for BOD, or 89% of the standard. The mean and standard deviation of SO₂ readings in backchecks were 64 and 108 parts per million (ppm), respectively, as against a standard of 40 ppm. A 0.131 standard deviation movement is thus 35% of the standard. Consistent with our earlier discussion of the distributions these specifications show that the change in pollutant reports shifts several plants from compliance to non-compliance and is economically significant, in the sense of representing a meaningful increase in reported pollution.

A.2 Truth-telling by Auditors

If the audit treatment caused plants to abate pollution then differences in audit report levels understate the treatment effect on reporting, as the treatment effect would be a composition of lower pollution levels with higher, more truthful reports.

Matched backcheck-audit differences in Table ??, panel B, show that auditors report more truthfully in the treatment group of plants. This panel is restricted to a smaller sample and uses the difference with backchecks, rather than the audit report alone, as the outcome variable. The difference between backcheck and audit readings is -0.304 standard deviations in the control, indicating a negative bias in reporting, but 0.210 standard deviations (standard error 0.0726 standard deviations) higher in the treatment for all pollutants.²⁷ The positive treatment effects mean that auditors in the treatment are getting closer to the true pollutant readings. The estimates for differences are somewhat larger than the estimates in audit levels both because audit reports rose over the course of the experiment, and, to a lesser extent, because backcheck readings are also falling at the same time. We discuss plant response to the treatments in Section V C below.

Unlike in the specifications in audit levels, the treatment coefficient in audit differences becomes somewhat smaller in specifications with auditor fixed effects, though not significantly so.

Columns (4) and (6) show this change to be driven entirely by air pollutants, as the

²⁶This is for plants where the effluent did not go on for further treatment.

²⁷In estimates of the specifications of Table ?? including the number of days between the audit and backcheck as a control variable, not shown, this lag is uniformly not significant. The treatment effect estimates become somewhat smaller, with the treatment effect for all pollutants 0.181 standard deviations (standard error 0.0789 standard deviations) as opposed to the 0.210 reported when not controlling for the backcheck lag.

coefficients for water pollutants are about the same in specifications with and without fixed effects. The improvement in the accuracy of auditor reporting is essentially the same within-auditor as in the base specification without fixed effects.

B The Role of Incentives

B.1 Regulatory Incentives for Plants

Table ?? shows that misreporting is greater for water than for air pollutants in the control group and the point estimates of treatment effects for water pollutants are slightly smaller, so that a reduced bias in reporting persists in the treatment group for water, but not for air. To what extent does this reflect the regulatory penalty structure?

In the auditor survey described in Section II C, we asked auditors what pollution measurements were viewed as most important by plants and GPCB. Auditors responded that pH and COD, both measures of water quality, were the most important to both parties. This perception is justified. In GPCB's own files, plants are most frequently cited for violations of standards for COD and BOD, both water pollutants.

To present more systematic evidence of the correlation between GPCB stringency and misreporting, in Figure 4 we separately plot the share of falsely compliant audit readings for each pollutant against the share of actions taken by GPCB that cited each pollutant for treatment and control. We measure violations by entering all actions taken by GPCB against plants in the study sample from 2007 through mid-2011, where each action includes a citation of reason. The plot shows that false compliance is larger in the control for precisely those pollutants that are most likely to be cited against plants. An increase of one percentage point in citation share leads to a 0.8 percentage point greater false compliance for a given pollutant, with a very strong fit. In the treatment group we continue to observe a positive relationship, though the slope coefficient falls to 0.6. Auditors are responsive to incentives for accuracy overall and also responsive to incentives to mute potentially incriminating reports at the level of the individual pollutant.

This strong relation between regulatory scrutiny and false compliance in the control, and the change in this relationship associated with treatment, suggest that auditors are not merely shirking in giving low reports but targeting their reporting to balance the need to maintain some veneer or honesty against the needs of client plants. As noted in the discussion of

Figure 2 in Section V, some pollutant readings are reported to be just compliant when in fact they are very low and easily compliant, which suggests some shirking in auditors forgoing sampling when a plant is not very polluting. If auditors' motive was only to avoid cost and effort, though, we would observe the greatest false reporting for the air pollutants that are most costly to sample. In fact, we see higher false reporting for water pollutants that are cheap to sample and analyze but may pose a high cost to firms if reported accurately.

B.2 Financial Incentives for Auditors

Our treatment consisted of several components, each of which may have independently influenced auditor reporting. In this section, we exploit non-experimental variation in the intensity of financial incentives to examine whether this component had an independent effect on reporting.

The treatment in year two included incentive pay for auditors based on the accuracy of their reports. Comparing the size of the treatment effects in the two years of the study may help to decompose the treatment into the effect of auditor independence, in assignments and means of payment, and the effects of direct incentives for accuracy. Table 5 presents estimates of the treatment effects for all pollutants with controls for time to isolate the effect of incentive pay. The first column duplicates the base specification of Table ?? for reference showing the effect of the dummy for the year being equal to 2010, when incentive pay was in effect. The second and third columns add an interaction of that dummy with the treatment, a linear trend in fractional years since January 1, 2009 and a separate linear trend for the treatment group, where pollution may be changing differentially. Column 2 shows reported audit levels as an outcome and column 3 compliance. In the levels specification in column 2, audit reports are higher in the treatment than in the control, as in the main specification of Table ?. The coefficient on Years X treatment shows a significant downward trend in pollution over time of -0.22 standard deviations per year, whereas the main time trend of 0.029 standard deviations per year is not different from zero, indicating no pollution trend in the control.

Against this trend, there is a positive interaction between incentive pay, the dummy for the year being 2010, and treatment. Audit reports remain a significant 0.257 standard deviations differentially higher in treatment in year two, when incentive pay was in effect, compared to

year one.²⁸ In column (3), by contrast, reported compliance surprisingly increases in the treatment in year two by 7.28 percentage points (standard error 3.43 percentage points)—the opposite of what we would expect for more truthful reporting. We interpret this finding as suggesting that the time trend is a poor control for the changes over time in compliance, a binary outcome that is sensitive to shifts in the mass of reports around the pollutant standard. In particular, there may be few marginally underreporting firms left for higher audit reports to shift out of compliance in year two.²⁹

Auditor reports were therefore higher in year two than in year one after controlling for pollutant trends in the treatment and control separately. Note that, within years, pollutants in the treatment are trending downwards. While the interpretation of the dummy for 2010 rests on a correct specification of the time trend, the result suggest a direct positive impact of incentive pay.

C Response of Polluting Plants

The experiment increased truth-telling by auditors. Since the purpose of the audit scheme is ultimately lower pollution emissions by plant, the natural next question is whether that change in auditors' behavior led to response by plants.

The model suggests that the change in audit reporting may have changed the plant's decision margin to whether to abate or incur a fine, rather than whether to abate or purchase a compliant report. Whether or not plants abate in response to the treatment then depends on the relation of abatement costs to fines. If the stakes are low and abatement costs are high, so that few plants have abatement costs less than fines, then we would expect no pollution response. Conversely, plants with abatement costs less than the fine expected for a report on their true level of pollution will reduce their pollution emissions in response to the audit treatment.

Table 6 shows regressions of pollution concentrations and compliance measured during the endline survey on audit treatment status.³⁰ We find evidence that plants reduce pollution

²⁸Adding quadratic and cubic trends and their interaction with treatment (results available from authors), the discrete jump in reporting in 2010 increase to 0.284 standard deviations (standard error 0.116 standard deviations)

²⁹Consistent with this interpretation, the coefficient on interaction of incentive pay and treatment is not as robust to the inclusion of extended time controls (not shown), with the point estimate nearly unchanged at 7.61 percentage points but the standard error rising to 4.73 percentage points.

³⁰The audit intervention was conducted concurrently with another treatment at the plant level. We restrict the pollution sample in this table to the subset of audit sample plants not subject to the other experimental

significantly in response to the audit treatment. In panel A, the units are standardized in the same manner as in the auditor accuracy tables above. On average firms subject to the audit treatment reduced pollution by a significant 0.165 standard deviations (standard error 0.0778 standard deviations) This reduction comes mostly from a large decrease of 0.223 standard deviations in water pollutant concentrations, shown in column (2), as opposed to air pollutants, shown in column (3), for which the point estimate is also negative but insignificant. This coefficient for all pollutants is due in part to several control plants with very high pollution readings, and decreases in magnitude to -0.143 standard deviations (standard error 0.068 standard deviations) and -0.114 (standard error 0.058 standard deviations) if we top-code readings above the 99.5 and 99 percentiles, respectively, of the pollutant distribution.³¹ The volume of effluent emitted, not shown, did not change in response to the experimental treatments, so these reductions in concentrations represent reductions in the total emitted effluent load—less water pollution—amongst treatment plants. In panel B, we report changes in compliance with the pollution standards. Compliance is estimated to have increased by a small and insignificant 2.68 percentage points. This suggests that pollution reductions were not coming from plants near the compliance threshold, but to plants with very high pollution levels.

Figure 5 shows that pollution reductions occurred almost exclusively in the right tail of the pollutant distribution. This figure shows the coefficients from quantile regressions of standardized endline pollutant levels on audit treatment and region fixed effects, ranging from the 0.05-quantile to the 0.95-quantile at 0.05-quantile intervals. The dark grey shaded area is a 90% confidence interval and the lighter grey a 95% confidence interval. While no individual quantile coefficient is significant at conventional levels, the point estimates show a clear pattern wherein the treatment reduced pollution more at higher quantiles of the pollution distribution. Up to the 0.75-quantile, point estimates are very close to zero, but from the 0.80-quantile onwards point estimates sharply decrease to under -0.5 standard deviations at the 0.95-quantile.

This pattern of pollution reductions, like the pattern in auditor reporting, has a clear

intervention.

³¹There are however several pieces of corroborating evidence that these readings are genuine and should not be top-coded, which is why we present the non top coded specification as our main one. The plants in question have track records of non-compliance and other pollutant samples collected at the same plants, both before the high endline pollutant readings and contemporaneously, also show pollution levels well above the average.

economic rationale in the history of GPCB's actions at different levels of pollution. Figure 6 shows the actions resulting from pollutant violations at different levels above the compliance threshold, as measured during inspections, amongst audit sample plants over the three years beginning the year before the study and running through its end. Follow up actions by GPCP are classified in four groups increasing in severity from the bottom (dark grey bars) to the top (light grey bars): a letter is official but not legal correspondence to the firm noting the violation and possibly threatening some action, a citation is a legal notice requiring a response from the firm, a closure is a warning that the plant will be closed unless a violation is immediately remedied, and a disconnection is an order to the utility that a plant's power be disconnected. All of these actions were coded based on complete administrative records of plant interactions with the regulator. All violating plants receive some citation, in accord with the letter of the regulation. Plants polluting at above 100% but less than 150% of the standard receive the most severe actions, closure warning and disconnection, less than 10% of the time, which increases to over 70% of the time for plants polluting at more than 1000% of the standard. The most costly sanctions are in practice reserved for the right tail of the pollutant distribution. This relationship between high pollution and likely penalties explains why accurate reporting induces high-polluting firms to clean up. All else equal, these firms have a much higher expected fine, and therefore when the audit treatment shifts their decision margin to abate or defy the regulation altogether, they abate.

The reform of auditor reporting thus caused plants working with independent auditors to reduce pollution emissions. This decrease comes from the water pollutants that are priorities for GPCB and which were the original spur for the development of the audit scheme. Reductions are concentrated entirely in high-polluting plants that have a greater incentive to abate given the observed history of follow up actions. A natural interpretation is that more accurate pollution reading in the audit reports led polluting plants to believe that sanctions were now more likely, spurring cleaning up efforts. It is also possible that there is a direct effect of auditors' increased effort, in the form of more relevant advice on what plants could do to clean up.

VI Conclusion

Third-party environmental auditors report more accurately when working under a system which insures financial independence from plants, monitoring for accuracy, and enough funding to actually perform the work. Client plants, in turn, respond to improved auditing by reducing pollution output.

In control plants, misreporting of compliance just beneath the threshold for compliant levels is ubiquitous, and stronger for the pollutants that the regulator cares most about, as indicated by citations of pollutants in violations. Readings in the treatment group are much less likely to falsely show compliance and on average are significantly higher and more accurate for a standardized pollutant index of water and air pollutants. By the end of the second year of the experiment, reports in the treatment group were unbiased. Additional, non-experimental analysis suggest that the financial incentives independently contributed to the change in auditor reporting.

Plants randomly assigned to the audit treatment responded to tighter monitoring by reducing their pollution output as measured about six months after the end of the audit intervention. This reduction in pollution was on the same scale as the observed change in audit reporting, though many plants in the treatment remained out of compliance with tight regulatory standards. Here again, the effects are stronger for water than air pollutants.

The results observed in this experiment can be interpreted in the context of a simple model of plant and auditor behavior. The two main results of the experiment are that auditor reporting became almost completely accurate and that, while actual pollution levels by plants declined on average, there were still many plants out of compliance. This is consistent with a model where there is heterogeneity in abatement costs, or in expected penalties, amongst plants, many plants have abatement costs greater than the anticipated cost of fines for high reports, and auditors had a significant continuation value from participating in the audit market. The experimental treatments were strong, in the sense that they moved the value of buying falsely low reports sufficiently that the relevant decision margin for plants was between abatement and open non-compliance. The large change in auditor reporting towards accurate reports argues that for many plants the expected fine was less than abatement cost. The observed reduction in pollution levels shows that for a significant fraction of plants,

abatement was less costly than the expected costs associated with a high audit report.

We attribute these results to greater independence of auditors from client plants and incentives for accurate reporting. The audit reforms raised the cost for auditors of misreporting and, with assignment to plants, removed the potential benefit of ensuring continued business. Several alternative explanations for why treatment auditors reported more accurately can be ruled out. The treatment effects are not experimental Hawthorne effects of studying auditors, or of auditors anticipating that reputation might matter more under revised audit regulations in the future, as the treatment effects are observed even comparing the behavior of the same audit plants under different incentive systems, and response to the experiment *per se* would occur at the auditor level. The differences in behavior are also not attributable to the simple fact that the price of the audit was raised by the treatment. Receiving a larger payment, in the absence of any changes in the cost of auditing effort or the benefits of accurate reporting, rational auditors would pocket the difference rather than spend it to improve audit quality. Nor are the results merely indicative of reduced shirking, as the level of misreporting and treatment response for different pollutants vary with the regulatory scrutiny each pollutant receives.

These results are encouraging in the context of environmental regulation in India. One of the basic challenges of regulation is gathering accurate information on the performance of regulated plants in the face of agency problems either in third-party reporting or within the regulator itself.³² The studied reform shows it is feasible for regulators to collect accurate information on pollution and use it to enforce regulatory standards. Where this information begets enforcement, for especially high readings, plants appear to respect the threat of action by reducing pollution.

The audit scheme in Gujarat was introduced to emulate “external audit[s] in the fields of company law and income tax” (Gujarat High Court, 1996) and succeeded all too well. While the core issue of a conflict of interest is common to this study’s setting and those of corporate auditing or credit ratings, there are limitations to how well the results may transpose to those settings. One limitation is that in other regulatory and competitive environments auditors might have stronger returns to a reputation for being accurate. As pointed out by Ronen

³²Like the environmental audit system in Gujarat, the national system for Environmental Impact Assessments (EIA) has foundered on this problem, because, as put by a former Minister of Environment, “[T]he person who is putting up the project will be preparing the report” (The Hindu, 2011).

(2010) and Bolton et al. (2011), though, reputational incentives can cut both ways. Auditors may want to develop a reputation for leniency to attract more business, as long as there are enough naive investors or other clients to accept their recommendations at face value. Moreover, while audit companies are large and may want to internalize the reputational externalities of misreporting, it is less clear that individual managers or analysts have the same incentives (Bar-Isaac and Shapiro, 2011). The evidence on the quality of corporate financial auditing is not very convincing, as quality is measured *ex post* by instances where the SEC decided to investigate, but the financial crisis gave a window into the quality of the credit ratings. The massive downgrades that had to take place suggest that credit ratings were greatly inflated, in particular in the boom years (Ashcraft et al., 2010).

Despite the low quality of reporting at baseline, we note that the environmental audit market we study did have some value of a reputation for quality. If there had been no value of reputation then we would not have observed a treatment effect at all, nor would we observe that auditors strategically misreport rather than just reporting all readings as compliant. We can measure the accuracy of auditors using auditor fixed effects in a regression of backcheck less audit differences on treatment. When we then regress the prices that auditors charge in the control group on their measured accuracy, we find that a one standard deviation increase in accuracy allows auditors to charge INR 14,630 *more* than less accurate peers, a large premium. Even in a broken market there is a niche for reliable information intermediaries.

Another factor that may make it difficult to adopt a similar system outside of the environmental field is that annual rotation of auditors may be infeasible for the financial audits or credit ratings of large plants, where there may be large fixed costs in understanding a plant's finances. Modifying the form of rotation to guarantee auditors would work with a client for, say, five years before being reassigned, unless found to be grossly inadequate under regulatory review, would serve the same purpose, of severing the link between auditors' work and plant satisfaction, while reducing the importance of fixed costs. While the details of a modified third-party audit system would clearly need to vary across settings, our study shows that, in one context, conflicts of interest matter enormously, and reforms that make auditors genuinely independent can bring about truthful reporting and greater compliance.

VII References

References

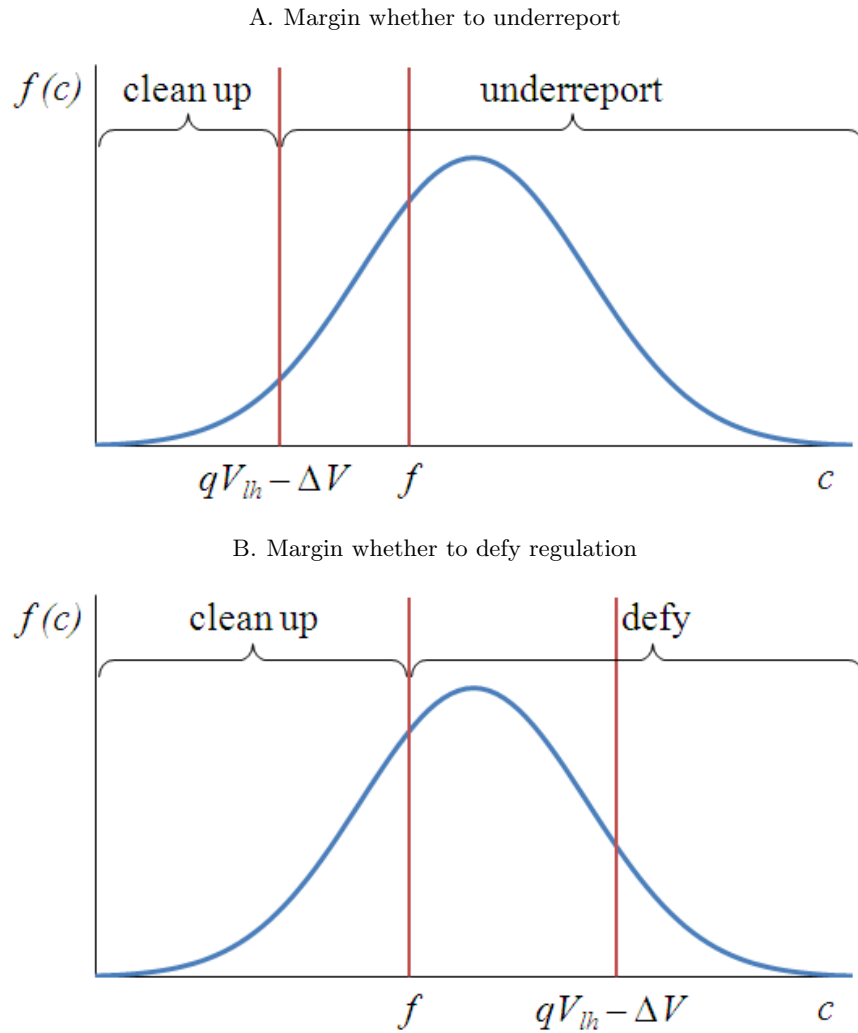
- Ashcraft, A., Goldsmith-Pinkham, P., and Vickery, J. (2010). Mbs ratings and the mortgage credit boom. *Federal Reserve Bank of New York Staff Reports*, (449).
- Bar-Isaac, H. and Shapiro, J. (2011). Credit ratings accuracy and analyst incentives. *American Economic Review, Papers & Proceedings*, 101(3):120–24.
- Bhattacharyya, S. (2011). The clean development mechanism. *Energy Economics*, pages 623–645.
- Bolton, P., Freixas, X., and Shapiro, J. (2011). The credit ratings game. *The Journal of Finance*, Forthcoming.
- Boyd, C. E. (2000). *Water Quality: An Introduction*. Klumer Academic Publishers.
- Central Pollution Control Board (2007). Annual report. Technical report, http://cpcb.nic.in/upload/AnnualReports/AnnualReport_34_Annual-Report-06-07.pdf.
- Central Pollution Control Board (2009a). Annual report. Technical report, http://cpcb.nic.in/upload/AnnualReports/AnnualReport_37_ANNUAL.REPORT-08-09.pdf.
- Central Pollution Control Board (2009b). Comprehensive environmental assessment of industrial clusters. Technical report, http://www.cpcb.nic.in/upload/NewItems/NewItem_152_Final-Book_2.pdf.
- Chakravorty, S. (2003). Industrial location in post-reform india: patterns of inter-regional divergence and intra-regional convergence. *The Journal of Development Studies*, 40(2):120–152.
- Chen, Y., Ebenstein, A., Greenstone, M., and Li, H. (2010). The impact of air pollution on life expectancy: Evidence from china’s huai river policy. *Mimeo*.
- Currie, J. and Walker, R. (2011). Traffic congestion and infant health: Evidence from e-zpass. *American Economic Journal: Applied Economics*, 3(1).

- Dranove, D. and Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. *Journal of Economic Literature*, 48(4).
- Duflo, E., Hanna, R., and Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American Economic Review*, 102:1241–78.
- Griffin, J. and Tang, D. (2011). Did credit rating agencies make unbiased assumptions on cdos? *American Economic Review, Papers & Proceedings*, 101(3):125–30.
- Gujarat High Court (1996). Environmental audit scheme for industries manufacturing specified products. *Special Civil Application*, 770/1995.
- Gujarat High Court (2010). Gujarat dyestuff manufacturers association vs. state of gujarat. *Special Civil Application*, 22609/2005.
- Hanna, R. and Greenstone, M. (2011). Environmental regulations, air and water pollution and infant mortality in india. *CEEPR Working Paper*, 2011(14).
- Hanna, R. and Oliva, P. (2011). The effect of pollution on labor supply: Evidence from a natural experiment in mexico city. *National Bureau of Economic Research Working Paper*, (17302).
- Hatanaka, M., Bain, C., and Busch, L. (2005). Third-party certification in the global agrifood system. *Food Policy*, 30(3):354–369.
- IHD-WHO Working Group (1978). *Water Quality Surveys*. World Health Organization.
- International Monetary Fund (2009). Global financial stability report.
- Kreps, D. and Wilson, R. (1982). Reputation and imperfect information. *Journal of economic theory*, 27(2):253–279.
- Olken, B. (2007). Monitoring corruption: Evidence from a field experiment in indonesia. *Journal of Political Economy*, 115(2).
- Olken, B. and Pande, R. (2012). Corruption in developing countries. *Annual Review of Economics*, 4.

- Potoski, M. and Prakash, A. (2005). Green clubs and voluntary governance: Iso 14001 and firms' regulatory compliance. *American Journal of Political Science*, 49(2):235–248.
- Raynolds, L., Murray, D., and Heller, A. (2007). Regulating sustainability in the coffee sector: A comparative analysis of third-party environmental and social certification initiatives. *Agriculture and Human Values*, 24(2):147–163.
- Ronen, J. (2010). Corporate audits and how to fix them. *The Journal of Economic Perspectives*, 24(2):189–210.
- Strobl, G. and Xia, H. (2011). The issuer-pays rating model and ratings inflation: Evidence from corporate credit ratings. *Mimeo, University of North Carolina*.
- The Hindu (2011). Present system of environmental impact assessment has flaws, says ramesh. March 19, 2011.
- U.S. Environmental Protection Agency (2010). National ambient air quality standards.
- Waite, T. D. (1984). *Principles of Water Quality*. Academic Press.
- White, L. (2010). Markets: The credit rating agencies. *The Journal of Economic Perspectives*, 24(2):211–226.
- World Health Organization (2006). Who air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide. *World Health Organization*.

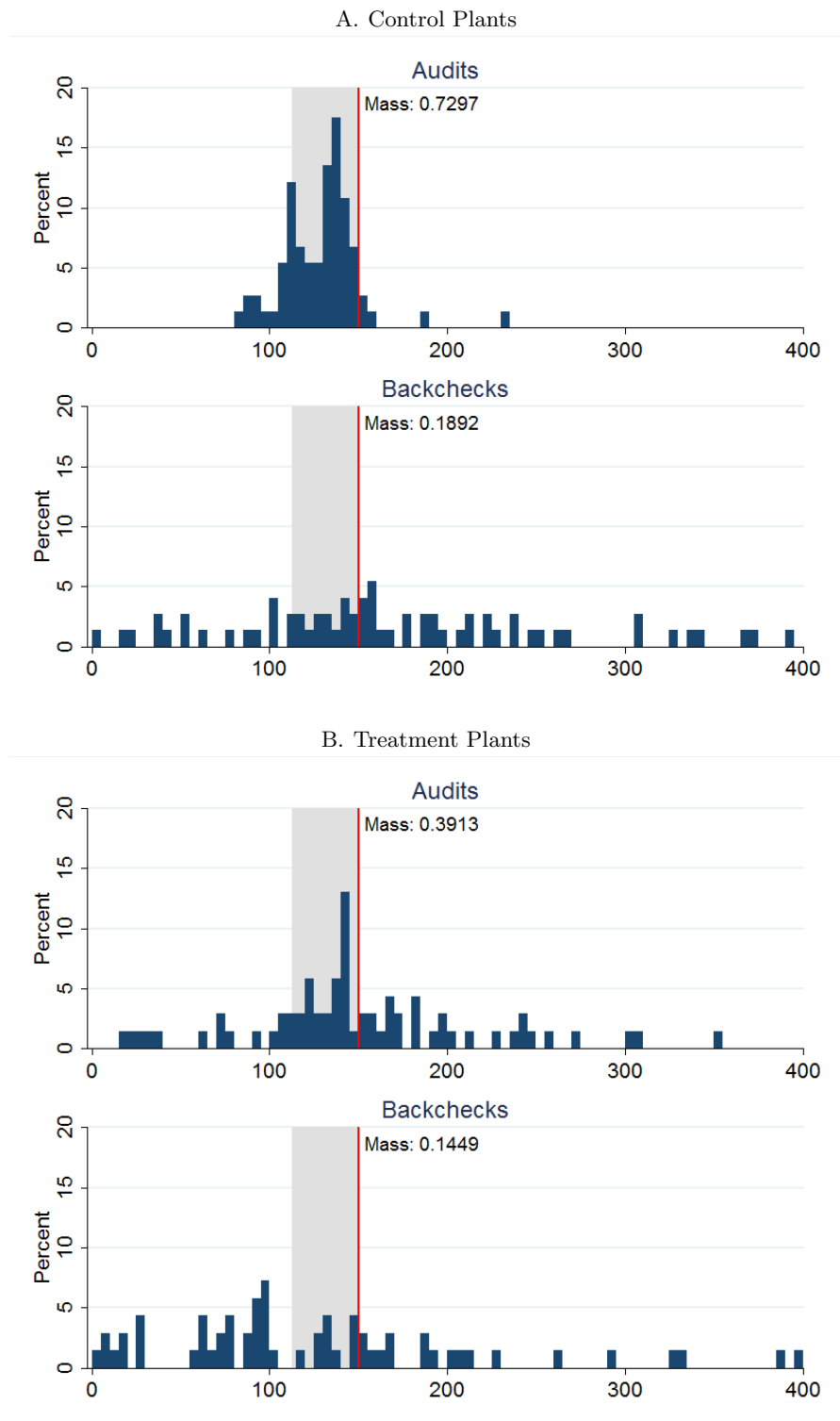
VIII Figures

Figure 1: Equilibrium Outcomes by Abatement Cost



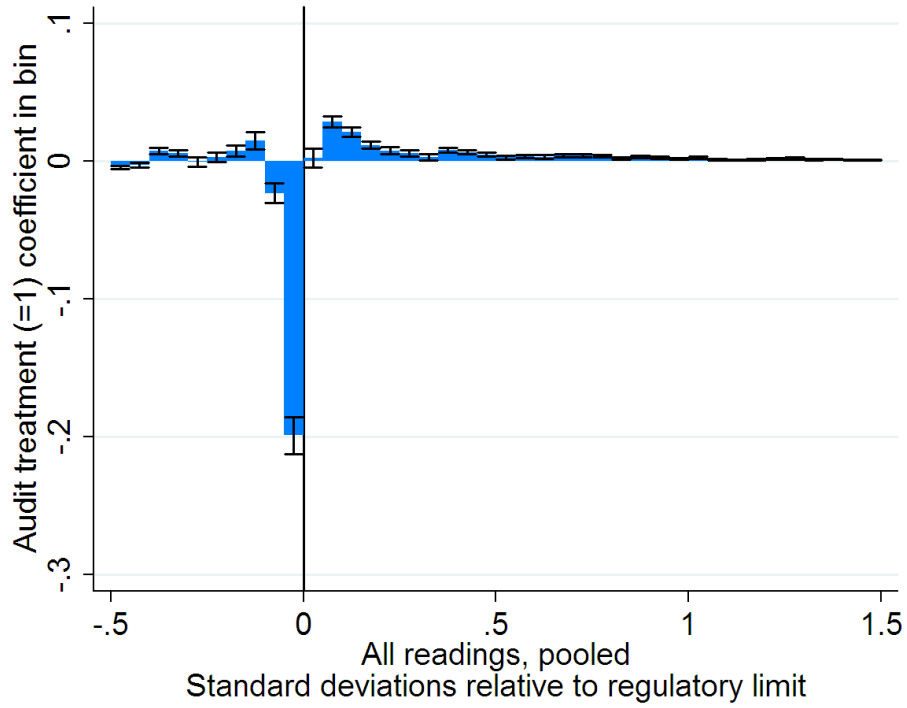
The figure shows possible outcomes in the model of Section III for a firm with $p = p_h$ above the regulatory standard. The distribution represents abatement costs c . The vertical lines on each graph represent the additional payment necessary to induce an auditor to falsely report low $qV_{lh} - \Delta V$ and the fine for a reported violation f .

Figure 2: Audit and Backcheck Readings for Suspended Particulate Matter (SPM, mg/Nm³), Midline



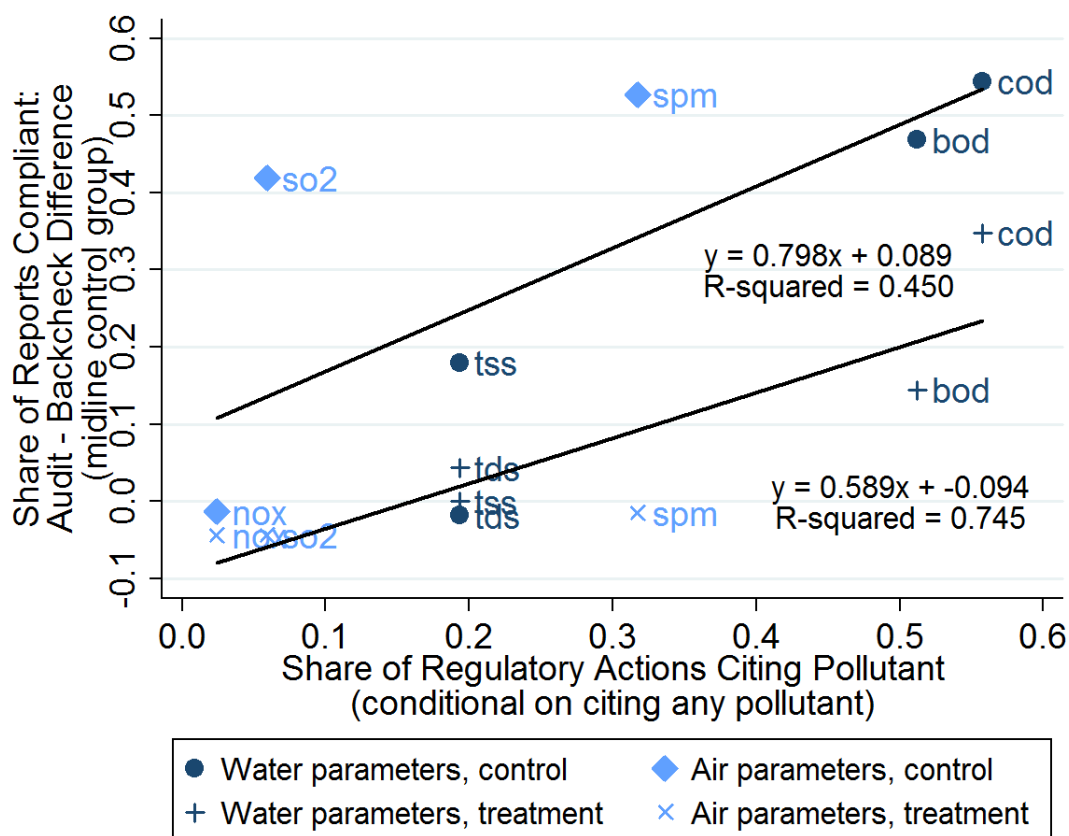
The figure shows distributions of pollutant concentrations for Suspended Particulate Matter in boiler-stack samples taken during the midline survey. Panel A shows the distributions of readings at control plants from audits and backchecks, respectively, and Panel B readings at treatment plants from the same two sources.

Figure 3: Audit Treatment Effect in Density Bins, All Pollutants



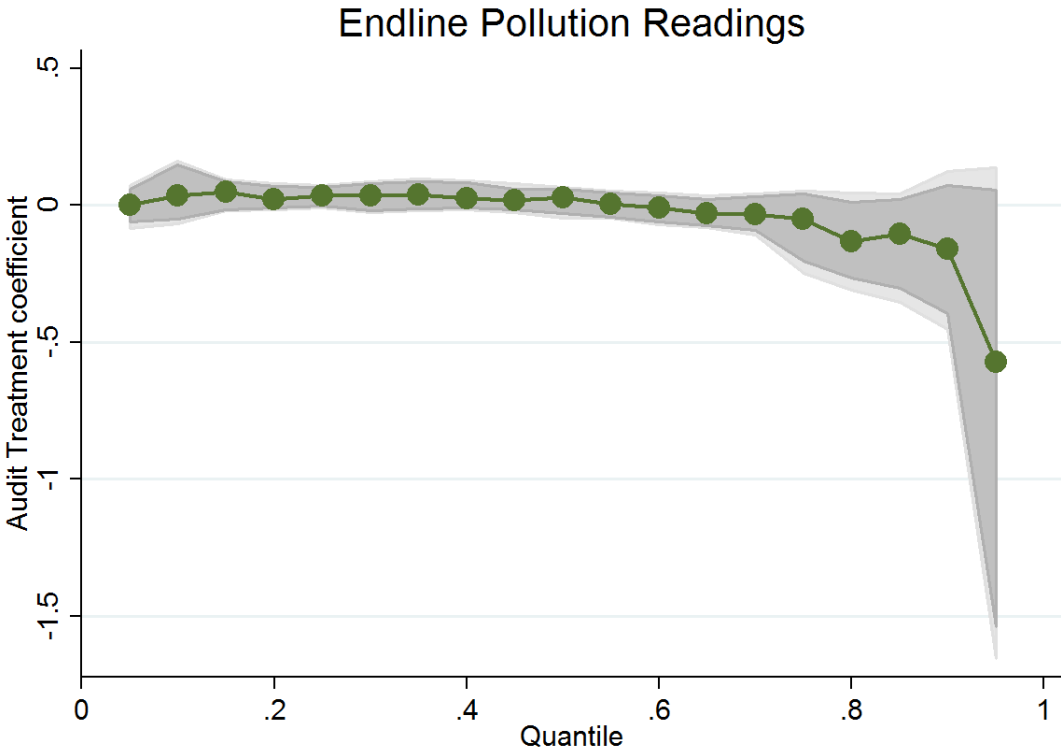
The figure reports point estimates and standard errors from OLS regressions where the dependent variables are indicators for a pollutant reading being within a given density bin and the independent variable is audit treatment. Pollutants included are $Water = \{NH_3-N, BOD, COD, TDS, TSS\}$ and $Air = \{SO_2, NO_x, SPM\}$ with $All = Water \cup Air$. Pollutants are standardized by subtracting the regulatory standard for each pollutant and dividing by the standard deviation in backchecks. Density bins are 0.05 standard deviations wide.

Figure 4: False Compliance against Pollutant Citations by Regulator, Control



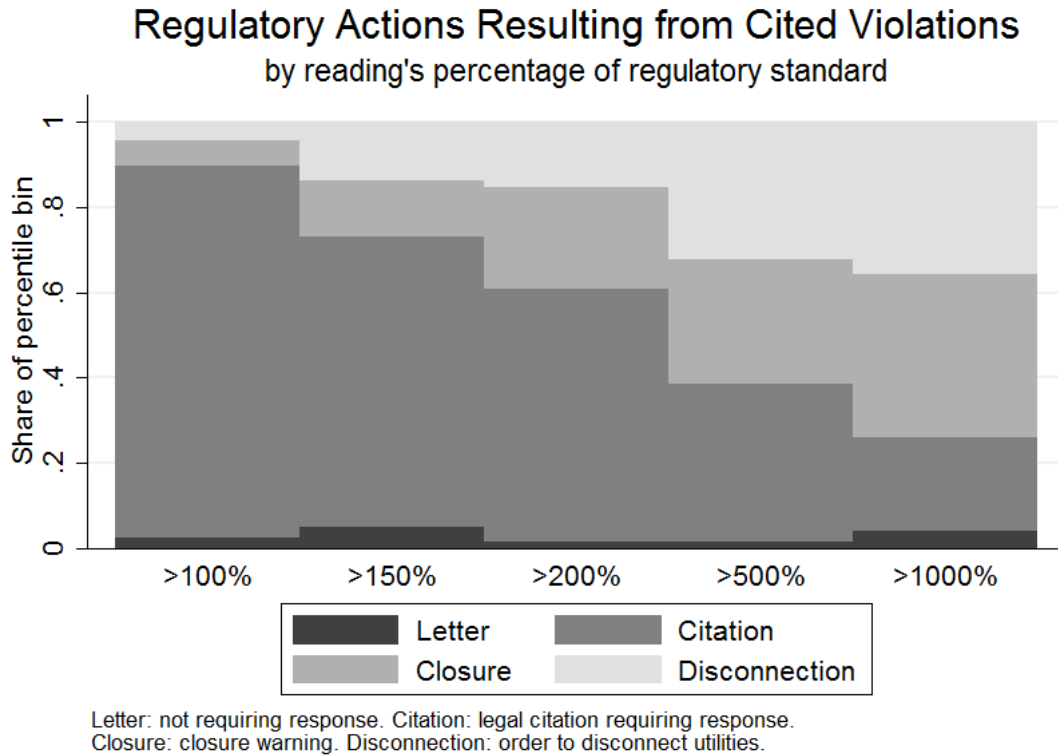
The figure reports a scatter plot for different pollutants of the difference between the share of firms reported compliant in audits versus backchecks, in the treatment and control groups, against the share of regulatory citations citing that pollutant, conditional on the citation citing any pollutant. The compliance difference is the same as the difference between the audit report coefficient in Table 3 and the corresponding backcheck mean level of compliance disaggregated by pollutant. Regulatory citations are from a complete review of all regulatory correspondence to sample plants from 2007 through 2011. The variable is constructed as the share of all regulatory actions (notices of violations or of pending enforcement) citing a given pollutant conditional on an action citing any pollutant.

Figure 5: Quantile Treatment Effects of Audit on Endline Pollution



The figure reports estimates from quantile regressions of standardized endline pollution for all pollutants on a dummy for audit treatment assignment in the inspection control sample, analogous to the specifications in Table ???. The quantiles are from 0.05-quantile to the 0.95-quantile at 0.05-quantile intervals. The confidence intervals shown are at the 95% and 90% level from a cluster-bootstrap at the plant level.

Figure 6: Regulatory Actions by Degree of Violation



The figure reports the regulatory responses to pollution readings measured at different levels of non-compliance during regulatory inspections for audit sample plants over the three years beginning one year prior to the study. Pollutant readings are shown in bins of readings at least a given percentage above the regulatory standard. The bars indicate the type of regulatory action taken in response to a given reading. Actions increase in severity from bottom (dark bars) to top (light bars): a letter is official but not legal correspondence to the firm noting the violation and possibly threatening action, a citation is a legal regulatory notice requiring a response from the firm, a closure is a warning that the plant will be closed unless a violation is remedied, and a disconnection is an order to the utility that a plant's power be disconnected. All of these actions were coded based on complete administrative records of plant interactions with the regulator.

IX Tables

Table 1: Audit Treatment Covariate Balance

	Treatment (1)	Control (2)	Difference (3)
<i>Panel A: Plant Characteristics</i>			
Capital investment Rs. 50m to Rs. 100m (=1)	0.092 [0.29]	0.14 [0.35]	-0.049 (0.033)
Located in industrial estate (=1)	0.57 [0.50]	0.53 [0.50]	0.037 (0.051)
Textiles (=1)	0.88 [0.33]	0.93 [0.26]	-0.046 (0.030)
Effluent to common treatment (=1)	0.41 [0.49]	0.35 [0.48]	0.067 (0.050)
Waste water generated (kl / day)	420.5 [315.9]	394.6 [323.4]	25.9 (33.0)
Lignite used as fuel (=1)	0.71 [0.45]	0.77 [0.42]	-0.058 (0.045)
Diesel used as fuel (=1)	0.29 [0.45]	0.25 [0.43]	0.037 (0.046)
Air emissions from flue gas (=1)	0.85 [0.35]	0.87 [0.33]	-0.021 (0.035)
Air emissions from boiler (=1)	0.93 [0.26]	0.92 [0.27]	0.0079 (0.027)
Bag filter installed (=1)	0.24 [0.43]	0.34 [0.47]	-0.096** (0.047)
Cyclone installed (=1)	0.087 [0.28]	0.079 [0.27]	0.0084 (0.029)
Scrubber installed (=1)	0.41 [0.49]	0.41 [0.49]	-0.0060 (0.051)
<i>Panel B: Regulatory Interactions in Year Prior to Study</i>			
Whether audit submitted (=1)	0.82 [0.38]	0.81 [0.39]	0.0091 (0.040)
Any equipment mandated (=1)	0.42 [0.50]	0.49 [0.50]	-0.068 (0.051)
Any inspection conducted (=1)	0.79 [0.41]	0.78 [0.42]	0.0079 (0.043)
Any citation issued (=1)	0.28 [0.45]	0.24 [0.43]	0.036 (0.045)
Any water citation issued (=1)	0.12 [0.33]	0.12 [0.33]	-0.00085 (0.034)
Any air citation issued (=1)	0.027 [0.16]	0.0052 [0.072]	0.022* (0.013)
Any utility disconnection (=1)	0.098 [0.30]	0.094 [0.29]	0.0036 (0.031)
Any bank guarantee posted (=1)	0.033 [0.18]	0.026 [0.16]	0.0064 (0.017)
Observations	184	191	

All variables from administrative data prior to the experiment. Firms in audit sample that submitted an audit report in either year 1 or year 2. Columns (1) and (2) show means with standard deviations in brackets. Column (3) shows differences between treatment and control with standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Submission of Audit Reports

	Treatment	Control	Difference
<i>Panel A: 2009</i>			
Audit submitted	163	177	
Total plants	233	240	
Share submitted	0.70	0.74	-0.038 (0.041)
<i>Panel B: 2010</i>			
Audit submitted	164	153	
Total plants	233	240	
Share submitted	0.70	0.64	0.066 (0.043)

Audit report submission to the regulator for audit sample over the two years of the experiment. Column (3) shows differences between treatment and control submission rates with standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Pollution Reports in Audits Relative to Backchecks
by Treatment Status

	(1) All	(2) Water	(3) Air
<i>Panel A: Report Between 75% and 100% of Regulatory Standard</i>			
Audit report X Audit treatment	-0.185*** (0.0339)	-0.212*** (0.0439)	-0.143*** (0.0462)
Audit report (=1)	0.270*** (0.0252)	0.297*** (0.0336)	0.230*** (0.0326)
Audit treatment (=1)	-0.0102 (0.0183)	-0.0206 (0.0247)	0.00555 (0.0239)
Control mean	0.232	0.259	0.191
<i>Panel B: Report Below Regulatory Standard</i>			
Audit report X Treatment	-0.234*** (0.0388)	-0.166*** (0.0501)	-0.345*** (0.0555)
Audit report (=1)	0.288*** (0.0233)	0.273*** (0.0327)	0.311*** (0.0322)
Audit treatment (=1)	0.0578* (0.0339)	0.00748 (0.0477)	0.145*** (0.0407)
Control mean	0.701	0.674	0.741
Observations	2236	1378	858

Regressions include region fixed effects. Standard errors clustered at the plant level in parentheses. Pollution samples from final-stage effluent outlet for water and boiler-stack for air. Pollutants included are $Water = \{NH_3-N, BOD, COD, TDS, TSS\}$ and $Air = \{SO_2, NO_x, SPM\}$ with $All = Water \cup Air$. Sample of matched pollutant pairs from audit reports submitted to the regulator and corresponding backchecks from the midline survey. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Audit Treatment Effects on Auditor Reporting

	All (1)	All (2)	Water (3)	Water (4)	Air (5)	Air (6)
<i>Panel A: Audit Report Levels</i>						
Audit treatment (=1)	0.103*** (0.0354)	0.131*** (0.0384)	0.117** (0.0527)	0.131** (0.0593)	0.0852*** (0.0214)	0.132*** (0.0216)
Auditor fixed effects	No	Yes	No	Yes	No	Yes
Control mean	-0.291	-0.291	-0.350	-0.350	-0.194	-0.194
Observations	13172	13172	8373	8373	4799	4799
<i>Panel B: Differences, Midline Sample</i>						
Audit treatment (=1)	0.210*** (0.0726)	0.153 (0.0986)	0.152 (0.102)	0.156 (0.138)	0.312*** (0.0825)	0.166* (0.0951)
Auditor fixed effects	No	Yes	No	Yes	No	Yes
Control mean	-0.304	-0.304	-0.354	-0.354	-0.225	-0.225
Observations	1118	1118	689	689	429	429

Regressions include region fixed effects and, in Panel A, year fixed effects. Standard errors clustered at the plant level in parentheses. Pollution samples from final-stage effluent outlet for water and boiler-stack for air. Pollutants included are $Water = \{NH_3-N, BOD, COD, TDS, TSS\}$ and $Air = \{SO_2, NO_x, SPM\}$ with $All = Water \cup Air$. Panel A sample of audit reports that reached the regulator over the two years of the experiment with audit report levels as the outcome variable. Panel B midline survey sample with audit reports less backchecks as the outcome. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Audit Reports on Treatment Status Over Time

	All (1)	All (2)	All (3)
Audit treatment (=1)	0.103*** (0.0354)	0.218*** (0.0627)	-0.150*** (0.0278)
Incentive pay (year=2010)	0.0561** (0.0273)	0.00187 (0.0325)	-0.104*** (0.0210)
Incentive pay X treatment		0.257*** (0.0920)	0.0728** (0.0343)
Years (fractional) from Jan 1, 2009		0.0290 (0.0226)	0.0706*** (0.0185)
Years (fractional) X treatment		-0.220*** (0.0677)	-0.0382 (0.0299)
Region fixed effects	Yes	Yes	Yes
Observations	13172	13172	13172

Regressions include region fixed effects. Standard errors clustered at the plant level in parentheses. Pollution samples from final-stage effluent outlet for water and boiler-stack for air. Pollutants included are $Water = \{NH_3-N, BOD, COD, TDS, TSS\}$ and $Air = \{SO_2, NO_x, SPM\}$ with $All = Water \cup Air$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Endline Pollutant Concentrations on Treatment Status, Inspection Control

	All (1)	Water (2)	Air (3)
<i>Panel A: Pollution Concentrations</i>			
Audit treatment assigned (=1)	-0.165** (0.0778)	-0.223* (0.121)	-0.0528 (0.0566)
Control mean	0.030	0.036	0.022
Observations	1439	860	579
<i>Panel B: Regulatory Compliance</i>			
Audit treatment assigned (=1)	0.0268 (0.0274)	0.0387 (0.0393)	0.00175 (0.0282)
Control mean	0.573	0.516	0.656
Observations	1439	860	579

Regressions include region fixed effects. Standard errors clustered at the plant level in parentheses. Pollution samples from final-stage effluent outlet for water and boiler-stack for air. Pollutants included are $Water = \{NH_3-N, BOD, COD, TDS, TSS\}$ and $Air = \{SO_2, NO_x, SPM\}$ with $All = Water \cup Air$. Endline survey data from audit sample in control group of treatment that cross-cut audit treatment. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

X Data Appendix

Table 7: Description of Pollutants

Pollutant	Description
Panel A: Water Pollutants	
Biochemical Oxygen Demand (BOD)	A measure of the amount of dissolved oxygen consumed by microscopic organisms in a confined sample of water. The BOD and volume of an effluent determine the oxygen demand that will be imposed on receiving waters (Boyd, 2000). The demand for oxygen from effluent may deplete available molecular oxygen, precluding other biological processes, such as marine plants or life, that require oxygen (Waite, 1984).
Chemical Oxygen Demand (COD)	A measure of the oxygen demand of the organic matter in a sample as determined by oxidation of the organic matter with potassium dichromate and sulfuric acid. Often used as a proxy for BOD in determining the oxygen demand of effluent.
Total Dissolved Solids (TDS)	Primarily inorganic substances dissolved in water, including calcium, magnesium, sodium, potassium, iron, zinc, copper, manganese, etc. Water with high dissolved solids is said to be mineralized and decreases the survival of plant and animal life, degrades the taste of water, corrodes plumbing and limits use of water for irrigation (Boyd, 2000; IHD-WHO Working Group, 1978). Depending on the composition of solids TDS may have adverse health effects on people with cardiac disease or high blood pressure.
Total Suspended Solids (TSS)	Organic and inorganic or mineral particles too large to be dissolved but small enough to remain suspended against gravity in an effluent (Boyd, 2000). Contribute to turbidity and color of water and proxy for adverse effects from individual solid components.
Ammonia Nitrogen ($\text{NH}_3\text{-N}$)	The nitrogen contained in unionized ammonia and ammonium. Though nitrogen is a vital nutrient, some forms of ammonia nitrogen are toxic to aquatic life (Boyd, 2000). The toxicity of ammonia nitrogen increases with decreasing dissolved oxygen concentrations.
Panel B: Air Pollutants	
Sulfur Dioxide (SO_2)	A reactive oxide of sulfur. Short-term exposure has been linked to adverse respiratory effects particularly damaging for asthmatics. SO_2 also contributes to formation of fine particles (World Health Organization, 2006).
Nitrogen Oxides (NO_x)	A group of reactive gases including nitrous acid, nitric acid and NO_2 . Nitrogen oxides are toxic at high concentrations and contribute to formation of ozone and fine particles, which are detrimental to health (World Health Organization, 2006).
Suspended Particulate Matter (SPM)	A mixture of small particules and liquid droplets with a number of components, including acids, organic chemicals, metals and soil or dust (U.S. Environmental Protection Agency, 2010). Particulate matter affects respiratory and cardiovascular health and has been shown to increase infant mortality and shorten lifespans (World Health Organization, 2006; Currie and Walker, 2011; Chen et al., 2010).

XI Model Appendix

The following is a subgame-perfect Nash equilibrium with regulation:

1. Firms abate if and only if $c \leq \min\{f, qV_{hl} - \Delta V\}$.
2. Firms offer payments

$$w(\hat{p}|p) = \begin{cases} a - \Delta V + qV_{lh} + \varepsilon & qV_{hl} - \Delta V \leq f, \hat{p} = p_l, p = p_h \\ a & otherwise \end{cases} .$$

3. Auditor submits report

$$\hat{p} = \begin{cases} p_h & p = p_l, \quad w_{ll} < w_{hl} - qV_0 & \cup & w_{hl} \geq a \\ p_l & p = p_l, \quad w_{ll} \geq w_{hl} - qV_0 & \cup & w_{ll} \geq a \\ p_h & p = p_h, \quad w_{lh} < w_{hh} - \Delta V + qV_{lh} & \cup & w_{hh} \geq a \\ p_l & p = p_h, \quad w_{lh} \geq w_{hh} - \Delta V + qV_{lh} & \cup & w_{lh} \geq a - \Delta V + qV_{lh} \end{cases} .$$

If none of these conditions are met, the auditor declines to submit a report.

4. Regulator imposes a fine f if $\hat{p} = p_h$ is observed and disbars auditors observed to have reported falsely.

The definition of an equilibrium with regulation is one in which the regulator imposes a fine f if $\hat{p} = p_h$ is observed and disbars auditors observed to have reported falsely. This equilibrium is unique in this set up to the off-equilibrium path payments of the plant to the auditor and therefore all other equilibria with regulation are payoff equivalent.

Consider the auditor's actions. In the subgame with $p = p_l$, the auditor receives a higher payoff from (truly) reporting $\hat{p} = p_l$ if $w_{ll} - a + V_0 \geq w_{hl} - a + (1 - q)V_0$ and is willing to do so if $w_{ll} - a + V_0 \geq V_0$. These conditions are equivalent to $w_{ll} \geq w_{hl} - qV_0, w_{hl} \geq a$ given that $\Delta V \equiv V_{lh} - V_0$. In the subgame with $p = p_h$. The auditor receives a higher payoff from (falsely) reporting $\hat{p} = p_l$ if $w_{lh} - a + (1 - q)V_{lh} \geq w_{hh} - a + V_0$ and is willing to do so if $w_{lh} - a + (1 - q)V_{lh} > V_0$. These conditions are equivalent to $w_{lh} \geq w_{hh} - \Delta V + qV_{lh}$ and $w_{lh} \geq a - \Delta V + qV_{lh}$.

Consider the plant's actions in the subgame where it has chosen $p = p_l$. Inducing low reporting requires $w_{ll} - a + V_0 \leq w_{hl} - a + (1 - q)V_0$ or $w_{ll} \leq w_{hl} - qV_0$. Inducing a high

report is optimal if $w_{ll} \geq w_{hl} + f$ but doing so requires a $w_{ll} \leq w_{hl} - qV_0$, a contradiction. The plant thus never induces a high report. For a low report it offers payment at cost. The payment for a high report can be any such that $w_{ll} = a > w_{hl} - qV_0 \rightarrow w_{hl} < a + qV_0$; we let $w_{hl} = a$ above.

Consider the plant's actions in the subgame where it has chosen $p = p_h$. Given the strategies of the auditor and regulators, inducing low reporting requires $w_{lh} - a + (1 - q)V_{lh} \geq w_{hh} - a + V_0$, $w_{lh} - a + (1 - q)V_{lh} \geq w_{hh} - a + V_0$. Inducing false reporting is profitable if $w_{lh} \leq w_{hh} + f$ and is possible if $w_{lh} \geq w_{hh} - \Delta V + qV_{lh}$. The plant will not pay any more than necessary, so inducing false reporting is worthwhile if $w_{hh} - \Delta V + qV_{lh} \leq f + w_{hh} \Leftrightarrow qV_{lh} - \Delta V \leq f$ which is the condition given. The plant will set $w_{lh} \geq a - \Delta V + qV_{lh}$, which is the lowest payment at which the auditor will submit. If the plant is not inducing false reporting it sets $w_{hh} = a$ to induce submission and $w_{lh} \leq a$.

Now consider the plant's choice of whether to abate. The outcome of the subgame without abatement is that total plant costs are $\min\{f, qV_{lh} - \Delta V\}$. In the subgame with abatement, abatement c is the only cost. Abateing is therefore rational if $c \geq \min\{f, qV_{lh} - \Delta V\}$ as given.