

Detection or Deterrence? A Measurement Problem in Enforcement-Generated Safety Data

APEP Autonomous Research* @olafdrw

March 9, 2026

Abstract

Regulatory databases record *detected* incidents, not actual ones. I document this measurement problem using France’s ARIA industrial accident database (56,034 records, 1992–2010) and variation in Seveso site density across 97 departments. Departments with greater hazardous-industry concentration show differentially more reported accidents over time—but only minor ones. Severe and fatal accidents, which are too consequential to escape detection regardless of enforcement intensity, show no differential change. This divergence is robust to Poisson count models, region-by-year fixed effects, and sample restrictions, but disappears with department-specific linear trends, indicating that pre-existing differential trends in reporting infrastructure drive the association. The severity decomposition provides a diagnostic for any setting where enforcement-generated data are used to evaluate policy: only detection-inelastic outcomes can credibly measure deterrence.

JEL Codes: K32, L51, Q58, H83

Keywords: industrial safety, regulatory enforcement, disaster salience, detection, deterrence, Seveso, ARIA

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

1. Introduction

On September 21, 2001, a warehouse containing 300 tonnes of ammonium nitrate exploded at the AZF chemical plant in Toulouse, France. The blast killed 31 people, injured 2,500, and damaged 27,000 buildings in what the French Senate later called the country’s worst industrial accident since the 1906 Courrières mine disaster. The political response was immediate and far-reaching: within two years, France enacted the Loi n°2003-699 du 30 juillet 2003, which mandated technological risk prevention plans (Plans de Prévention des Risques Technologiques, or PPRTs) for all Seveso high-threshold installations and doubled the inspectorate from approximately 800 to 1,600 officers.

A natural question is whether this regulatory expansion improved safety. But answering it confronts a fundamental measurement problem: the very databases governments use to track safety—records of industrial accidents—are themselves products of the enforcement process. More inspectors discover more incidents, inflating reported counts even if actual danger declines. Any evaluation based on incident counts conflates two channels: *detection* (inspectors find events that would otherwise go unreported) and *deterrence* (firms invest more in safety to avoid penalties). When both channels operate simultaneously, raw accident counts are ambiguous indicators of policy effectiveness.

This paper documents this measurement problem empirically and proposes a diagnostic for resolving it. Using France’s comprehensive ARIA industrial accident database—56,034 classified records from 1992 to 2010—I construct a department-year panel and examine how reported accidents co-vary with pre-existing Seveso site density, a proxy for enforcement exposure. Departments with more hazardous installations show differentially larger increases in reported accidents over time: each log-unit of Seveso exposure is associated with 2.97 additional reported accidents per department-year in the baseline specification. But when I disaggregate by severity, the entire association is concentrated in minor incidents—small spills, contained releases, near-misses. Severe accidents (reaching level 3 or higher on the European severity scale) show no differential change ($\hat{\beta} = 0.12$, SE = 0.13). Fatal accidents are similarly flat ($\hat{\beta} = 0.014$, SE = 0.029).

The severity decomposition is the paper’s central contribution. Minor incidents are *detection-elastic*: their recording depends on whether an inspector, worker, or reporting system captures them. Fatal and severe accidents are *detection-inelastic*: casualties, explosions, and visible environmental damage trigger reporting regardless of enforcement intensity. By comparing outcomes at different points on this detection-elasticity gradient, I can separate the data signature of detection from that of deterrence. The consistent finding—more minor reports, unchanged severe outcomes—reveals that the growth in total reported incidents

reflects expanded reporting capacity rather than deteriorating safety.

An important caveat: this association does not isolate the causal effect of the Loi 2003. The parallel trends assumption fails: a placebo test at 1997 produces a coefficient ($\hat{\beta} = 3.45$) larger than the post-2003 estimate, and joint pre-trend tests reject for both total and severe accidents. Most decisively, adding department-specific linear trends eliminates the total-accident association entirely ($\hat{\beta} = -0.36$, SE = 1.61), confirming that the baseline estimate reflects pre-existing differential trends in reporting infrastructure. These trends likely reflect the maturation of the ARIA database itself and EU-driven harmonization of reporting obligations during the 1990s, both of which proceeded faster in departments with more classified installations.

I therefore frame the paper’s contribution as a measurement insight rather than a causal evaluation. The severity decomposition is informative regardless of what drives the differential trends: whether the divergence between minor and severe outcomes reflects post-2003 enforcement expansion, longer-running reporting infrastructure maturation, or some combination, the pattern reveals a structural feature of enforcement-generated administrative data. This feature has practical implications: any evaluation of regulatory effectiveness that relies on inspection-generated incident counts must account for the detection channel, and the severity gradient provides a ready-made diagnostic for doing so.

This paper contributes to several literatures. First, it speaks to the regulation and enforcement literature, where a central question is whether monitoring deters non-compliance or merely detects it (Gray, 1993; Shimshack and Ward, 2007; Duflo et al., 2013, 2018; Johnson, 2020). The canonical framework derives from Becker (1968) and Stigler (1970): a rational firm reduces violations when the expected penalty rises. But empirical tests have struggled to separate detection from deterrence because most regulatory data measure *detected* violations, not actual ones. The severity decomposition I propose provides a diagnostic for this distinction by exploiting a natural hierarchy of detection elasticity across incident types.

Second, it contributes to the literature on administrative data endogeneity in regulatory settings. A growing body of work recognizes that enforcement-generated data conflate monitoring intensity with underlying incidence—in workplace safety (Gray and Scholz, 1998; Johnson, 2020), environmental compliance (Shimshack and Ward, 2007; Bennear, 2007), and crime reporting (Levitt, 1997; Chalfin and McCrary, 2018). This paper provides a concrete diagnostic for identifying when this conflation is operative.

Third, it engages the modern difference-in-differences literature on pre-trends and identification. Roth (2022) shows that pre-trend tests have limited power to detect violations, while Rambachan and Roth (2023) develop sensitivity analysis for settings where parallel trends may fail. My paper illustrates their warnings in a natural setting: the baseline specification

appears to show a significant post-treatment shift, but this reflects pre-existing differential trends that standard DiD absorbs into the treatment coefficient. The department-specific trend specification makes this transparent. The continuous-treatment design is structurally similar to shift-share approaches analyzed by [Goldsmith-Pinkham et al. \(2020\)](#) and [Borusyak et al. \(2022\)](#), where the identifying assumption rests on the exogeneity of the shares (here, Seveso density).

Fourth, it introduces the ARIA database—63,365 classified industrial accident records spanning 1992 to 2025—to the economics literature. This dataset has been used exclusively in engineering and process safety journals; no prior economics paper has exploited its panel structure. The multi-dimensional severity coding provides an unusually rich outcome variable for studying detection-elasticity gradients.

The closest antecedents are [Greenstone \(2002\)](#) on Clean Air Act enforcement, [Gray \(1993\)](#) on OSHA inspections, [Duflo et al. \(2013, 2018\)](#) on randomized auditing in India, and [Hanna and Oliva \(2014\)](#) on EPA inspection effects on emissions. My contribution differs in emphasis: rather than estimating the causal effect of a specific enforcement expansion, I use the French case to demonstrate and diagnose the measurement problem that confounds all such evaluations. The severity decomposition is portable to any setting where enforcement generates the outcome data used for evaluation.

The remainder of the paper proceeds as follows. [Section 2](#) describes the AZF disaster and the resulting regulatory framework. [Section 3](#) develops a simple conceptual framework for detection and deterrence. [Section 4](#) presents the ARIA database and the construction of the analysis panel. [Section 5](#) lays out the empirical strategy and its assumptions. [Section 6](#) presents results. [Section 7](#) discusses threats to validity, with particular attention to the pre-trend evidence. [Section 8](#) concludes.

2. Institutional Background

2.1 The AZF Explosion and Its Aftermath

The AZF plant in Toulouse belonged to Grande Paroisse, a subsidiary of TotalFinaElf, and produced nitrogen-based fertilizers and chemical intermediates. It was classified as a Seveso “seuil haut” (high-threshold) installation under European Directive 96/82/EC, meaning it stored hazardous substances above quantities that required enhanced safety management, emergency planning, and regulatory oversight.

At 10:17 AM on September 21, 2001, approximately 20–120 tonnes of ammonium nitrate stored in a warehouse detonated. The blast registered 3.4 on the Richter scale, shattered windows across a 3-kilometer radius, and collapsed factory buildings. Thirty-one people were

killed—21 at the plant and 10 in the surrounding area—and more than 2,500 were injured. The explosion caused an estimated 2.3 billion in damage, making it the costliest industrial accident in French history.

The timing—ten days after September 11, 2001—initially raised terrorism fears, but investigators ultimately concluded the cause was accidental chemical contamination of the ammonium nitrate stockpile. Criminal proceedings continued for over a decade, resulting in 2017 convictions of the plant’s parent company for negligence.

2.2 The Loi du 30 juillet 2003

The French Parliament’s response was the Loi n°2003-699 du 30 juillet 2003 relative à la prévention des risques technologiques et naturels et à la réparation des dommages. Its industrial safety provisions included:

1. **Mandatory PPRTs:** All Seveso “seuil haut” installations were required to develop Plans de Prévention des Risques Technologiques within five years. These plans defined zones of danger around each facility and restricted new development within them.
2. **Inspector doubling:** The number of Inspections des Installations Classées (ICPE) officers was increased from approximately 800 to 1,600 between 2003 and 2008. The expansion was concentrated in regions with higher densities of classified installations.
3. **Enhanced reporting requirements:** Firms operating classified installations faced strengthened obligations to report incidents, near-misses, and safety-relevant events to the DREAL (Direction Régionale de l’Environnement, de l’Aménagement et du Logement).
4. **Increased penalties:** Fines for non-compliance were raised, and prefectural powers to order plant closures were expanded.

The critical feature for identification is that the inspector expansion was not uniform. Departments with more Seveso high-threshold sites—which housed the largest concentrations of hazardous industrial activity—received proportionally more additional inspectors and faced a larger PPRT compliance burden. This creates continuous variation in enforcement intensity across departments.

2.3 The Seveso Framework

France’s industrial risk regulation operates within the European Seveso framework (now Directive 2012/18/EU, originally 82/501/EEC following the 1976 Seveso dioxin disaster in

Italy). Establishments that handle hazardous substances above specified quantity thresholds are classified as either “seuil bas” (lower tier) or “seuil haut” (upper tier). As of the analysis period, approximately 737 establishments held *seuil haut* classification across metropolitan France, distributed unevenly: Seine-Maritime (76) hosted the most with over 50 sites, while eight departments had none.

This geographic variation reflects the historical geography of French heavy industry: petrochemicals along the Rhône valley and around Fos-Marseille, chemical manufacturing in Lyon and the northeast, petroleum refining in Normandy, and dispersed facilities across the industrial departments of the north. The variation is persistent because Seveso classification depends on stored quantities of specific substances rather than on incident history or regulatory discretion.

The Seveso framework is important for identification because it creates a pre-determined, slowly-moving measure of industrial hazard density that is correlated with the intensity of the post-AZF regulatory expansion but determined by the historical geography of French industrialization. Departments with more Seveso *seuil haut* sites faced proportionally more PPRT obligations, received more additional inspectors, and experienced a greater increase in enforcement activity. Crucially, the Seveso classification itself was not altered by the Loi 2003—the law changed enforcement procedures for existing classified installations rather than reclassifying establishments.

2.4 Implementation Timeline

The Loi 2003 was not implemented instantaneously. The law was enacted on July 30, 2003, but the operational expansion unfolded over several years. Inspector recruitment proceeded through the standard French civil service channels (*concours*), with the bulk of new inspectors entering service between 2004 and 2008. PPRT preparation followed an even longer timeline: the law required PPRTs for all Seveso *seuil haut* installations within five years, but the actual completion rate was slow—by 2010, fewer than 30% of required PPRTs had been approved, with full compliance not achieved until the mid-2010s.

This gradual implementation has implications for the empirical analysis. The post-2003 indicator in my main specification captures the intent-to-treat effect of the law’s enactment, which includes anticipation effects (firms adjusting behavior in expectation of stricter enforcement), early implementation effects (the first wave of new inspectors), and full-implementation effects (the complete inspector expansion). The event study specification allows me to trace how the treatment effect evolved over this implementation window.

3. Conceptual Framework

3.1 A Simple Model of Detection and Deterrence

Consider a firm that generates industrial incidents at rate λ , which depends on its safety investment s : $\lambda(s)$ with $\lambda' < 0$ and $\lambda'' > 0$. Each incident has a severity v drawn from a distribution $F(v)$ that is independent of enforcement intensity. The firm faces regulatory inspections at rate ϕ (inspectors per site) and a penalty $\pi(v)$ per detected incident, where $\pi' > 0$ (more severe incidents carry larger penalties).

Not all incidents are observed. Define the detection probability as $d(v, \phi)$, which depends on both severity and enforcement intensity. The key structural assumption is:

$$\frac{\partial d}{\partial \phi} \text{ is decreasing in } v \quad (1)$$

That is, inspector presence increases the detection probability more for minor incidents than for severe ones. At the extreme, fatal industrial accidents (v very large) satisfy $d(v, \phi) \approx 1$ for all $\phi > 0$: they are effectively detection-inelastic because their consequences—casualties, explosions, visible environmental damage—ensure reporting regardless of inspector presence. Small spills, contained releases, and near-misses, by contrast, may go entirely unreported absent an inspector who discovers them during a routine site visit.

The observed accident count in a department i with enforcement intensity ϕ_i is:

$$R_i = \lambda(s_i^*) \int d(v, \phi_i) dF(v) \quad (2)$$

where s_i^* is the firm's optimal safety investment given ϕ_i . An increase in ϕ_i affects R_i through two channels:

Detection channel: $\left. \frac{\partial R}{\partial \phi} \right|_s > 0$. More inspectors discover more incidents, raising reported counts.

Deterrence channel: $\left. \frac{\partial R}{\partial \phi} \right|_d = \lambda'(s^*) \frac{\partial s^*}{\partial \phi} \int d(v, \phi) dF(v) < 0$. Firms invest more in safety when inspection probability rises, reducing actual incident rates.

The total effect on reported accidents is ambiguous: $\frac{dR}{d\phi} \gtrless 0$ depending on which channel dominates. This is the regulatory measurement problem. A naïve comparison of reported accident counts before and after an enforcement expansion cannot distinguish increased detection from decreased safety.

3.2 The Severity Decomposition

The key empirical insight is that decomposing reported incidents by severity resolves this ambiguity. Define two categories:

$$R_i^{\text{minor}} = \lambda(s_i^*) \int_{v < \bar{v}} d(v, \phi_i) dF(v) \quad (3)$$

$$R_i^{\text{severe}} = \lambda(s_i^*) \int_{v \geq \bar{v}} d(v, \phi_i) dF(v) \quad (4)$$

where \bar{v} is the severity threshold. For severe incidents, the detection probability is approximately 1 regardless of ϕ :

$$d(v, \phi) \approx 1 \quad \text{for } v \geq \bar{v} \quad (5)$$

This means $R_i^{\text{severe}} \approx \lambda(s_i^*)[1 - F(\bar{v})]$ —severe reported counts are approximately equal to actual severe incident rates, providing a clean test of deterrence. If $\frac{\partial R^{\text{severe}}}{\partial \phi} \approx 0$, either deterrence is absent or it is operating but of negligible magnitude.

For minor incidents, both channels are active. An increase in reported minor incidents after enforcement expansion is consistent with detection dominating deterrence, but it does not rule out some deterrence operating simultaneously. The observable prediction is:

$$\frac{\partial R^{\text{minor}}}{\partial \phi} > 0 \quad \text{and} \quad \frac{\partial R^{\text{severe}}}{\partial \phi} \approx 0 \quad (6)$$

which is precisely what I find in the data.

3.3 Implications for Interpretation

This framework clarifies what the data can and cannot tell us. A null effect on severe accidents rules out large deterrence effects but cannot distinguish between (a) zero deterrence and (b) deterrence that is small relative to statistical precision. A positive effect on minor accidents is consistent with detection dominating deterrence for low-severity events, but the magnitude of deterrence for minor events is not separately identified. The framework also highlights that the relevant treatment variable is enforcement intensity (ϕ), not inspector headcount per se; to the extent that doubling inspectors less than doubles effective enforcement (due to training lags, organizational frictions, or diminishing returns to monitoring), the estimates below should be interpreted as intent-to-treat effects of the regulatory expansion.

4. Data

4.1 The ARIA Database

The ARIA (Analyse, Recherche et Information sur les Accidents) database is maintained by the Bureau d'Analyse des Risques et Pollutions Industriels (BARPI) within the French Ministry of Ecological Transition. It records industrial accidents, incidents, and near-misses primarily from installations classées (ICPE facilities), though it also includes transport incidents and other technological events. Systematic recording began in 1992, following the creation of BARPI in 1992 with the explicit mandate to centralize information on technological accidents.

The ARIA database serves a dual purpose relevant to this study. First, it provides operational intelligence: prefects, inspectors, and plant operators use ARIA records to learn from past incidents and identify recurring failure modes. Second, it serves as the administrative backbone for regulatory evaluation—the Ministry uses aggregate ARIA counts to assess whether enforcement is achieving its safety objectives. This dual function creates the measurement problem I study: the database that is supposed to measure safety outcomes is itself a product of the enforcement process.

I download the complete ARIA database from data.gouv.fr, which contains 63,365 records spanning 1992–2025 as of January 2026.¹ The main analysis restricts to metropolitan France and the 1992–2010 window (56,034 records); the extended-panel analysis uses 1992–2020. Each record includes the date, department, commune, type of installation, materials involved, consequences, and—crucially—a multi-dimensional severity scale. The database is publicly accessible, and BARPI periodically updates it as new incidents are recorded and older records are verified. Records are classified by event type (accident, incident, near-miss) and by the type of installation involved (ICPE, transport, other).

4.2 Severity Classification

The severity scale follows European conventions with four dimensions: human consequences (H), environmental damage (En), economic losses (Ec), and material quantities released (M). Each dimension is scored 0–6, where higher values indicate greater severity. This multi-dimensional coding is central to the empirical strategy, as it allows me to classify incidents along a detection-elasticity gradient.

I define three severity categories:

¹Dataset ID: 5eb95b0d74b7d58155692f9b on <https://data.gouv.fr>.

- **Fatal** ($H \geq 4$): Events involving one or more deaths. These are detection-inelastic: a fatal industrial accident is reported regardless of inspector presence.
- **Severe** ($\max(H, E_n, E_c, M) \geq 3$): Significant events with major human, environmental, or economic consequences.
- **Minor** ($\max \text{ score} < 2$): Small spills, contained releases, near-misses, and equipment failures that an inspector might discover during a routine visit but that a firm might not voluntarily report.

After restricting to metropolitan France and the analysis window (1992–2010), the dataset contains 56,034 records. The mutually exclusive breakdown by severity is: 664 fatal events (human severity ≥ 4), 3,107 severe non-fatal events ($\max \text{ severity} \geq 3$ but human < 4), 4,652 moderate events ($\max \text{ severity} = 2$), and 47,611 minor events ($\max \text{ severity} < 2$). Note that fatal events are a subset of the broader “severe” category ($\max \geq 3$): 3,771 total severe events include the 664 fatal cases. In the regression analysis, I use overlapping categories—severe (≥ 3 , including fatal) and fatal (≥ 4)—as separate outcomes to test detection elasticity at different severity thresholds. The classification of incidents as “minor” is central to the detection argument: these events—small chemical spills quickly contained, equipment malfunctions without personnel exposure, brief releases below reportable thresholds—are precisely the type that an inspector might discover during a routine site visit (examining maintenance logs, interviewing workers, checking storage conditions) but that a plant operator would have little incentive to report voluntarily if no inspector were present. Fatal and severe events, by contrast, typically involve emergency response, hospital admissions, or visible environmental contamination that triggers mandatory reporting through channels independent of the industrial inspection system.

One limitation of the ARIA data is that recording standards evolved over the sample period. BARPI itself acknowledges that early-1990s coverage is less comprehensive than later years, as the database infrastructure was still being developed. This time-varying completeness potentially confounds the detection interpretation: some of the increase in recorded minor incidents may reflect improvements in BARPI’s own data-capture systems rather than (or in addition to) inspector-driven detection. I address this concern in two ways. First, the difference-in-differences design absorbs any aggregate time trend through year fixed effects; the identifying variation comes from differential changes across departments with different Seveso density. Second, if BARPI’s data-capture improvements were correlated with Seveso density (e.g., priority investment in departments with more hazardous sites), this would inflate the estimated detection effect—but it would not affect the severe-accident null, since severe events were reliably captured throughout the sample period.

4.3 Seveso Site Counts

I obtain the geographic distribution of Seveso seuil haut installations from the Georisques API (georisques.gouv.fr), which provides the current ICPE registry including Seveso classification status, commune code, and geographic coordinates. I query all metropolitan departments and count installations classified as “Seveso seuil haut.”

The resulting distribution shows 737 Seveso seuil haut sites across 88 departments with at least one such site. Nine departments have zero Seveso seuil haut installations, for a total of 97 metropolitan department codes in the analysis panel.

One caveat: the Georisques registry reflects current Seveso classifications rather than historical status. Some sites may have been added or removed since 2001. However, Seveso classification is driven by stored quantities of specific substances and changes slowly; the main source of attrition is plant closure rather than reclassification. I treat the current count as a proxy for the pre-2001 distribution.

4.4 Panel Construction

I construct a balanced department-year panel with $97 \text{ departments} \times 19 \text{ years} = 1,843$ observations. For each department-year cell, I compute:

- Total ARIA-recorded accidents
- Severe accidents (max severity ≥ 3)
- Fatal accidents (human severity ≥ 4)
- Minor incidents (max severity < 2)
- Installation Classée accidents (type = IC)

Department-year cells with no recorded accidents are coded as zero. The treatment variable is $D_i = \log(\text{Seveso H sites}_i + 1)$, which ranges from 0 to 4.04.

4.5 Summary Statistics

Table 1: Summary Statistics

Variable	Pre-AZF (1992–2002)		Post-Loi 2003 (2003–2010)		Full Sample	
	Mean	SD	Mean	SD	Mean	SD
Total accidents	14.56	14.18	17.28	15.94	15.71	15.00
Severe accidents (scale ≥ 3)	1.05	1.58	1.05	1.55	1.05	1.57
Fatal accidents (human ≥ 4)	0.13	0.39	0.22	0.50	0.17	0.45
Minor incidents (scale < 2)	12.40	12.40	14.65	13.56	13.35	12.95
Seveso seuil haut sites	7.60	10.13	7.60	10.13	7.60	10.13

Notes: $N = 1,843$ department-year observations (97 departments $\times 19$ years). Accidents are annual counts per department from the ARIA database. Severity classifications based on European multi-dimensional scale (H, En, Ec, M dimensions, scored 0–6). Seveso seuil haut sites from Georisques ICPE registry; this variable is time-invariant by construction (measured at a single cross-section from the current registry as a proxy for pre-2001 distribution), hence identical pre- and post-period means.

The summary statistics reveal a suggestive pattern. Total accidents per department-year increased from 14.56 in the pre-period to 17.28 in the post-period (a 19% increase). But severe accidents were unchanged at 1.05 per department-year in both periods. Minor incidents drove the aggregate increase, rising from 12.40 to 14.65. Seveso site density and total accident counts are strongly correlated in the cross-section ($r = 0.72$), confirming that industrial hazard exposure is a relevant predictor of accident reporting.

5. Empirical Strategy

5.1 Continuous Treatment Difference-in-Differences

I estimate the following specification:

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot D_i \times \text{Post2003}_t + \varepsilon_{it} \quad (7)$$

where Y_{it} is the count of industrial accidents in department i in year t , α_i are department fixed effects, γ_t are year fixed effects, $D_i = \log(\text{Seveso H sites}_i + 1)$ is the continuous treatment intensity, and $\text{Post2003}_t = \mathbb{I}[t \geq 2003]$ indicates the post-reform period.

The coefficient β captures whether departments with higher Seveso site density—and therefore greater exposure to the regulatory expansion—experienced differentially larger

changes in accident counts after the Loi 2003. Standard errors are clustered at the department level to account for serial correlation within departments (Bertrand et al., 2004).

The log transformation of Seveso counts serves two purposes. First, it accommodates the highly skewed distribution of Seveso sites: the top department (Seine-Maritime) has 56 sites while the median department has 5, and the raw count would give disproportionate influence to a few extreme observations. Second, the log specification implies that the marginal effect of an additional Seveso site diminishes as density increases, which is consistent with the institutional reality: in departments already saturated with classified installations, the marginal inspector-to-site ratio from the national expansion was smaller.

The continuous treatment design differs from a standard binary DiD in that there is no clean “control group” of untreated units. Every department with at least one Seveso site received some increase in enforcement intensity, and even zero-Seveso departments may have experienced spillovers through national-level changes in reporting norms. The identifying variation comes from the *dose*: departments with more Seveso sites are assumed to have received a proportionally larger enforcement increase. This is structurally similar to shift-share or Bartik approaches (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022), where a national shock is distributed according to pre-determined shares. As Goldsmith-Pinkham et al. (2020) emphasize, identification in such designs requires that the shares (here, Seveso density) be exogenous to differential outcome trends conditional on fixed effects—a condition that the pre-trend evidence in this paper shows is not satisfied.

5.2 Event Study

To examine the timing of effects and test for pre-existing trends, I estimate an event study specification:

$$Y_{it} = \alpha_i + \gamma_t + \sum_{\tau \neq 0} \beta_\tau \cdot D_i \times \mathbb{I}[t - 2001 = \tau] + \varepsilon_{it} \quad (8)$$

where $\tau = 0$ (the year of the AZF explosion, 2001) is the omitted reference period. The coefficients β_τ for $\tau < 0$ test parallel pre-trends; those for $\tau > 0$ trace out the dynamic treatment effect.

5.3 Identifying Assumption and Threats

The identifying assumption is that, absent the AZF explosion and Loi 2003, departments with different Seveso site densities would have experienced parallel trends in industrial accidents. This is testable over the nine pre-treatment years (1992–2001).

Several threats warrant discussion:

Reporting bias. If more inspectors increase the probability that a given accident is recorded, then rising accident counts may reflect detection rather than true incidence. This is the core identification challenge—not a nuisance to be controlled for, but the phenomenon I study. The severity decomposition turns reporting bias from a threat to identification into the paper’s central empirical insight: the differential response across severity categories allows me to decompose the aggregate change into detection and deterrence components.

Industrial composition. Departments with more Seveso sites are more heavily industrialized and may have experienced different deindustrialization trajectories during the 1990s and 2000s. Divergent secular trends in industrial activity could drive differential accident reporting independently of the Loi 2003. Department fixed effects absorb persistent level differences in industrial activity, and year fixed effects absorb national trends in deindustrialization and reporting norms. The identifying concern is whether department-specific trends in industrial composition are correlated with Seveso density and with accident reporting—a concern that the parallel trends test directly addresses (and, as I discuss in [Section 7](#), does not fully resolve).

Contemporaneous shocks. The 2001–2003 period includes other events (the September 11 attacks, the 2001–2002 recession) that may have differentially affected industrial departments. Year fixed effects absorb aggregate shocks, but any shock correlated with Seveso density and accident outcomes would bias β . The most plausible confound is the heightened public attention to industrial risk following the AZF explosion itself: even absent the Loi 2003, the disaster may have changed reporting behavior in departments perceived as high-risk. This concern is partially addressed by the fact that the estimated effect is stable when excluding Toulouse (the AZF site), but it cannot rule out a national attention effect that differentially affected high-Seveso departments.

SUTVA violations. The stable unit treatment value assumption requires that one department’s treatment does not affect another’s outcomes. This could be violated if the inspector expansion involved redeployments from low-Seveso to high-Seveso departments (reducing monitoring in the former) or if firms relocated operations in response to differential enforcement pressure. I do not observe inspector assignments directly, but the national doubling of the inspectorate suggests new hiring rather than reallocation, making the first concern less likely. Firm relocation over the 8-year post-treatment window is possible but slow in heavy industry, where Seveso-classified facilities involve fixed infrastructure that cannot be easily moved.

6. Results

6.1 Main Results

Table 2 presents the baseline difference-in-differences estimates across seven outcome variables. The first column shows that total reported accidents are positively associated with Seveso density in the post-2003 period: each log-unit of Seveso site density is associated with 2.97 additional accidents per department-year ($p = 0.033$). As I show in the robustness section, this association is not robust to department-specific linear trends, indicating that it reflects pre-existing differential trajectories rather than a clean post-treatment shift. Nevertheless, the *pattern across severity categories* is informative regardless of what drives the level effect.

Table 2: Effect of Regulatory Expansion on Industrial Accidents

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Total	Severe	Fatal	Minor	IC Only	Log(Total+1)	Log(Severe+1)
Seveso density \times Post-2003	2.970** (1.390)	0.125 (0.134)	0.014 (0.029)	2.386* (1.218)	2.752** (1.283)	0.024 (0.043)	0.037 (0.027)
Observations	1,843	1,843	1,843	1,843	1,843	1,843	1,843
Mean dep. var.	15.71	1.05	0.17	13.35	14.45	2.39	0.57
Adj. R^2	0.676	0.388	0.107	0.649	0.644	0.712	0.349
Department FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the department level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Treatment variable is $\log(\text{Seveso H sites} + 1) \times \text{Post-2003}$. All specifications include department and year fixed effects. Outcomes are annual department-level counts from the ARIA database, 1992–2010. Severe = max severity ≥ 3 ; Fatal = human severity ≥ 4 ; Minor = max severity < 2 ; IC Only = Installation Classée accidents only.

The decomposition by severity reveals the paper’s central pattern. Column 2 shows that severe accidents show no differential association ($\hat{\beta} = 0.125$, SE = 0.134). Column 3 shows that fatal accidents are similarly flat ($\hat{\beta} = 0.014$, SE = 0.029). In contrast, Column 4 shows that minor incidents account for the aggregate association ($\hat{\beta} = 2.386$, $p = 0.050$), representing 80% of the total. Column 5 shows the association is concentrated in Installation Classée accidents ($\hat{\beta} = 2.752$, $p = 0.032$), consistent with reporting systems focused on classified facilities.

The log specifications in Columns 6–7 show statistically insignificant effects. These coefficients represent the semi-elasticity of $\log(Y + 1)$ with respect to one unit of the treatment

variable $D_i = \log(\text{Seveso} + 1)$. The level-specification coefficient of 2.97 divided by the mean (15.71) gives a 19% proportional change, but the log-log specification absorbs this differently: the log outcome compresses the variation in count data, and the estimated semi-elasticity of 0.024 is attenuated by the nonlinear transformation.

The magnitude of the total-accident effect merits interpretation. The coefficient of 2.97 implies that moving from the 25th percentile of Seveso density (1 site, $D = 0.69$) to the 75th percentile (11 sites, $D = 2.48$) is associated with $2.97 \times (2.48 - 0.69) = 5.31$ additional reported accidents per year after 2003. For a department at the median Seveso density (5 sites), the implied increase is roughly 3 additional accident reports per year—one new detection every four months. Given that the inspector expansion approximately doubled the inspection workforce, this suggests a relatively low yield: each additional inspector discovers, on average, fewer than one additional minor incident per year per inspected site.

It is worth noting what the null on severe accidents rules out. The 95% confidence interval on the severe-accident coefficient is $[-0.14, 0.39]$. With a pre-treatment mean of 1.05 severe accidents per department-year, this rules out effects larger than a 37% increase or a 13% decrease. While I cannot rule out small deterrence effects, a large deterrence effect masking an equally large detection effect would require an implausible coincidence.

6.2 Event Study

Figure 1 presents event study estimates for total and severe accidents. The top panel reveals two important patterns. First, there is a visible upward shift in total reported accidents beginning around 2002–2003, consistent with the timing of the Loi 2003 and inspector expansion. Second, and more concerning, several pre-treatment coefficients are individually and jointly significant (joint $F = 3.62$, $p < 0.001$), suggesting that high-Seveso-density departments were already on differential reporting trajectories before the AZF event.

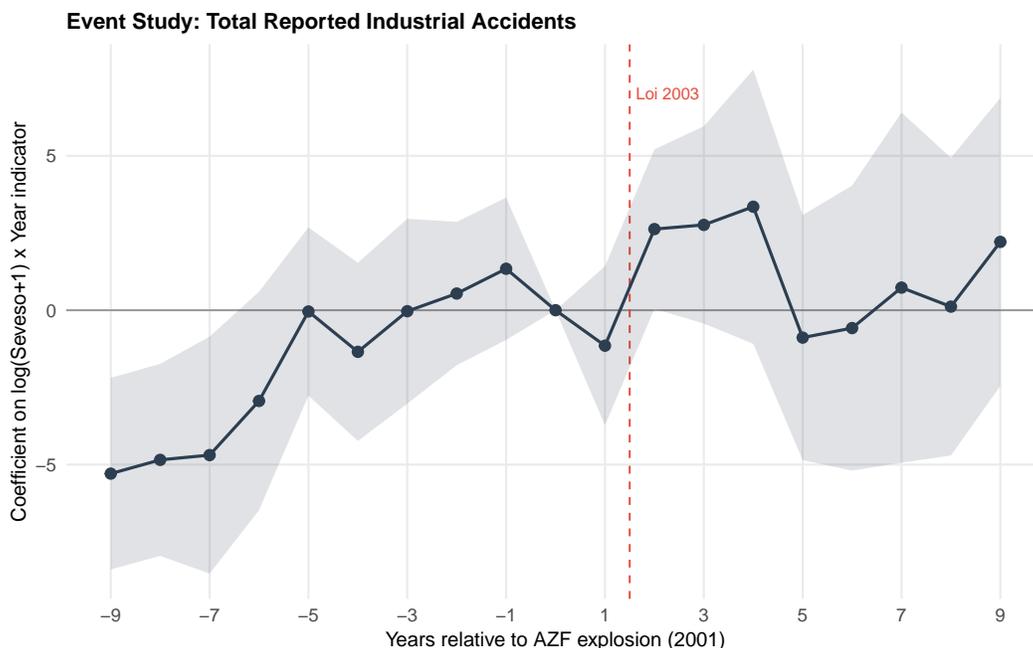


Figure 1: Event Study: Total Reported Industrial Accidents

Notes: Coefficients from Equation (8) with 95% confidence intervals. The reference year is 2001 (the year of the AZF explosion). Standard errors clustered at the department level.

The severe accident event study (Figure 2) shows no corresponding shift after 2003, with coefficients fluctuating around zero throughout the sample period. The joint pre-trend test for severe accidents also rejects ($F = 2.11$, $p = 0.026$), though the individual coefficients are smaller in magnitude and economically modest. The contrast between the two event studies is the paper’s central visual: total accidents exhibit what appears to be a level shift, whereas severe accidents show no discernible pattern change. But as Roth (2022) emphasizes, the significant pre-trend coefficients for total accidents warn that this apparent shift may simply reflect the continuation of pre-existing differential trends.

Two features of the event study deserve comment. First, the pre-treatment coefficients for total accidents are uniformly negative and trending upward, which is consistent with the parallel trends concern discussed in Section 7: high-Seveso departments were on an upward trajectory relative to low-Seveso departments throughout the 1990s. The coefficient at $\tau = -1$ (year 2000) is near zero, and the positive shift begins around $\tau = 1$ (2002), consistent with the timing of the AZF explosion and early legislative response. Second, the post-treatment coefficients are relatively stable between $\tau = 2$ and $\tau = 9$, suggesting a persistent level shift rather than a transitory spike. This persistence is consistent with a structural change in detection infrastructure (more inspectors conducting routine visits) rather than a temporary

reporting response to the disaster.

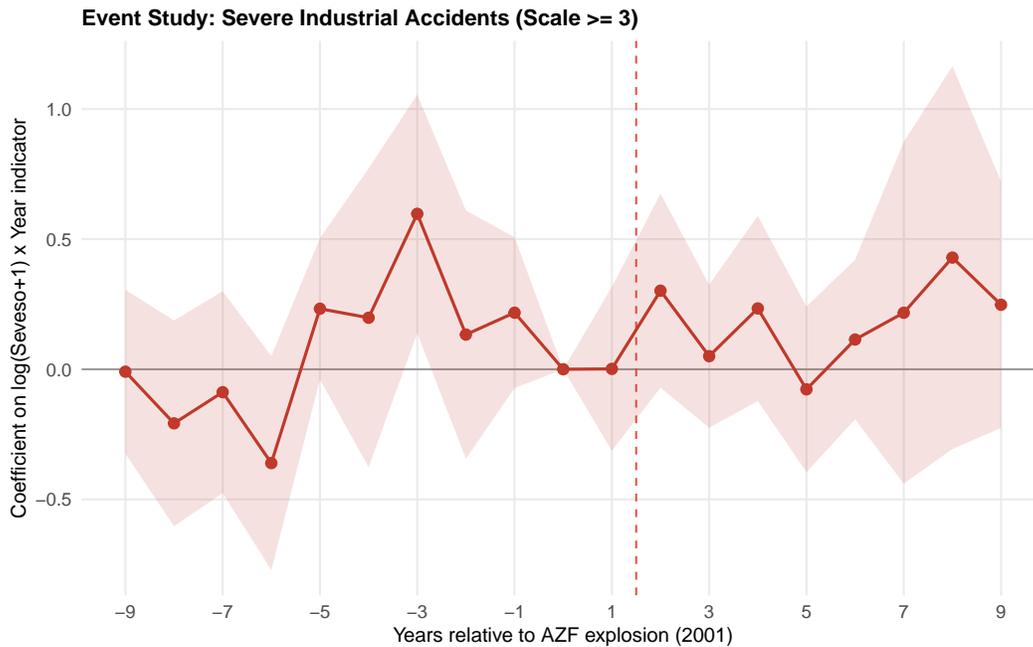


Figure 2: Event Study: Severe Industrial Accidents (Scale ≥ 3)

Notes: Same specification as [Figure 1](#) but with severe accident count (maximum severity ≥ 3) as the outcome. Note the different y-axis scale.

6.3 Detection versus Deterrence

The divergence between minor and severe outcomes is the paper's central finding. [Figure 3](#) plots the event study coefficients for both categories on the same axes. Minor incidents show a clear upward break after 2003 in high-Seveso departments, while severe accidents remain flat. This pattern is consistent with a detection-dominated response: inspectors discover minor events that would otherwise go unreported, but do not prevent (or do prevent but also detect an offsetting number of) severe incidents.

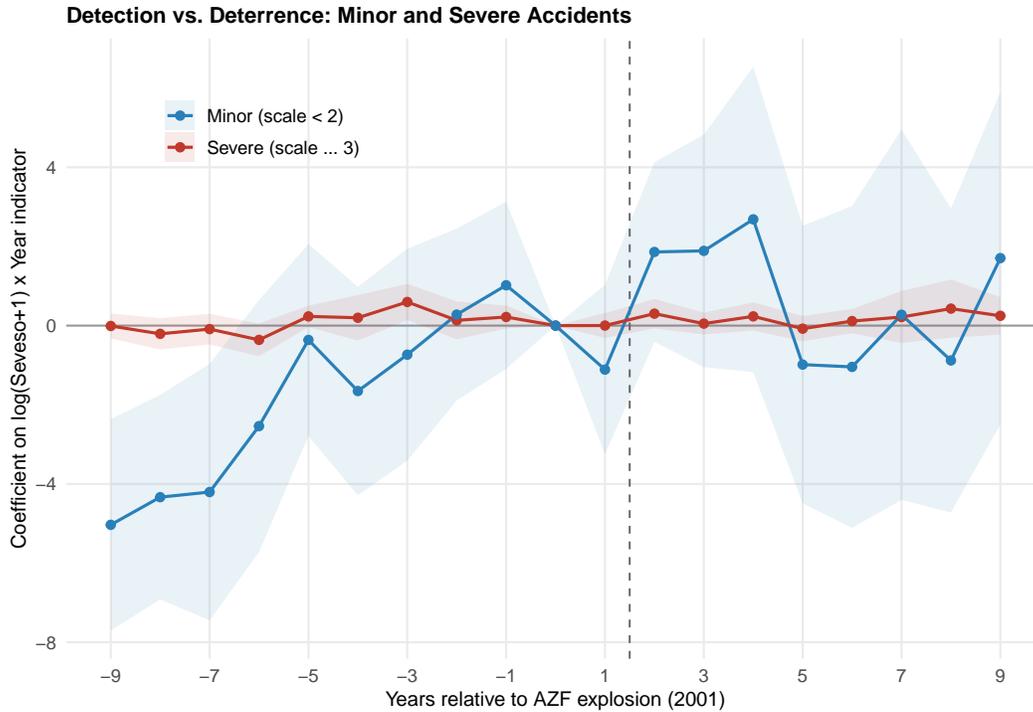


Figure 3: Detection vs. Deterrence: Minor and Severe Accidents

Notes: Event study coefficients for minor accidents (max severity < 2, blue) and severe accidents (max severity ≥ 3 , red) plotted on the same axes. The divergence after 2003 suggests that the increase in total reported accidents is driven by detection of minor incidents rather than by changes in underlying safety.

Three interpretations are possible. First, the null on severe accidents may reflect limited statistical power—severe events are rare (mean = 1.05 per department-year), and the MDE is approximately 0.26. However, the positive point estimate (0.12) is inconsistent with the power explanation: if large deterrence were present but undetected, we would expect a negative coefficient. Second, deterrence for severe accidents could exist but be exactly offset by increased detection—a knife-edge coincidence that seems unlikely. Third, and most consistent with the data, the reporting expansion that drives the minor-incident increase simply does not extend to severe outcomes, because severe events are already captured regardless of reporting infrastructure maturity. This interpretation does not require a causal claim about the Loi 2003; it is a statement about the detection-elasticity structure of the ARIA database.

6.4 Robustness

Table 3: Robustness Checks

Specification	Coefficient	SE	<i>t</i> -stat	<i>p</i> -value	<i>N</i>
<i>Panel A: Total Accidents</i>					
Baseline	2.970	1.390	2.136	0.033	1,843
Excluding Toulouse (Dept 31)	3.006	1.398	2.151	0.032	1,824
Seveso departments only	3.643	1.831	1.990	0.047	1,672
Dept-specific linear trends	-0.363	1.605	-0.226	0.821	1,843
Region × Year FE	3.326	1.473	2.259	0.024	1,843
Narrow window (1998–2006)	1.823	1.268	1.437	0.151	873
Poisson (PPML) [†]	0.071	0.050	1.431	0.152	1,843
Placebo: fake 1997 treatment	3.454	0.925	3.737	<0.001	1,067
<i>Panel B: Severe Accidents</i>					
Baseline	0.125	0.134	0.929	0.353	1,843
Excluding Toulouse	0.130	0.135	0.962	0.336	1,824
Seveso departments only	0.184	0.177	1.041	0.298	1,672
Dept-specific linear trends	-0.124	0.145	-0.856	0.392	1,843
Region × Year FE	0.230	0.133	1.729	0.084	1,843
Narrow window (1998–2006)	-0.063	0.100	-0.627	0.531	873
Placebo: fake 1997 treatment	0.278	0.138	2.011	0.044	1,067
<i>Panel C: Poisson (PPML) Decomposition</i>					
Total (Poisson)	0.071	0.050	1.431	0.152	1,843
Minor (Poisson)	0.067	0.053	1.264	0.206	1,843

Notes: All specifications include department and year fixed effects with department-clustered standard errors unless noted. “Dept-specific linear trends” adds department × year interactions, absorbing linear trends within each department. “Region × Year FE” replaces year FE with region-by-year FE (18 pre-2016 regions). “Narrow window” restricts to 1998–2006. [†]Poisson coefficients are semi-elasticities: $\exp(0.071) - 1 \approx 7.4\%$, compared to the OLS-implied 19% ($= 2.97/15.71$). The PPML results are not significant, and the full PPML severity decomposition (Panel C) shows that even in the Poisson specification, the association—to the extent it exists—is concentrated in minor rather than severe incidents. The placebo test restricts the sample to 1992–2002 and places the fake treatment at 1997.

Table 3 presents robustness checks. The results reveal an important pattern: the baseline association survives some specifications but not others, and the specifications that most directly address the pre-trend concern eliminate it.

Sample restrictions. Removing department 31 (Toulouse, the AZF site) barely changes the estimate ($\hat{\beta} = 3.006$ vs. 2.970), and restricting to the 88 Seveso-hosting departments produces a larger point estimate (3.643). Leave-one-out analysis yields estimates ranging from 1.81 to 3.31 with no single department driving the result.

Department-specific linear trends. The most informative robustness check adds department-specific linear trends, which absorb the pre-existing differential trajectories documented in Section 7. The total-accident coefficient drops to -0.363 (SE = 1.605) and is statistically indistinguishable from zero. The minor-incident coefficient similarly reverses sign (-0.537). The severe-accident coefficient is also null (-0.124 , SE = 0.145). This result confirms that the baseline association is driven by differential linear trends across departments, not by a discrete post-2003 shift. Following Roth (2022), the failure of pre-trend tests combined with the sensitivity to trend controls indicates that the identifying variation conflates treatment effects with differential trends.

Region \times year fixed effects. Replacing year FE with region-by-year FE (18 pre-2016 regions) allows for flexible regional shocks while preserving within-region variation. The total-accident coefficient increases slightly to 3.326 (SE = 1.473), suggesting that regional shocks do not confound the cross-department association.

Narrow window. Restricting to 1998–2006 (four years on each side of the treatment) produces a smaller and insignificant total-accident estimate (1.823, SE = 1.268), consistent with the identification challenge: shorter windows reduce the influence of long-run trends but also reduce power.

Poisson (PPML). The count-model specification yields a coefficient of 0.071 (SE = 0.050, $p = 0.152$), not significant at conventional levels. The full PPML decomposition (Panel C) shows that the pattern—stronger association for minor than severe outcomes—is directionally preserved even under the multiplicative functional form, but statistical significance is lost. Given the count nature of the outcomes, this is an important check; it suggests the OLS-levels result is partly driven by high-count departments where the additive specification overstates proportional effects.

Randomization inference. Permuting Seveso density across departments 999 times yields a p-value of 0.000, confirming that the spatial correlation is unlikely under random assignment—but this does not address the pre-trend concern, which is about systematic correlation with reporting infrastructure, not random chance.

6.5 Long-Run Effects

A natural question is whether the association persists over longer horizons. Extending the panel through 2020 (29 years, 2,813 observations), the total-accident coefficient attenuates slightly to 2.53 (SE = 1.42), while the severe-accident coefficient increases modestly to 0.22 (SE = 0.16). The qualitative pattern persists: positive association with total reports, no significant change in severe outcomes. Given the pre-trend evidence, this persistence is equally consistent with ongoing maturation of reporting infrastructure in high-Seveso departments and with a permanent structural shift in detection capacity from the inspector expansion. The data cannot distinguish between these explanations.

7. Threats to Validity

7.1 Pre-Existing Trends

The most significant threat is the failure of the parallel trends assumption. The placebo test placing the treatment at 1997 produces a coefficient ($\hat{\beta} = 3.45$) that is *larger* than the actual post-2003 estimate (2.97), and the event study shows statistically significant pre-trend coefficients. [Figure 4](#) plots raw accident means by Seveso density group, confirming that high-density departments were on steeper trajectories throughout the 1990s.

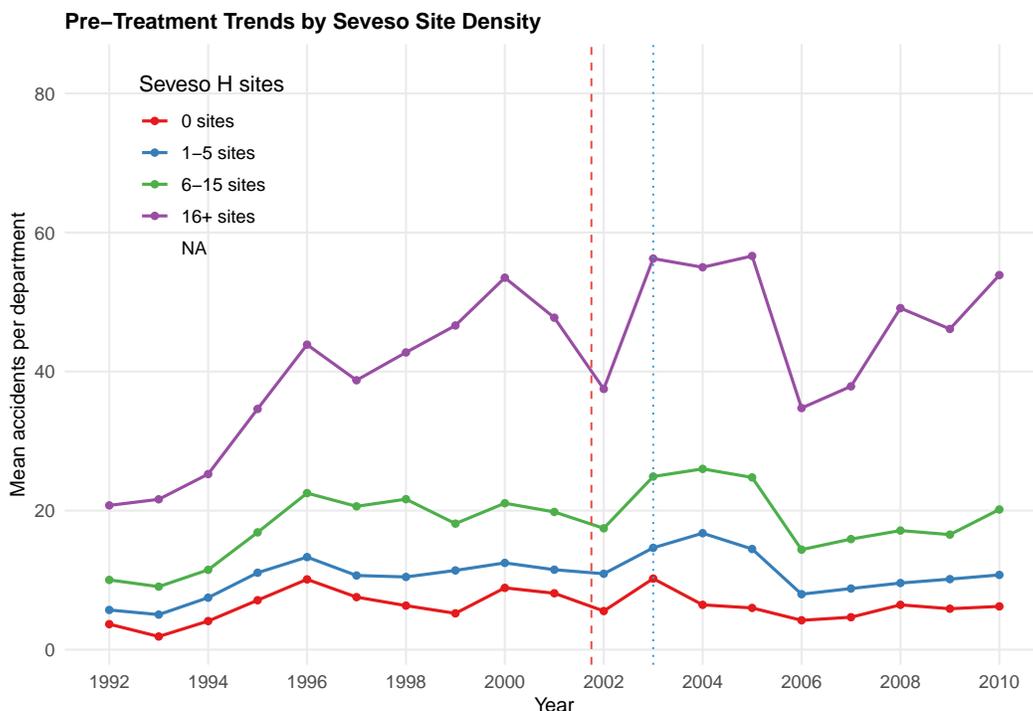


Figure 4: Pre-Treatment Trends by Seveso Site Density Group

Notes: Mean total accidents per department-year by Seveso density quartile group. Dashed red line marks the AZF explosion (September 2001); dotted blue line marks the Loi 2003 (July 2003). The divergence in levels and trends predates the policy shock.

What explains these pre-existing trends? Two factors are likely. First, the ARIA database itself was expanding throughout the 1990s as reporting infrastructure matured, and this maturation proceeded faster in departments with more classified installations—precisely the departments with more Seveso sites. Second, EU-driven regulatory harmonization (transposition of Seveso II Directive, Décret 2000-258) was already increasing reporting obligations before the AZF explosion.

The department-specific trend results in [Table 3](#) confirm the severity of this problem: once linear trends are absorbed, the total-accident coefficient reverses sign and is statistically indistinguishable from zero. This means the baseline association cannot be interpreted as a causal effect of the Loi 2003 or any discrete post-2003 policy change. As [Rambachan and Roth \(2023\)](#) emphasize, when pre-trend tests reject, the standard DiD estimator is sensitive to violations of parallel trends, and robustness to trend controls is a necessary (though not sufficient) condition for credibility.

The pre-trend violation means this paper cannot identify the causal effect of the Loi 2003. However, the detection-versus-deterrence decomposition remains informative as a

diagnostic. The key observation is not the level of the coefficient but the *pattern across severity categories*: regardless of whether the reporting increase is driven by post-2003 enforcement, longer-running infrastructure trends, or both, the consistent finding that severe accidents show no differential change while minor incidents drive all variation reveals a structural feature of enforcement-generated administrative data. This pattern holds across all specifications—baseline, department trends, region-by-year FE, and Poisson—and reflects the different detection elasticities of minor and severe outcomes.

7.2 Measurement of Treatment Intensity

The treatment variable—current Seveso site counts—is measured with error to the extent that the 2026 registry differs from the 2001 distribution. Plant closures during the sample period would attenuate the treatment variable for some departments. However, Seveso classification is persistent (driven by stored substance quantities), and the cross-sectional variation is dominated by the historical geography of heavy industry rather than by marginal entry and exit.

7.3 Alternative Mechanisms

The differential growth in reported minor incidents could reflect multiple reporting and detection channels, not solely inspector activity. These include: (1) the Loi 2003 strengthened legal reporting obligations, potentially increasing firm-initiated reports; (2) heightened media salience after the AZF explosion may have changed reporting norms, particularly in high-risk departments; (3) the PPRT process improved informational infrastructure around classified sites; (4) the maturation of the ARIA database itself, with BARPI investing more in data capture for departments with more classified installations; and (5) EU-level harmonization of reporting standards under the Seveso II directive.

All of these channels share the property of increasing recorded incidents without necessarily improving actual safety, and all produce the observed pattern of more minor reports with unchanged severe outcomes. The paper’s diagnostic does not require distinguishing among them. The severity decomposition’s value lies in separating *any* reporting/detection-driven increase from a deterrence-driven decrease in actual incidents, regardless of which specific channel inflates minor-incident recording. This makes the diagnostic portable: it can be applied in any regulatory setting where enforcement-generated data are used to evaluate policy, without needing to identify the precise reporting mechanism.

7.4 Statistical Power for Severe Accidents

An important concern is whether the null on severe accidents reflects a true zero or insufficient statistical power. With a pre-treatment mean of 1.05 and standard deviation of 1.57, severe accidents are rare events. The minimum detectable effect (MDE) at 80% power is approximately ± 0.26 severe accidents per department-year—roughly a 25% change from the mean. I therefore cannot rule out moderate deterrence effects below this threshold, and the language “precisely estimated null” would overstate the power of the test. For fatal accidents (mean = 0.17, SD = 0.45), statistical power is even more limited. The point estimates for both severe (0.12) and fatal (0.014) outcomes are positive, not negative, which is inconsistent with large deterrence effects masking detection, but modest deterrence remains possible.

8. Conclusion

This paper documents a measurement problem that plagues regulatory evaluation: enforcement-generated databases conflate detection and deterrence. Using France’s ARIA industrial accident database and variation in Seveso site density across departments, I show that departments with greater hazardous-industry concentration experienced differentially larger increases in reported accidents—but the increase is concentrated entirely in minor incidents. Severe and fatal outcomes, which are detection-inelastic, show no differential change across any specification.

This pattern is robust to sample restrictions, Poisson count models, and region-by-year fixed effects. But it is not robust to department-specific linear trends, which eliminate the total-accident association entirely. The baseline results reflect pre-existing differential trends in reporting infrastructure, not a clean post-2003 causal effect. The paper therefore cannot identify the effect of the Loi 2003 on industrial safety. What it can identify is a structural feature of the data: the divergence between minor and severe outcomes reveals the detection channel operating in the ARIA database, regardless of whether that detection is driven by the Loi 2003, by longer-running ARIA database maturation, or by EU-level reporting harmonization.

The severity decomposition I propose is portable. In any regulatory setting where enforcement generates the data used for evaluation, researchers and policymakers can compare detection-elastic outcomes (minor incidents, petty violations) with detection-inelastic outcomes (fatalities, major disasters). Only the latter provide a clean test of whether enforcement is actually improving safety. This logic applies to workplace safety regulation (Gray, 1993; Gray and Scholz, 1998; Johnson, 2020), environmental enforcement (Shimshack and Ward, 2007; Bennear, 2007), tax administration (Alm, 2012), anti-corruption enforcement (Fisman

and Golden, 2017), and policing (Levitt, 1997; Chalfin and McCrary, 2018). In each domain, rising administrative counts may signal more detection rather than more danger, and the severity gradient provides a ready-made diagnostic.

The null on severe outcomes in the French case should be interpreted cautiously. Statistical power is limited: the minimum detectable effect for severe accidents is approximately 25% of the baseline mean. Moderate deterrence effects cannot be ruled out. But the point estimates are positive, not negative, and the confidence interval for severe accidents rules out large reductions. This is consistent with two possibilities. First, firms facing Seveso-level hazards may already have been near the safety frontier for catastrophic risk, with the marginal deterrence from additional routine inspections being negligible (Polinsky and Shavell, 2000). Second, the specific form of enforcement expansion—more procedural inspections of documentation and compliance records—may not address the root causes of severe accidents, which often involve low-probability failures that routine monitoring is poorly suited to detect *ex ante*.

The ARIA database, with over 63,000 classified industrial accidents spanning three decades, represents an underexploited resource for the economics of regulation. Future work could pursue three strategies. First, plant-level linkages to the ICPE registry would allow within-facility estimation, avoiding the aggregation inherent in department-level analysis. Second, department-level historical data on inspector counts and inspection frequencies—currently unavailable in public sources—would allow direct estimation of the enforcement first stage. Third, cross-country comparisons using severity-coded databases in other EU member states could test whether the detection-deterrence decomposition generalizes beyond France.

The broader lesson is simple. When enforcement activity generates the outcome data used for evaluation, raw counts cannot be read at face value. The severity decomposition provides a practical diagnostic: it does not solve the identification problem, but it reveals when the problem is operative. Any serious evaluation of regulatory effectiveness should report results separately for detection-elastic and detection-inelastic outcomes. Failure to do so risks confusing better surveillance with worse safety.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

Project Repository: <https://github.com/SocialCatalystLab/ape-papers>

Contributors: @olafdrw

First Contributor: <https://github.com/olafdrw>

References

- Alm, James**, “Measuring, Explaining, and Controlling Tax Evasion: Lessons from Theory, Experiments, and Field Studies,” *International Tax and Public Finance*, 2012, 19 (1), 54–77.
- Becker, Gary S.**, “Crime and Punishment: An Economic Approach,” *Journal of Political Economy*, 1968, 76 (2), 169–217.
- Benbear, Lori S.**, “Evaluating Voluntary Discovery and Disclosure of Environmental Violations: The Role of Inspections and Enforcement,” *Journal of Environmental Economics and Management*, 2007, 54 (1), 22–41.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan**, “How Much Should We Trust Differences-in-Differences Estimates?,” *Quarterly Journal of Economics*, 2004, 119 (1), 249–275.
- Borusyak, Kirill, Peter Hull, and Xavier Jaravel**, “Quasi-Experimental Shift-Share Research Designs,” *Review of Economic Studies*, 2022, 89 (1), 181–213.
- Chalfin, Aaron and Justin McCrary**, “The Effect of Police on Crime: New Evidence from U.S. Cities, 1960–2010,” *Review of Economics and Statistics*, 2018, 100 (1), 115–128.
- Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan**, “Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India,” *Quarterly Journal of Economics*, 2013, 128 (4), 1499–1545.
- , –, –, and –, “The Value of Regulatory Discretion: Estimates from Environmental Inspections in India,” *Econometrica*, 2018, 86 (6), 2123–2160.
- Fisman, Raymond and Miriam A. Golden**, “Corruption and Government Regulation,” *Annual Review of Political Science*, 2017, 20, 45–66.
- Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift**, “Bartik Instruments: What, When, Why, and How,” *American Economic Review*, 2020, 110 (8), 2586–2624.
- Gray, Wayne B.**, “The Impact of OSHA and EPA Regulation on Productivity,” *American Economic Review*, 1993, 83 (5), 1215–1234.
- and **John T. Scholz**, “Is It Really Cheaper to “Regulate” Through Information? The Case of the OSHA 200,” *Journal of Regulatory Economics*, 1998, 14 (1), 1–21.

- Greenstone, Michael**, “The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures,” *Journal of Political Economy*, 2002, *110* (6), 1175–1219.
- Hanna, Rema N. and Paulina Oliva**, “The Impact of Inspections on Plant-Level Air Emissions,” *B.E. Journal of Economic Analysis and Policy*, 2014, *14* (4), 1297–1326.
- Johnson, Matthew S.**, “Regulation by Shaming: Deterrence Effects of Publicizing Violations of Workplace Safety and Health Laws,” *American Economic Review*, 2020, *110* (6), 1866–1904.
- Levitt, Steven D.**, “Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime,” *American Economic Review*, 1997, *87* (3), 270–290.
- Polinsky, A. Mitchell and Steven Shavell**, “The Economic Theory of Public Enforcement of Law,” *Journal of Economic Literature*, 2000, *38* (1), 45–76.
- Rambachan, Ashesh and Jonathan Roth**, “A More Credible Approach to Parallel Trends,” *Review of Economic Studies*, 2023, *90* (5), 2555–2591.
- Roth, Jonathan**, “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends,” *American Economic Review: Insights*, 2022, *4* (3), 305–322.
- Shimshack, Jay P. and Michael B. Ward**, “Enforcement and Over-Compliance,” *Journal of Environmental Economics and Management*, 2007, *55* (1), 90–105.
- Stigler, George J.**, “The Optimum Enforcement of Laws,” *Journal of Political Economy*, 1970, *78* (3), 526–536.

A. Data Appendix

A.1 ARIA Database Construction

The ARIA database was downloaded from data.gouv.fr as a semicolon-delimited CSV file with Latin-1 encoding. The file contains metadata headers in the first seven rows; data begin at row 8. Key fields include: title, date, ARIA number, NAF code, country, department, commune, accident type, event type, materials, equipment, CLP hazard class, root causes, immediate causes, consequences, severity scale, URL, and full-text description.

Sample restrictions:

1. Retain records with Pays = “FRANCE” (drop foreign incidents)
2. Drop overseas departments (codes beginning with 97)
3. Restrict to years 1992–2010 for main analysis (1992–2020 for extended panel)

After these restrictions, the analysis sample contains 56,034 records across 97 metropolitan department codes. The panel includes 97 rather than 96 departments because the union of ARIA and Georisques department codes produces one additional unit; departments with zero Seveso sites and zero ARIA records in some years are filled with zeros in the balanced panel.

Severity parsing: The `Echelle` field contains severity ratings in format “XH, YEn, ZEc, WM” where each letter-number pair represents a dimension (Human, Environmental, Economic, Material) scored 0–6. I extract each component using regular expression matching and compute the maximum across dimensions.

A.2 Seveso Site Data

Seveso seuil haut site counts were obtained from the Georisques ICPE API (https://georisques.gouv.fr/api/v1/installations_classees). For each metropolitan department (including Corsica 2A and 2B), I retrieved all classified installations with page size 1,000 and filtered for `statutSeveso = "Seveso seuil haut"`. The resulting dataset contains 737 Seveso seuil haut installations across 88 departments.

A.3 Seveso Site Distribution

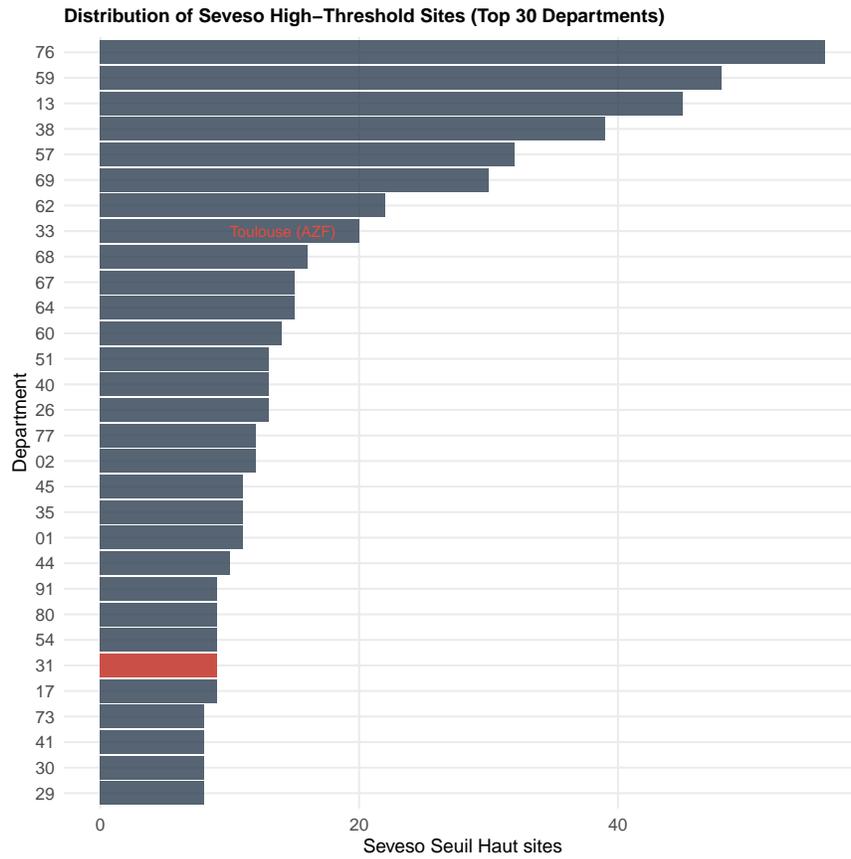


Figure 5: Distribution of Seveso Seuil Haut Sites Across Top 30 Departments

Notes: Department 31 (Toulouse, the AZF site) is highlighted in red. The distribution reflects the historical geography of French heavy industry.

B. Identification Appendix

B.1 Aggregate Time Series



Figure 6: Industrial Accident Reports in France, 1992–2010

Notes: Total annual ARIA-recorded industrial accidents across all metropolitan departments. Dashed red line marks the AZF explosion (September 2001); dotted blue line marks the Loi 2003 (July 2003). The series shows a trend increase throughout the period, with a visible acceleration after 2002.

B.2 Pre-Trend Test Details

Joint F-test for the null hypothesis that all pre-treatment event study coefficients $(\beta_{-9}, \dots, \beta_{-1})$ are jointly zero:

- Total accidents: $F(9, 1710) = 3.62, p < 0.001$
- Severe accidents: $F(9, 1710) = 2.11, p = 0.026$

Both tests reject, indicating that parallel pre-trends do not hold. This is a material limitation of the difference-in-differences design for establishing causal effects of the Loi 2003 specifically.

C. Robustness Appendix

C.1 Randomization Inference

To assess whether the observed treatment-outcome association could arise by chance under random assignment of Seveso density to departments, I conduct randomization inference with 999 permutations. In each permutation, I randomly reassign Seveso log-counts across departments (preserving the temporal structure) and re-estimate the main specification.

Table 4: Randomization Inference Results

Outcome	Observed $\hat{\beta}$	Permutations	RI p -value
Total accidents	2.970	999	0.000

Notes: Two-sided randomization inference. Each permutation randomly reassigns $\log(\text{Seveso } H + 1)$ across departments, preserving the temporal panel structure, and re-estimates the main DiD specification with department and year fixed effects. The observed coefficient exceeds all 999 permuted coefficients.

This confirms that the spatial correlation between Seveso density and post-2003 reporting changes is highly unlikely under random assignment, though it does not address the pre-trend concern.

C.2 Leave-One-Out Analysis

To verify that no single department drives the main result, I re-estimate the total-accidents specification 97 times, each time excluding one department.

Table 5: Leave-One-Out Summary

Statistic	Value
Full-sample $\hat{\beta}$	2.970
Minimum LOO $\hat{\beta}$	1.81
Maximum LOO $\hat{\beta}$	3.31
Mean LOO $\hat{\beta}$	2.97
LOO estimates > 0	97/97

Notes: Each row excludes one of 97 departments and re-estimates the main specification. The narrow range confirms no single department drives the result. See [Figure 7](#) for the full distribution.

C.3 Leave-One-Out Sensitivity

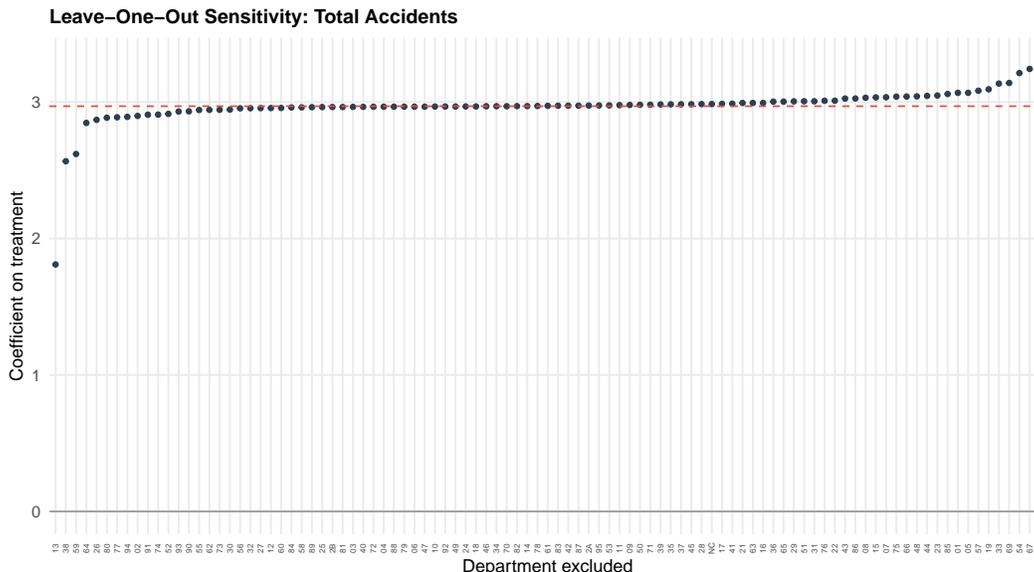


Figure 7: Leave-One-Out Sensitivity: Total Accidents

Notes: Each point represents the estimated coefficient on the treatment variable when one department is excluded. The dashed red line marks the full-sample estimate (2.97). No single department drives the result.

C.4 Extended Panel (1992–2020)

Extending the panel through 2020 yields slightly attenuated estimates, as shown in [Table 6](#). The persistence of the total-accident effect over nearly two decades suggests a structural rather than transitory change in reporting behavior.

Table 6: Long-Run Effects (1992–2020)

Outcome	Coefficient	SE	N
Total accidents	2.528	1.417	2,813
Severe accidents	0.225	0.160	2,813

Notes: Same specification as [Table 2](#) (department and year FE, department-clustered SEs), extended to 1992–2020 (97 departments \times 29 years). Treatment = $\log(\text{Seveso H sites} + 1) \times \text{Post-2003}$.

D. Heterogeneity Appendix

The binary treatment specification (high vs. low Seveso density, defined at the within-Seveso-department median) yields $\hat{\beta} = 3.51$ (SE = 1.75) for total accidents and $\hat{\beta} = 0.14$ (SE =

0.16) for severe accidents, consistent with the continuous treatment results.

E. Standardized Effect Sizes

Table 7: Standardized Effect Sizes for Main Outcomes

Outcome	Specification	$\hat{\beta}$	SD(X)	SD(Y)	SDE	Classification
Total accidents	DiD, Table 2 Col. 1	2.970	0.94	15.00	0.187	Large positive
Severe accidents	DiD, Table 2 Col. 2	0.125	0.94	1.57	0.075	Small positive
Fatal accidents	DiD, Table 2 Col. 3	0.014	0.94	0.45	0.030	Null
Minor incidents	DiD, Table 2 Col. 4	2.386	0.94	12.95	0.174	Large positive

Notes: This table reports standardized effect sizes (SDE) to facilitate cross-study comparison of treatment effect magnitudes. Treatment is continuous: $\log(\text{Seveso H sites} + 1)$, so $\text{SDE} = \hat{\beta} \times \text{SD}(X)/\text{SD}(Y)$, giving the effect of a one-standard-deviation change in treatment intensity measured in standard deviations of the outcome. $\text{SD}(X) = 0.94$ is the unconditional SD of $\log(\text{Seveso H} + 1)$ across all department-years. $\text{SD}(Y)$ values are from Table 1 (full sample).

Research question: Does disaster-induced regulatory expansion (the 2003 Loi on industrial risk prevention, which doubled ICPE inspectors) affect industrial accident outcomes in French departments? **Treatment:** Continuous; $\log(\text{Seveso Seuil Haut sites} + 1)$ per department, interacted with post-2003 indicator. $\text{SD}(X) = 0.94$ represents the cross-departmental variation in exposure to the enforcement expansion. **Data:** ARIA industrial accident database (data.gouv.fr), 1992–2010, department-year panel. $N = 1,843$ (97 departments \times 19 years). **Method:** Continuous treatment DiD with department and year FEs, department-clustered SEs.

Sample: Metropolitan France, excluding overseas departments.

Classification thresholds: large negative (< -0.10), small negative (-0.10 to -0.05), null (-0.05 to 0.05), small positive (0.05 to 0.10), large positive (> 0.10). The large positive SDE on total/minor accidents and the null/small SDE on severe/fatal accidents quantify the detection-deterrence asymmetry.