Technology as Tool of Government: Evidence from Satellite-based Environmental Enforcement in India

Aditya Dasgupta*

September 15, 2025

Abstract

Developing countries struggle with enforcement, ranging from tax collection to environmental governance. By deploying new monitoring technologies, can governments improve enforcement? This paper provides cautionary evidence from a natural experiment in the usage of satellites for environmental enforcement in India. When the party governing the city of Delhi won power in Punjab, an agricultural state responsible for much of the farm fire smoke that causes air pollution in the capital, the new government deployed satellite fire monitoring to put pressure on local bureaucracy to enforce laws against crop burning. Using a border spatial regression discontinuity design, I show that detected fires fell sharply in Punjab. However, 80 percent of this decline was driven by farmers shifting burn times to evade satellite detection, likely in collusion with local officials. I estimate "hidden fires" using a machine-learning algorithm that imputes the true number of fires from burn scars, which cannot be concealed. A promising technological solution unraveled due to collusive gaming by local bureaucrats and polluters, highlighting Goodhart's law as a cautionary principle for deploying technology as a tool of government.

*Associate Professor of Political Science and Director, Political Economy of Agriculture and Rural Societies Lab. University of California, Merced. Mail: 5200 North Lake Road, Merced, CA 95343. Email: adasgupta3@ucmerced.edu

1 Introduction

Developing countries struggle with enforcement, ranging from tax collection to environmental governance, in part due to due to monitoring problems. However, recent technological advances enable analysts to monitor ground-level outcomes in new ways, such as using cameras or cell phones to track field-level outcomes (Duflo et al 2012), using forensic algorithms to infer procurement corruption (Bandiera et al 2009), or using satellite imagery to monitor environmental harms (Burgess et al 2012). A natural question is whether these technological advances can be deployed by governments to improve monitoring of violations and therefore also enforcement? This paper provides cautionary evidence from a natural experiment in the usage of satellite monitoring for environmental enforcement in India.

This paper studies these issues in the context of crop-residue ("stubble") burning in northern India, a major cause of air pollution in cities. In 2022, the Aam Aadmi Party (AAP) won the state election in Punjab while also governing Delhi, a megacity ranked among the most polluted in the world. Transboundary political alignment created new incentives for the Punjab government to enforce long ignored laws gainst crop burning, which contributes significantly to air pollution in Delhi in the winter season. Punjab authorities began to rely on thermal detection of fires from the NASA MODIS and VIIRS satellite sensors to target enforcement. Measured by these sensors, detected fires fell to their lowest levels in years. Nonetheless, air quality in Delhi remained dire, spurring speculation that farmers had learned to burn outside the satellites' detection windows.

I leverage a spatial border discontinuity design to estimate the causal effect of this enforcement push and its unintended consequences. Comparing change in detected fires in grid cells just inside Punjab to adjacent cells across the border, I document a sharp decline in detected fires in Punjab relative to neighboring areas across the border. I then show that this apparent success was largely illusory. Using burned-area imagery, which records lasting burn scars from fires that cannot be concealed, I train a machine learning algorithm (a convolutional neural network) on pre-2022 data to predict the number of fires implied by burned area. The gap between predicted and detected fires—"hidden" fires—grows substantially in Punjab after cross-border political alignment, consistent with strategic time-shifting of burns to evade thermal detection during satellite overpass windows. Finally, I present a simple model in which local officials, facing political pressure to show progress and social pressure to avoid conflict with farmers, prefer to share evasion-relevant information rather than enforce, yielding a collusive equilibrium in which the metric improves while the underlying outcome changes little.

The paper makes four key contributions to the literature on environmental governance, technology adoption, and political economy. First, it provides causal evidence that cross-jurisdictional political incentives can reshape environmental enforcement (see e.g. Sigman, 2002; Lipscomb and Mobarak, 2007; DiPoppa and Gulzar 2024). Second, it combines a natural experiment with remote-sensing forensics – using a novel machine-learning algorithm to estimate hidden fires – to provide evidence that most of the observed reduction in burning was due to gaming as opposed to an actual reduction in pollution, contributing to a literature on corruption forensics (see e.g. Jacob & Levitt, 2003; Olken, 2007). Third, it articulates and tests a mechanism—bureaucrat-polluter collusion—that helps explain how farmers learned to game the satellites and why a technologically sophisticated monitoring intervention backfired. Fourth, it contributes to the growing literature on the unintended consequences of performance measurement in public administration by showing how Goodhart's law operates in the domain of environmental governance.

The remainder proceeds as follows. Section 2 describes the policy and geographic setting and outlines the theoretical framework. Section 3 details data and measurement, including the CNN approach to impute fires from burn scars. Section 4 sets out the empirical strategy. Section 5 presents main results and robustness checks, followed by Section 6, which develops a simple model of collusive gaming and derives comparative statics. Section 7 concludes.

2 Background and Setting

North india faces one of the world's most severe air pollution crises, with cities like Delhi regularly recording particulate matter concentrations (PM2.5 etc) that exceed World Health Organization guidelines by an order of magnitude. A major contributor to this crisis is seasonal agricultural burning in Punjab and neighboring agricultural states, where farmers burn rice crop residue to quickly clear fields for wheat planting. The vast majority of fires are concentrated in the agricultural state of Punjab alone, with most of the burning place following the Kharif/winter harvest season from mid-October to mid-November. The scale is enormous—tens of thousands of fire detections occur during this period, creating smoke plumes visible from space that contribute to a regional haze affecting air quality across the Indo-Gangetic Plain.

Despite decades of policy interventions, including subsidies for mechanical alternatives and penalties for violators, the problem proved remarkably persistent due to weak enforcement. State governments lacked strong incentives to crack down on burning because farmers represent a powerful voter bloc in Punjab politics. Additionally, the costs of enforcement were borne locally by Punjab while the air quality benefits primarily accrued to Delhi, representing a classic problem of failure to enforce laws against environmental harms as a result of transboundary externalities (Greenstone and Hanna, 2014; DiPoppa and Gulzar, 2024).

This dynamic changed dramatically following the 2022 Punjab state election, when the Aam Aadmi Party (AAP)—which had governed Delhi since 2015—won a decisive victory in Punjab. For the first time, a single party controlled both the source region (Punjab) and the most affected sink region (Delhi). Air pollution consistently ranks as a an important concern for Delhi voters, and the annual winter smog season generates intense media coverage and public criticism, making it both a policy priority and electoral liability for AAP.

The AAP government in Punjab moved quickly to demonstrate action, announcing a "zero tolerance" policy toward burning and committing to use satellite technology for real-time violation detection, with data collected from MODIS and VIIRS satellite by the Punjab Remote Sensing Centre (PRSC) in Ludhiana and the Punjab Pollution Control Board (PPCB), which then shares specific location data with district heads for action against violators. The state established action plans with specific targets, including aims for ”at least 50% reduction in fire counts in Punjab during 2023 compared to last year.”Local district officials were given explicit targets for reducing detected fires.

This enforcement system represented a textbook application of performance measurement in public administration—using objective, technology-based metrics to create accountability and drive behavioral change.The results appeared remarkable. By winter 2024, satellite-detected fires in Punjab fell to their lowest recorded levels. Government officials celebrated these improvements as evidence of successful policy implementation, and the decline was widely reported in media as a major environmental victory. Figure 1 displays temporal trends in detected fires for Punjab using both MODIS and VIIRS sensors. The data reveal clear seasonal patterns, with fire activity peaking during the post-harvest period from October to December. Following political alignment in 2022, detected fires fall sharply in Punjab. Notably, in 2024 the number of detected fires in the winter season reached an all-time low since satellite data has been available, and appeared to be lower than the number of fires detected in the spring – a reversal of the normal pattern.

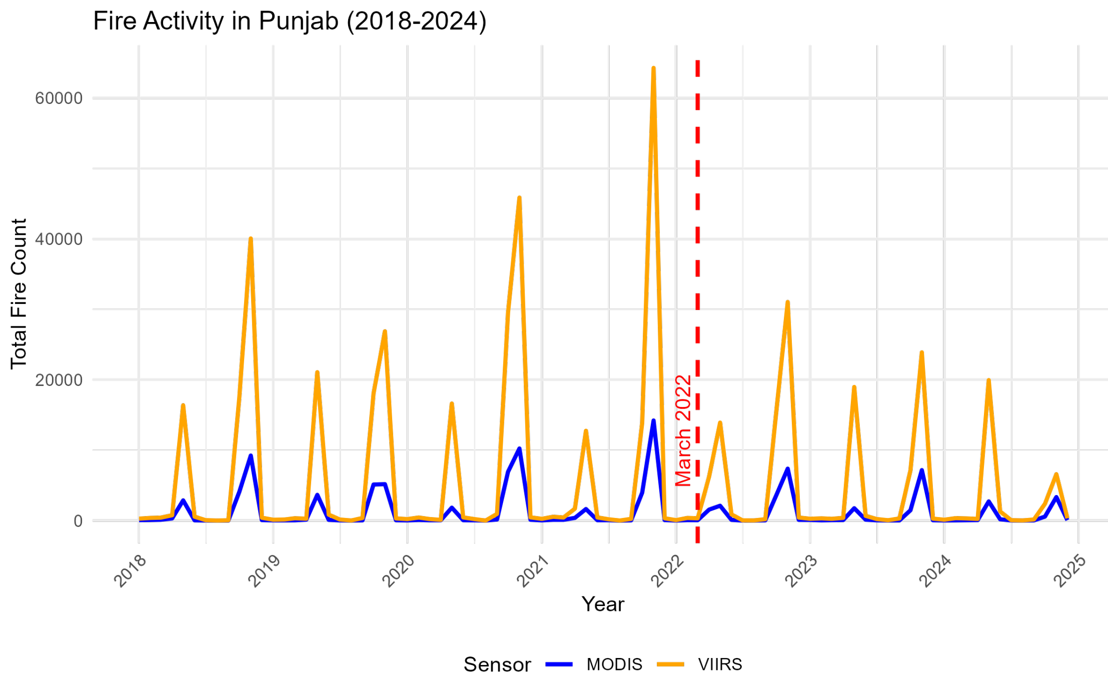


Figure 1: Fires in Punjab over time (MODIS and VIIRS). The figure shows monthly fire detections in Punjab from 2010 to 2023. Detected fires from both sensors show sharp declines following political alignment in 2022. The vertical dashed line indicates the timing of political alignment.

However, a troubling mystery emerged: despite the dramatic reduction in detected fires, air pollution levels in Delhi remained stubbornly high during the 2024 burning season. Air quality indices showed little improvement over previous years, raising questions about the relationship between satellite measurements and actual environmental outcomes. Rumors began circulating that farmers were adapting to the monitoring system by burning in the early evening to avoid satellite detection. Since MODIS and VIIRS overpasses over northern India occur at approximately 10:30 AM and 1:30 PM daily, and a typical field burn lasts only 2-3 hours, farmers could complete burning in the later afternoon and early evening to avoid thermal detection entirely.

This apparent disconnect between measured performance and environmental outcomes reflects a classic example of Goodhart's law—the principle that "when a measure becomes a target, it ceases to be a good measure." Originally formulated in monetary economics, Goodhart's law warns that any statistical regularity will collapse once pressure is placed upon it for control purposes. In environmental governance contexts, this suggests that using easily measured proxies as enforcement targets may create incentives for actors to optimize the measure rather than the underlying environmental outcome.

The Punjab case satisfied all key conditions for Goodhart's law to apply: thermal fire detections provided an imperfect proxy for true burning activity, political pressure created strong incentives to improve measured performance, the predictable nature of satellite overpasses enabled gaming through timing manipulation, and authorities had limited capacity to detect such evasion without alternative measurement approaches. Critically, local bureaucrats—facing intense pressure to show reductions in detected fires while maintaining farmer support—had both motive and opportunity to facilitate such gaming. Rather than engaging in costly and politically difficult enforcement against burning, officials could plausibly ease pressure by informally advising farmers about satellite overpass times and encouraging nighttime burning. This would allow officials to meet their performance targets while avoiding direct confrontation with their agricultural constituencies, representing a rational response to the incentive structure created by the satellite-based monitoring system.

3 Data and Measurement

3.1 Fire Detection Data

The analysis focuses on a border region spanning Punjab and its three neighboring states: Haryana to the south and east, Rajasthan to the south, and Himachal Pradesh to the north and northeast. This region encompasses the core agricultural areas where rice-wheat systems predominate and farm burning is most intensive.

Following standard practice in spatial analysis, I aggregate all observations to uniform square grid cells measuring $10\,\text{km} \times 10\,\text{km}$. This grid resolution balances several considerations: it is fine enough to capture local variation in burning patterns while being coarse enough to reduce noise from measurement error in satellite products. The resulting dataset contains 1,333 grid cells within 100 kilometers of the Punjab border.

For each grid cell, I construct time series of fire activity at the grid-cell-month level spanning the period 2010-2024, though in the empirical analysis the key outcome measure is on change in fires from the pre-alignment period (2020-2021) to the post-alignment period (2023-2024), where the measure for each grid cell is the total number of computer fires over these time periods.

The primary outcome measures use thermal fire detections from MODIS and VIIRS satellites. I obtain these data from NASA's Fire Information for Resource Management System (FIRMS), which provides near real-time fire detection products with quality-assured processing.

For each sensor, I aggregate individual fire detections to grid cell-season totals, distinguishing between the spring fire season (February-May) and the post-harvest winter fire season (October-January). The winter season receives primary attention as it corresponds to the rice harvest period when burning contributes most significantly to Delhi's air pollution.

3.2 Estimating "Hidden Fires"

A key innovation of this study involves using burned area measurements to infer "hidden" fires that evade thermal detection. Burned area products capture the cumulative footprint of fire activity over longer time periods and are less susceptible to timing-based evasion strategies.

I use the MODIS burned area product (MCD64A1), which provides monthly maps of burned areas at 500-meter spatial resolution. Unlike thermal detections that capture instantaneous fire activity, burned area measurements record the persistent signatures of fire that remain visible for weeks or months after burning occurs.

To convert burned area measurements into estimates of fire activity, I develop a machine learning approach using convolutional neural networks (CNNs). The intuition is that different spatial patterns of burned area correspond to different types and intensities of fire activity. By training a CNN on pre-treatment data where both thermal detections and burned area are available, I can learn a mapping function that predicts expected fire activity based on burned area patterns.

The CNN architecture uses burned area tiles as input and produces predicted fire density maps as output. The model is trained on data from 2010-2018, before political alignment created incentives for gaming. For each grid cell and time period in the post-treatment era, I apply the trained model to observed burned area patterns to generate "imputed" fire counts representing the expected level of thermal detections given the observed burning footprint.

The difference between imputed fires (based on burned area) and detected fires (from thermal sensors) provides an estimate of "hidden" fires that escaped thermal detection. This approach assumes that the relationship between burned area and fire activity learned from pre-treatment data provides a valid counterfactual for what thermal detections should be in the post-treatment period absent gaming.

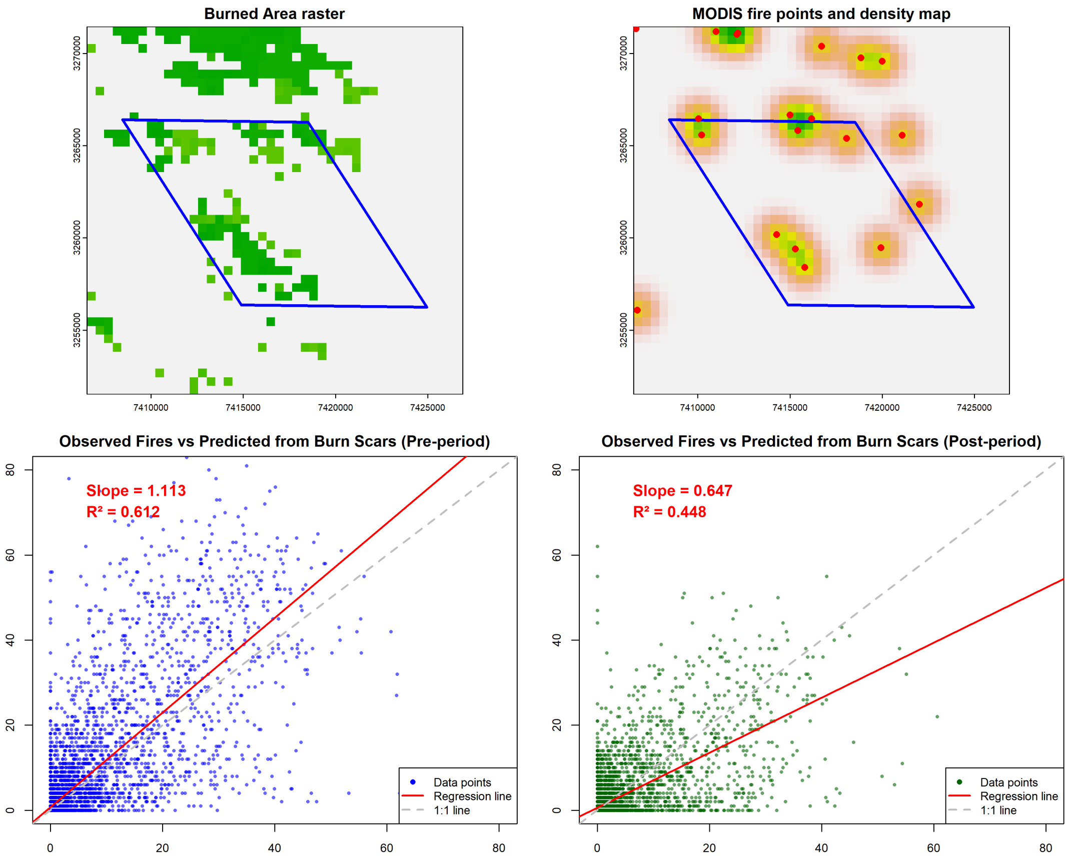


Figure 2: Top-left panel shows input data, burned area raster, overlaid over an example grid cell polygon. Top-right panel shows the density maps that are constructed from fire point coordinates. I train a convolutional neural network to predict the fires density map from burned area raster data. Bottom-left panel shows predictive performance of the panel in Punjab in the pre-period (2020-21). Bottom-right panel shows the predictive performence in the post-period, showing substantial decay in the performance of the model due to gaming/burn time-shifting.

4 Empirical Strategy

4.1 Spatial Regression Discontinuity Design

The empirical strategy exploits the sharp political boundary between Punjab and neighboring states to identify causal effects of enforcement pressure on fire activity.

The spatial RD design compares grid cells just inside Punjab to adjacent cells just outside

Punjab, leveraging the assumption that geographic and agricultural characteristics evolve smoothly across state boundaries while enforcement policies can change discontinuously at the border.

The running variable is distance to the Punjab border, defined as positive for cells inside Punjab and negative for cells outside Punjab. I implement the RD using local linear regression with a 100km bandwidth around the border.

The outcome measure is the change in fire counts from the pre-treatment period (2010-2021) to the post-treatment period (2023-2024). This temporal differencing focuses identification on changes that coincide with political alignment while controlling for time-invariant differences across the border.

Let ΔY_i denote the change in fire activity for grid cell i between the pre- and post-treatment periods. The specification is:

$$\Delta Y_i = \alpha + \tau \cdot \text{Punjab}_i + f(\text{dist}_i) + \delta_s + \varepsilon_i \tag{1}$$

where Punjab_i is an indicator for cells in Punjab, $f(\text{dist}_i)$ is a linear function of distance to the border that is allowed to vary on each side of each border segment, δ_s are border segment fixed effects (capturing distinct borders with Rajasthan, Himachal Pradesh, and Haryana). This equation is estimated restricting the sample to grid cells within a 100km bandwidth of the Punjab border. Additionally, I prune grid cells on or near the boundary itself, which may incorrectly record fires on the other side of the border due to satellite measurement error.

The coefficient τ identifies the causal effect of political alignment on the change in fire activity in Punjab relative to neighboring areas. The border segment fixed effects control for systematic differences between Punjab's borders with different states, while the segment-specific slopes allow the relationship between distance and fire changes to vary across different border regions.

The identifying assumption is that, absent political alignment, changes in fire activity would have evolved smoothly across the state boundary within each border segment. I test this assumption using placebo tests that examine whether similar discontinuities exist when the treatment year is artificially shifted to earlier periods.

4.2 Outcomes and Hypotheses

I estimate effects for three main outcome measures that allow me to distinguish between genuine reductions in burning and gaming through evasion:

Detected fires: Changes in thermal fire detections from MODIS and VIIRS satellites. If enforcement pressure reduces actual burning, detected fires should fall in Punjab relative to neighboring areas.

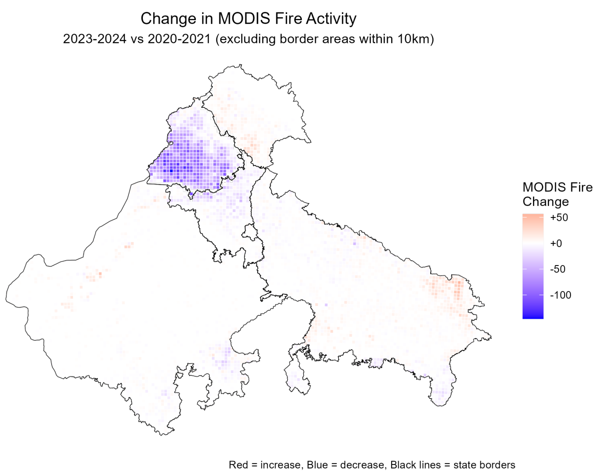
Hidden fires: Changes in the gap between imputed fires (from burned area CNN) and detected fires. If farmers shift burning to evade thermal detection without reducing actual burning, hidden fires should increase in Punjab.

Burned area: Changes in cumulative burned area from satellite imagery. If gaming primarily involves timing shifts rather than reductions in actual burning, burned area should show little change despite reductions in detected fires.

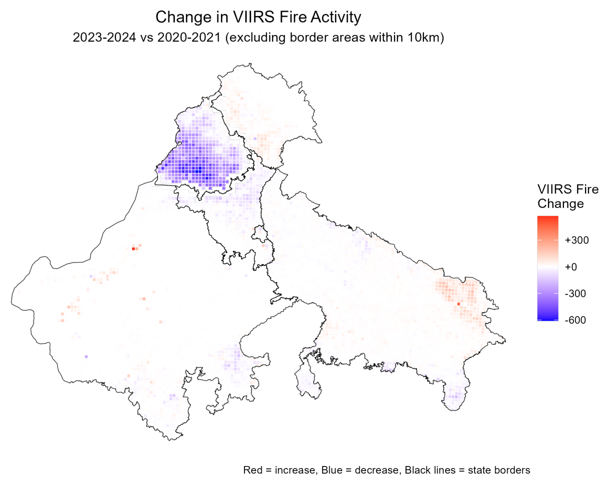
The combination of these measures allows me to test competing hypotheses about the mechanisms driving observed changes in detected fires. Genuine compliance should generate reductions in detected fires, no increase in hidden fires, and reductions in burned area. Gaming through timing shifts should generate reductions in detected fires, increases in hidden fires, and little change in burned area.

5 Results

Spatial patterns provide additional evidence of border-aligned changes in fire activity. Figure 3 displays maps of changes in detected fires between pre- and post-alignment periods for both MODIS and VIIRS sensors. The maps reveal pronounced reductions concentrated in Punjab, with much smaller changes in neighboring states. The sharp spatial contrast at state boundaries suggests that policy changes rather than region-wide environmental factors drive the observed patterns.



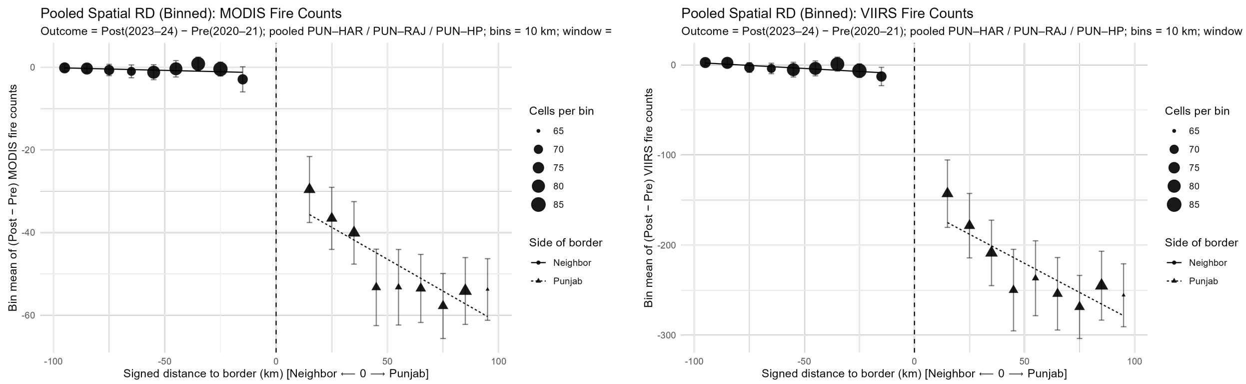
(a) MODIS fire change



(b) VIIRS fire change

Figure 3: Maps of change in detected fires by sensor (post-pre). The maps show the change in fire detections between the pre-alignment period (2020-2021) and post-alignment period (2023-2024). Red areas indicate reductions in detected fires, while blue areas indicate increases. The pronounced reductions in Punjab contrast sharply with neighboring states.

Figure 4 provides visual evidence of the discontinuities through binned scatter plots. The plots show clear jumps in fire reductions at the Punjab border for both sensors, with smooth trends on either side of the boundary. This pattern supports the identifying assumption that outcomes would evolve smoothly across the border absent policy changes.



(a) MODIS fires **(b)** VIIRS fires

Figure 4: Spatial RD plots for change in detected fires. The plots show binned scatter plots of changes in detected fires against distance to the Punjab border. Negative distances indicate locations outside Punjab, while positive distances indicate locations inside Punjab. The clear discontinuities at the border provide visual evidence of treatment effects.

5.1 Main Regression Discontinuity Results

Table 1 presents spatial RD estimates of the effect of political alignment on detected fires. The results confirm large and statistically significant reductions in both MODIS and VIIRS detections in Punjab relative to neighboring areas.

For MODIS fires, the pooled RD estimates indicate reductions ranging from 30 to 50 detected fires per $100\,\text{km}^2$ across different specifications. For VIIRS fires, the estimated effects are larger in magnitude, ranging from 180 to 260 fires per $100\,\text{km}^2$.

Table 1: Spatial RD estimates of effect of cross-border political alignment on change in detected fires

	Dependent variable:					
	Change in Detected Fires per 100 sq km grid-cell					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: MODIS fires						
Punjab (treatment)	−30.487***	−47.404***	−49.664***	−3.084	−5.143***	−25.344***
	(8.795)	(11.880)	(14.042)	(5.544)	(1.341)	(8.792)
Panel B: VIIRS fires						
Punjab (treatment)	−147.999***	−262.730***	−183.072***	−3.596	−9.915*	−138.084***
	(43.784)	(64.997)	(51.071)	(23.487)	(5.550)	(43.787)
Specification	Full	Haryana border	Rajasthan border	Himachal border	Spring fires	Winter fires
Observations	1,333	532	264	537	1,333	1,333

Notes: Spatial RD estimates on changes in detected fires per 100 km^2. Columns (1)-(4) show pooled effects and border-specific effects. Columns (5)-(6) show season-specific effects. Standard errors adjusted for spatial correlation within a 100 km radius using Conley adjustment. *** p<0.01, ** p<0.05, * p<0.1.

The results are robust across border segments and show particularly pronounced effects for winter fires, which are most relevant for Delhi's air quality. Effects for spring fires are much smaller and often statistically insignificant, consistent with the expectation that enforcement pressure focused on the post-harvest burning season.

5.2 Evidence of Gaming: Hidden Fires and Burned Area

The dramatic reductions in detected fires might suggest successful environmental enforcement. However, analysis of hidden fires and burned area reveals a more complex picture. Table 2 presents RD estimates for hidden fires, constructed as the difference between machine learning-imputed fires and detected fires.

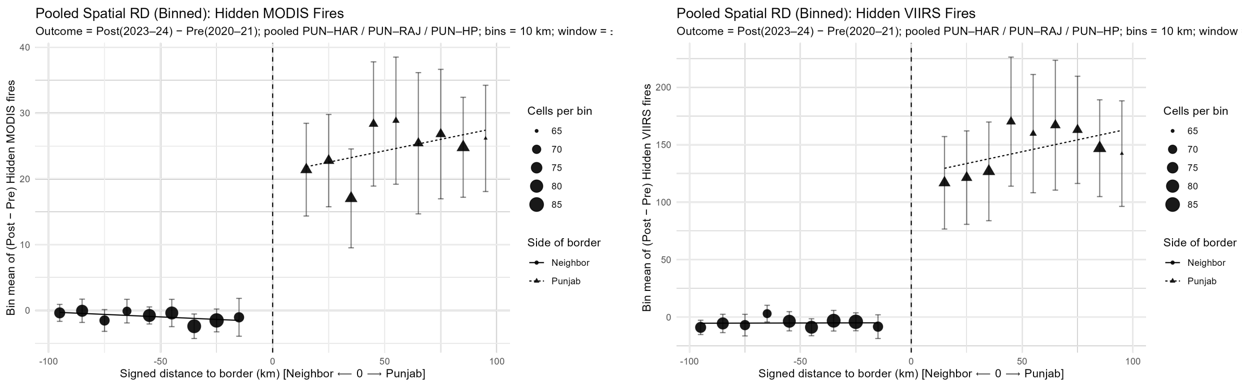
The results show large and statistically significant increases in hidden fires in Punjab following political alignment. For MODIS-based measures, hidden fires increase by 20 to 50 fires per 100 km^2, while VIIRS-based measures show increases of 125 to 225 fires per 100 km^2. These increases are roughly 70-80 percent as large as the reductions in detected fires, suggesting that most of the apparent improvement in enforcement reflects gaming rather than genuine compliance.

Table 2: Spatial RD estimates of effect of cross-border political alignment on change in hidden fires

	Dependent variable:					
	Average Change in Hidden Fires per 100 sq km grid-cell					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: MODIS fires						
Punjab (treatment)	22.777***	32.171***	47.470***	1.430	4.131***	18.646**
	(7.717)	(11.387)	(13.988)	(3.748)	(0.819)	(7.542)
Panel B: VIIRS fires						
Punjab (treatment)	127.955**	226.002***	207.647***	−17.147	2.016	125.939**
	(50.446)	(82.968)	(61.805)	(22.632)	(5.623)	(49.363)
Specification	Full	Haryana border	Rajasthan border	Himachal border	Spring fires	Winter fires
Observations	1,333	532	264	537	1,333	1,333

Notes: Spatial RD estimates for hidden-fire outcomes constructed as imputed fires (from burned-area CNN) minus detected fires. Standard errors use Conley adjustment with 100 km radius. *** p<0.01, ** p<0.05, * p<0.1.

Figure 5 provides visual evidence of these increases through binned scatter plots. The plots show clear discontinuous increases in hidden fires at the Punjab border, mirroring the pattern observed for detected fire reductions but with opposite sign.



(a) MODIS hidden fires **(b)** VIIRS hidden fires

Figure 5: Spatial RD plots for change in hidden fires (imputed minus detected). The plots show discontinuous increases in hidden fires at the Punjab border, providing evidence that farmers shifted burning to evade thermal detection rather than reducing actual burning.

Analysis of burned area provides further evidence against genuine compliance. RD estimates for changes in cumulative burned area (see Online Appendix) are small in magnitude and generally statistically insignificant. This finding is inconsistent with genuine reductions in burning activity, which should generate corresponding reductions in burned area. Instead, the pattern suggests that farmers maintained similar levels of actual burning while shifting the timing to evade satellite detection.

5.3 Robustness Tests and Placebo Analysis

The identification strategy relies on the assumption that changes in fire activity would evolve smoothly across the Punjab border absent political alignment. I test this assumption using placebo analyses that examine whether similar discontinuities exist when the treatment year is artificially shifted to earlier periods.

Figure 6 displays results from placebo tests for detected fires. Each point represents an RD estimate from a regression that treats a particular year as the treatment year, with 95 percent confidence intervals shown. The results demonstrate that estimated effects are close to zero and statistically insignificant for all placebo years, providing strong support for the identifying assumption.

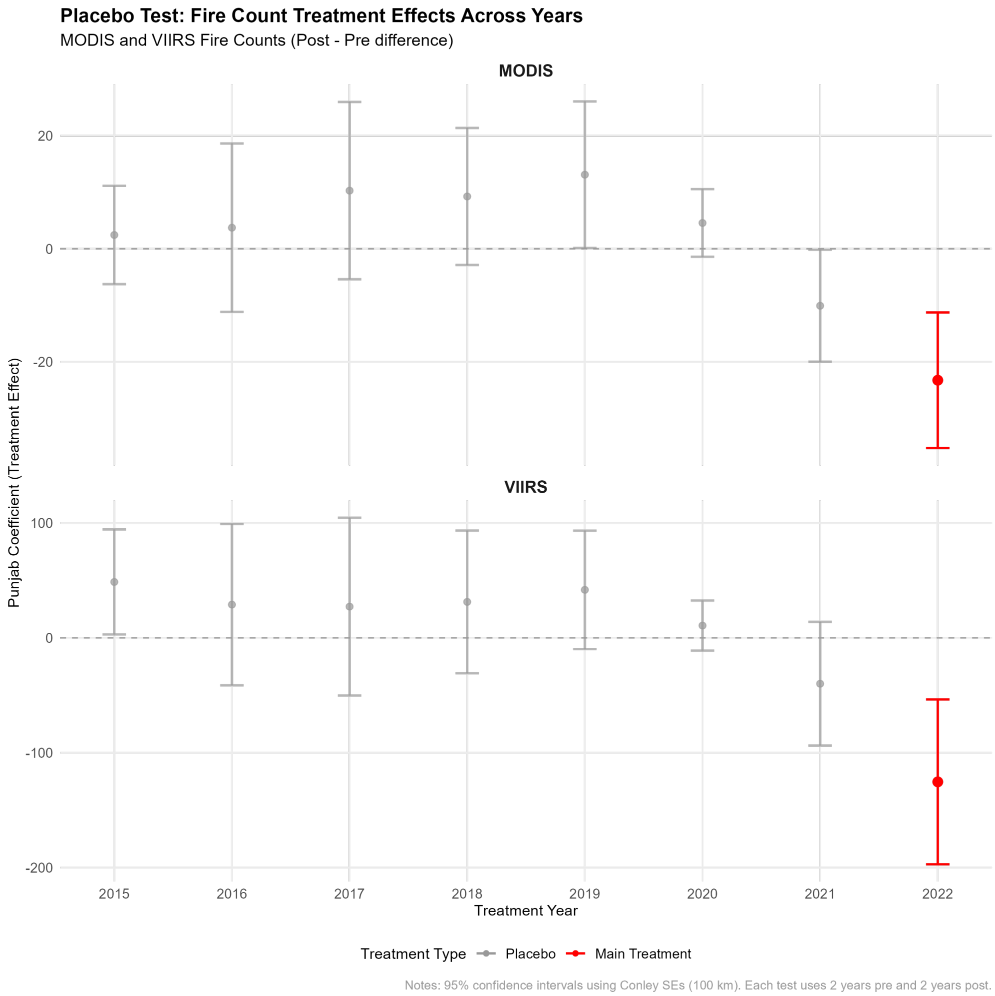


Figure 6: Placebo tests shifting treatment year backward: detected fires. Points show RD estimates treating each year as the treatment year; lines denote 95% confidence intervals. The large negative effects in 2022 (actual treatment year) contrast with near-zero effects in placebo years.

Figure 7 presents corresponding results for hidden fires, showing that increases in hidden fires are similarly concentrated in the actual treatment year rather than placebo years.

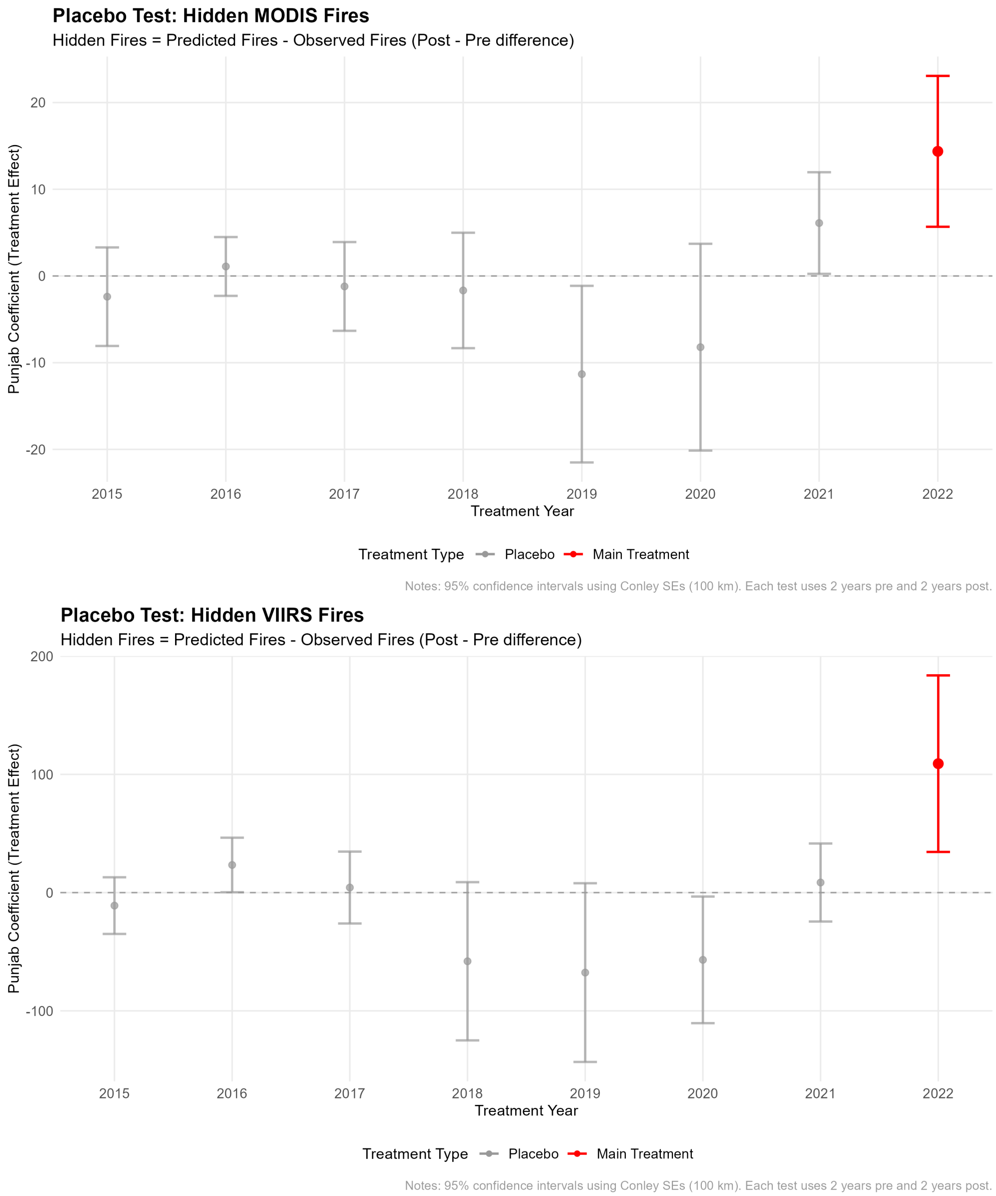


Figure 7: Placebo tests shifting treatment year backward: hidden fires. The positive effects on hidden fires are concentrated in 2022, consistent with gaming behavior emerging following political alignment.

6 Mechanism: A Simple Model of Collusive Gaming

The gaming mechanism likely operates through strategic timing of burning activities to avoid satellite detection windows. MODIS and VIIRS satellites have predictable overpass times that are publicly available, creating opportunities for farmers to time burning activities outside detection windows.

Yet an important puzzle is how farmers learned to game these overpass windows, which are unlikely to have been common knowledge among farmers and were not directly broadcast by the Punjab government. Below, I develop a simple model which suggests that this learning is consistent with information sharing between local officials and farmers. Local officials, facing pressure to show reductions in detected fires while maintaining social relationships with farmers, have incentives to share information about satellite overpass schedules that enables evasion without requiring costly enforcement actions.

6.1 Environment and Sequence

There is a politician P, jurisdictions $i = 1, \ldots, N$ each with a local bureaucrat B_i, and a continuum of farmers of measure M per jurisdiction. The timing:

S1 The politician chooses technology deployment $T \in \{0, 1\}$ and monitoring intensity $\theta \in [0, 1]$.

S2 Each bureaucrat B_i observes (T, θ) and chooses an enforcement effort $e_i \in [0, 1]$ and whether to share timing/evasion information $s_i \in \{0, 1\}$.

S3 Farmers in each jurisdiction observe (T, θ, e_i, s_i) and decide whether to burn crop residue or adopt clean alternatives.

6.2 Detection, Punishment, and Farmer Behavior

We separate (i) *detection* from (ii) *punishment bite* conditional on detection.

Detection. Let the jurisdiction-level detection probability be

$$q_i(T, \theta, s_i) \;=\; \eta + T\left[\theta(1 - \alpha s_i) - \eta\right], \quad \alpha \in (0, 1),\ \eta \in (0, 1).^{1} \tag{2}$$

[1] The parameter restrictions ensure $q_i \in [0, 1]$ for all feasible values of T, θ, and s_i.

When $T = 0$, detection equals the baseline η. When $T = 1$, higher θ raises detection unless bureaucrats share information ($s_i = 1$), which attenuates detection by a factor $(1 - \alpha)$.

Punishment bite. Conditional on detection, sanctions scale with bureaucratic enforcement effort via $r(e_i)$:

$$\text{sanction} \;=\; F \cdot r(e_i), \qquad r : [0,1] \to [0,1] \text{ increasing, with } r(0) = 0,\; r'(e) > 0. \tag{3}$$

A convenient benchmark is $r(e_i) = e_i$.

Farmers. Burning yields private benefit $B > 0$; the clean alternative normalizes to payoff 0.[2] A farmer burns iff expected sanction is below B:

$$B \;>\; q_i(T, \theta, s_i)\, F\, r(e_i). \tag{4}$$

Let B be heterogeneous with CDF G (continuous, strictly increasing) and density $g(\cdot) = G'(\cdot)$. Then the burning *share* in jurisdiction i is

$$b_i \;=\; 1 - G\big(q_i(T, \theta, s_i)\, F\, r(e_i)\big)\,. \tag{5}$$

6.3 Measurement, Payoffs, and Objectives

Measured performance. Detected fires in jurisdiction i are

$$D_i \;=\; M\, b_i\, q_i(T, \theta, s_i). \tag{6}$$

This is the key "optics" metric: lowering q_i via $s_i = 1$ shrinks D_i even if b_i remains high.

[2] If preferred, set clean payoff $-C$ and interpret B as gross benefit of burning; results are unchanged with B replaced by $B + C$.

Hidden fires. I define hidden fires as the difference between imputed and detected fires:

$$H_i = \widehat{Y}_i^{\text{imputed}} - Y_i^{\text{detected}} \tag{7}$$

where $\widehat{Y}_i^{\text{imputed}}$ represents the CNN-predicted fire count based on burned area patterns and Y_i^{detected} represents actual satellite thermal detections. Positive values of H_i indicate fires that escaped thermal detection, likely due to strategic timing.

Bureaucrat. B_i receives wage W, pays convex effort costs and a social cost of strict enforcement, and pays a cost $\kappa \geq 0$ for sharing information. They incur a penalty increasing in detected fires:

$$U_i^B \;=\; W \;-\; \frac{\psi e_i^2}{2} \;-\; \sigma e_i \;-\; \kappa s_i \;-\; \lambda L\, \phi(D_i), \tag{8}$$

with $\psi, \sigma, \lambda, L > 0$ and $\phi'(\cdot) > 0$, $\phi''(\cdot) \geq 0$.

Politician. The politician values both optics (low D_i) and true environmental outcomes (low b_i), and bears monitoring costs:

$$U^P \;=\; -\omega \sum_i D_i \;-\; \beta \sum_i b_i \;-\; \frac{c_1 \theta^2}{2} \;-\; c_2 T, \tag{9}$$

with $\omega, \beta, c_1, c_2 \geq 0$. Setting $\omega > 0$ captures reputational pressure on detected fires; $\beta > 0$ captures concern for true harm.

6.4 Equilibrium Analysis

The game is solved by backward induction.

Stage 3 (farmers). Given (T, θ, s_i, e_i), the burning share b_i is given by (5).

Stage 2 (bureaucrat). B_i chooses (s_i, e_i) to maximize utility, trading off enforcement costs against the penalty from detected fires $D_i = M\, b_i\, q_i$.

For any choice of s_i, the optimal enforcement $e_i^*(s_i)$ satisfies:

$$\psi e_i + \sigma = \lambda L \phi'(D_i) M q_i^2 F g(q_i F r(e_i)) r'(e_i) \tag{10}$$

where $g(\cdot) = G'(\cdot)$ is the density function.

- *Enforcement (no gaming):* $s_i = 0$. With $T = 1$, $q_i = \eta + T(\theta - \eta)$ is high. Raising e_i increases $r(e_i)$, raising expected sanctions and lowering b_i, thus reducing D_i; but effort is costly.
- *Gaming (info sharing):* $s_i = 1$. With $T = 1$, $q_i = \eta + T[(1-\alpha)\theta - \eta]$ is *directly* reduced. B_i can set e_i low (saving effort). Even if b_i rises, $D_i = M\, b_i\, q_i$ can still fall because q_i is smaller.

Proposition 1 (Gaming under tech deployment)**.** *Suppose $T = 1$, r is increasing, G is strictly increasing, and ϕ is increasing and convex. Then there exists a threshold κ^* such that for all $\kappa < \kappa^*$ and α sufficiently large, the unique best response of B_i is to* game*: $s_i^* = 1$ and e_i^* is minimal.*

Proof. Let $e^0 = e_i^*(s_i = 0)$ and $e^1 = e_i^*(s_i = 1)$ denote optimal enforcement under no gaming and gaming respectively. Let $q^0 = \eta + T(\theta - \eta)$ and $q^1 = \eta + T[(1-\alpha)\theta - \eta]$ denote the corresponding detection probabilities.

The utility difference from choosing gaming over enforcement is:

$$\Delta U = U_i^B(s_i = 1, e^1) - U_i^B(s_i = 0, e^0) \tag{11}$$

$$= \left[\frac{\psi (e^0)^2}{2} + \sigma e^0\right] - \left[\frac{\psi (e^1)^2}{2} + \sigma e^1\right] - \kappa \tag{12}$$

$$+ \lambda L \left[\phi(D_i^0) - \phi(D_i^1)\right] \tag{13}$$

where $D_i^j = Mb_i(q^j, e^j)q^j$ for $j \in \{0, 1\}$.

Since $q^1 < q^0$ (when $\alpha > 0$), we have $e^1 < e^0$ from the first-order condition (10), so the effort cost savings are strictly positive.

For the penalty difference, the effect of information sharing on detected fires is:

$$\frac{\partial D_i}{\partial s_i}\Big|_{s_i=0} = M\left[b_i(-\alpha T\theta) + q_i \frac{\partial b_i}{\partial q_i}(-\alpha T\theta)\right] \tag{14}$$

Since $\frac{\partial b_i}{\partial q_i} = -g(q_i Fr(e_i))Fr(e_i) < 0$, we have:

$$\frac{\partial D_i}{\partial s_i}\Big|_{s_i=0} = -M\alpha T\theta\left[b_i - q_i g(q_i Fr(e_i))Fr(e_i)\right] \tag{15}$$

For α sufficiently large, the direct reduction in detection probability dominates any increase in burning share, making $\frac{\partial D_i}{\partial s_i} < 0$. Therefore, $D_i^1 < D_i^0$, and $\phi(D_i^1) < \phi(D_i^0)$.

Thus, $\Delta U > 0$ when κ is sufficiently small, establishing the result. □ □

Stage 1 (politician). The politician chooses (T, θ) anticipating equilibrium responses. The first-order conditions are:

$$\frac{\partial U^P}{\partial \theta} = -\omega \sum_i \frac{\partial D_i}{\partial \theta} - \beta \sum_i \frac{\partial b_i}{\partial \theta} - c_1\theta = 0 \tag{16}$$

$$\frac{\partial U^P}{\partial T} = -\omega \sum_i \frac{\partial D_i}{\partial T} - \beta \sum_i \frac{\partial b_i}{\partial T} - c_2 \lesseqgtr 0 \tag{17}$$

If bureaucrats game ($s_i = 1$ for all i), then increasing θ can reduce $\sum_i D_i$ (improving optics) while having little effect on $\sum_i b_i$ (true burning). This creates a policy paradox where the politician may choose $T = 1$ even anticipating gaming when ω is large relative to β.

6.5 Comparative Statics and Predictions

P1. Hidden Fires Under Monitoring Technology: With T=1 and high α, detected fires D_i fall while true burning b_i changes little, generating "hidden fires" (a widening wedge between actual and detected burning).

P2. **Determinants of gaming:** Gaming is more likely when κ (info-sharing cost) is low, α (evasion efficacy) is high, enforcement costs (ψ, σ) are high, and bureaucratic penalties place weight on detected fires (large ϕ').

7 Conclusion

This paper provides systematic evidence of gaming behavior in response to technology-based environmental monitoring, with important implications for governance and policy design. The case of satellite-based enforcement of agricultural burning regulations in Punjab, India, demonstrates how even sophisticated monitoring systems can be undermined by strategic adaptation when actors have incentives and opportunities to game the system.

The empirical strategy combines a spatial regression discontinuity design with machine learning techniques to identify both the effects of enforcement pressure and the extent of gaming behavior. The findings reveal that apparent success in reducing satellite-detected fires was largely illusory, driven primarily by farmers shifting burning times to evade detection rather than reducing actual burning activity.

These results underscore the importance of Goodhart's law as a cautionary principle for technology adoption in government. When measures become targets, they may cease to be good measures, particularly in contexts where gaming is feasible and incentives for measured performance are strong. The theoretical model developed in the paper helps clarify when such gaming is most likely to emerge and suggests that bureaucratic incentives play a crucial role in determining implementation outcomes.

From a policy perspective, the findings highlight the need for more sophisticated approaches to technology-based governance that anticipate and guard against strategic responses. Effective systems must combine technological capabilities with appropriate institutional safeguards, incentive structures, and monitoring approaches that are robust to gaming.

More broadly, the study contributes to growing evidence that technology is not a panacea for governance problems in developing countries. While satellite monitoring and other advanced technologies offer powerful tools for oversight and enforcement, their effectiveness depends critically on the institutional and political contexts in which they are deployed. Understanding these interactions is essential for realizing the potential benefits of technology for governance while avoiding the pitfalls of naive techno-optimism.

References

Bandiera, O., Prat, A., & Valletti, T. (2009). Active and passive waste in government spending: Evidence from a policy experiment. *American Economic Review*, 99(4), 1278-1308.

Banerjee, A., Hanna, R., Kyle, J., Olken, B. A., & Sumarto, S. (2018). Tangible information and citizen empowerment: Identification cards and food subsidy programs in Indonesia. *Journal of Political Economy*, 126(2), 451-491.

Besley, T., & Ghatak, M. (2005). Competition and incentives with motivated agents. *American Economic Review*, 95(3), 616-636.

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

Dal Bó, E., Finan, F., & Rossi, M. A. (2013). Strengthening state capabilities: The role of financial incentives in the call to public service. *Quarterly Journal of Economics*, 128(3), 1169-1218.

Dipoppa, Gemma, & Saad Gulzar. (2024). "Bureaucrat incentives reduce crop burning and child mortality in South Asia." *Nature*, 634.8036: 1125-1131.

Duflo, E., Greenstone, M., Pande, R., & Ryan, N. (2013). Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *Quarterly Journal of Economics*, 128(4), 1499-1545.

Duflo, E., Hanna, R., & Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4), 1241-1278.

Finan, F., Olken, B. A., & Pande, R. (2017). The personnel economics of the developing state. In *Handbook of Economic Field Experiments* (Vol. 2, pp. 467-514). North-Holland.

Greenstone, M., & Hanna, R. (2014). Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review*, 104(10), 3038-3072.

Jacob, B. A., & Levitt, S. D. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics*, 118(3), 843-877.

Kahn, M. E., Li, P., & Zhao, D. (2015). Water pollution progress at borders: The role of changes in China's political promotion incentives. *American Economic Journal: Economic Policy*, 7(4), 71-103.

Lipscomb, M., & Mobarak, A. M. (2017). Decentralization and pollution spillovers: Evidence from the re-drawing of county borders in Brazil. *Review of Economic Studies*, 84(1), 464-502.

Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, 115(2), 200-249.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature*, 37(1), 7-63.

Sigman, H. (2002). International spillovers and water quality in rivers: Do countries free ride? *American Economic Review*, 92(4), 1152-1159.