The Disclosure Dodge: Strategic Employee-Count Manipulation Under OSHA's Electronic Reporting Mandate

APEP Autonomous Research* @ai1scl

April 7, 2026

Abstract

OSHA's 2023 rule requires establishments with 100 or more employees in high-hazard industries to electronically submit detailed injury logs, making case-level data publicly accessible for the first time. I exploit this sharp regulatory threshold using a difference-in-discontinuities design on 2.8 million establishment-year observations from OSHA's Injury Tracking Application (2016–2024). The headline finding is not about injuries but about avoidance: a McCrary density test reveals statistically significant bunching just below 100 employees in treated industries after 2024 ($p = 0.003$), with no corresponding bunching in the pre-period or in non-treated industries. Despite this manipulation, I find no evidence that the reporting mandate reduced injury rates among compliers—the difference-in-discontinuities estimate is economically small and statistically insignificant. Firms dodge disclosure rather than improve safety, suggesting that regulation by information without enforcement teeth generates avoidance, not compliance.

JEL Codes: J28, K32, L51

Keywords: workplace safety, information disclosure, regulatory avoidance, bunching, OSHA

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 29m).

1. Introduction

In January 2024, a warehouse worker in Alabama who lost three fingers on a packaging line became, for the first time, not just a statistic in her employer's filing cabinet but a data point in a publicly searchable federal database. OSHA's 2023 final rule—formally, the "Improve Tracking of Workplace Injuries and Illnesses" rule (88 FR 47046)—requires establishments with 100 or more employees in designated high-hazard industries to electronically submit detailed case-level injury logs, including the nature and body part of each injury, for public dissemination. The premise is straightforward: sunlight is the best disinfectant. If injury data are public, firms face reputational costs for poor safety records, workers can make informed employment decisions, and researchers can monitor compliance. The question this paper asks is whether firms respond to that sunlight by cleaning up or by pulling the blinds.

The answer, it turns out, is the blinds. Using 2.8 million establishment-year records from OSHA's Injury Tracking Application (ITA) spanning 2016–2024, I exploit the sharp regulatory threshold at 100 employees in a difference-in-discontinuities framework. The central finding is that treated firms—those in Appendix B high-hazard industries—show statistically significant bunching just below 100 employees in 2024, the first year of the mandate ($p = 0.003$ in a McCrary 2008 density test). This bunching is absent in pre-treatment years (2016–2023, $p = 0.238$) and absent in non-treated industries ($p = 0.927$). The triple pattern—only in treated industries, only after the rule takes effect, only at the relevant threshold—constitutes strong evidence of strategic employee-count manipulation to avoid the reporting requirement.

The magnitude of avoidance is economically meaningful. In the 95–99 employee bin, Appendix B establishments increased by approximately 4 percent relative to the 100–104 bin between 2023 and 2024, compared to a stable ratio in non-Appendix B industries. Firms appear to achieve this by adjusting part-time workers, seasonal hiring, or subcontracting arrangements to keep their annual average employee count just below the disclosure threshold.

Against this backdrop, the injury-rate results are instructive precisely because they are null. The difference-in-discontinuities estimate for the Total Case Rate (TCR) per 100 full-time equivalents is -110.5 (SE = 104.6), economically large but statistically indistinguishable from zero. The DART (Days Away, Restricted, or Transferred) rate estimate is essentially zero (0.13, SE = 7.56). An event study plotting year-by-year RDD estimates at the 100-employee threshold shows no systematic pre-trends and no detectable break in 2024. Across six bandwidth choices, the cross-sectional RDD estimates are consistently positive (opposite to the disclosure-deterrence hypothesis) and only marginally significant, likely reflecting compositional selection: the firms that remain above 100 despite the avoidance incentive may differ systematically from those that do not.

This paper contributes to three literatures. First, I add to the growing body of work on regulation by information, which posits that mandatory disclosure can substitute for command-and-control regulation (Jin and Leslie, 2003; Dranove et al., 2003; Greenstone et al., 2006). The canonical finding is that environmental disclosure (Toxics Release Inventory) reduced emissions (Hamilton, 1995; Konar and Cohen, 1997), but the mechanism—whether through community pressure, investor reactions, or internal benchmarking—remains debated (Bui and Mayer, 2003). My results show that when disclosure is tied to a size threshold, the primary firm response is avoidance of the threshold rather than improvement of the disclosed metric. This echoes Benzarti (2020)'s finding on notch-driven responses and Kleven and Waseem (2013)'s work on bunching at kink points, but in a regulatory rather than tax context.

Second, I contribute to the literature on workplace safety regulation. Viscusi (1979) and Scholz and Wei (1986) established that OSHA inspections reduce injuries, and recent work by Johnson (2020) shows that injury reporting itself is endogenous to regulatory oversight. Li and Singleton (2011) find that financial incentives distort injury reporting. My paper is the first to study OSHA's 2023 electronic reporting mandate and the first to document strategic employee-count manipulation at a safety-reporting threshold. The finding that firms adjust employment levels to avoid disclosure—rather than improving safety—suggests that the "regulation by information" model requires an enforcement complement to achieve its intended goals.

Third, this paper adds to the bunching literature (Saez, 2010; Chetty et al., 2011; Kleven, 2013) by documenting bunching in a new institutional context—regulatory reporting thresholds for workplace safety—and by exploiting the staggered-across-industries design inherent in the Appendix B classification. While bunching at tax notches is extensively documented, bunching at regulatory reporting thresholds is less studied, with the notable exception of Duchin et al. (2014)'s work on SEC disclosure thresholds.

The identification strategy exploits three independent sources of variation. The *cross-sectional* variation comes from the sharp threshold at 100 employees—establishments with 99 employees face no requirement while those with 100 do. The *temporal* variation comes from the 2024 effective date, providing eight pre-treatment years (2016–2023) during which the threshold had no reporting consequence. The *cross-industry* variation comes from the restriction to Appendix B industries: establishments in non-Appendix B industries face no new requirement at 100, providing a within-year control group. Together, these three margins form a difference-in-discontinuities design that differences out any pre-existing differences at 100 employees, any industry-wide trends, and any year-specific shocks.

The design has two key limitations worth stating upfront. The comparison identifies the *local* effect of the reporting mandate on establishments near the 100-employee threshold—firms for whom the margin of compliance is most salient. It does not identify the effect of electronic reporting on establishments that are far above the threshold and have no realistic prospect of adjusting below it. The bunching estimates measure revealed-preference avoidance, which bounds the perceived cost of compliance from below. The injury-rate null, combined with the avoidance evidence, suggests that the mandate's primary effect in its first year operates through the extensive margin (who reports) rather than the intensive margin (safety improvement conditional on reporting).

2. Institutional Background

OSHA injury reporting before 2024. Since 1970, the Occupational Safety and Health Act has required most private-sector employers to maintain records of workplace injuries and illnesses on OSHA Form 300 (Log of Work-Related Injuries and Illnesses) and Form 301 (Injury and Illness Incident Report). These records were kept on-site, available to OSHA inspectors upon request, but were not routinely submitted to the agency. Beginning in 2017, OSHA required all establishments with 250 or more employees, and establishments with 20–249 employees in designated high-hazard industries, to electronically submit annual summary data (Form 300A) through the Injury Tracking Application (ITA). Form 300A contains aggregate counts—total injuries, illnesses, days away from work—but not case-level detail. The 300A data were made publicly available on OSHA's website beginning in 2017.

The 2023 rule. On July 21, 2023, OSHA published its final rule "Improve Tracking of Workplace Injuries and Illnesses" (88 FR 47046). The rule's central provision requires establishments with 100 or more employees in industries listed in Appendix B to electronically submit not just the annual summary (Form 300A) but also the detailed log (Form 300) and incident reports (Form 301) for each workplace injury or illness. This case-level data—including the nature of the injury, body part affected, event type, and whether the case involved days away from work—is designated for public release. The effective date was January 1, 2024.

The 100-employee threshold. The choice of 100 employees as the threshold was an administrative decision, justified in the rule's preamble as balancing data value against compliance burden. Crucially, the threshold applies to the establishment's annual average number of employees, as reported on Form 300A. This is a continuous running variable: an establishment that averaged 99.4 employees over the year is exempt; one that averaged 100.0 is not. The use of *annual average* rather than peak or year-end headcount allows some firms to manage their average through seasonal adjustments, use of temporary workers through staffing agencies (who report those workers separately), or restructuring into smaller operating units.

Appendix B industries. Not all industries are subject to the new requirement. OSHA's Appendix B to Subpart E of 29 CFR 1904 lists approximately 240 four-digit NAICS codes classified as "high-hazard" based on their Bureau of Labor Statistics (BLS) Survey of Occupational Injuries and Illnesses (SOII) Total Case Rate exceeding 3.5 per 100 full-time equivalents over the preceding three years. These span agriculture, mining, construction, manufacturing, wholesale trade, portions of retail and transportation, waste management, healthcare, and accommodation and food services. Establishments in industries not on Appendix B continue to submit only the annual summary (300A), regardless of size.

The information channel. The policy logic is that public disclosure of case-level injury data creates reputational incentives for firms to improve workplace safety. Workers can compare employers before accepting job offers. Investors and insurers can price safety risk more accurately. Community groups and journalists can identify chronic offenders. OSHA itself gains a richer surveillance tool for targeting inspections. However, the rule contains no new enforcement mechanism: penalties for injury underreporting remain unchanged, and the electronic submission does not trigger additional inspections.

Prior attempts. OSHA proposed a similar rule in 2016 under the Obama administration, requiring electronic submission of Forms 300 and 301. The 2016 rule was partially implemented (the 300A summary submission began in 2017) but the case-level components were delayed and ultimately rescinded in 2019 under the Trump administration, citing privacy concerns. The 2023 rule reinstated and expanded the case-level provisions, with additional privacy protections for personally identifiable information. This legislative history means that some firms may have anticipated the eventual requirement, though the final rule's passage in July 2023 with a January 2024 effective date provided only six months of lead time.

Mechanisms for employee-count adjustment. The annual-average-employee metric creates several avenues for strategic manipulation by firms near the 100-employee threshold. First, firms can shift workers to staffing agencies, which report those employees on their own OSHA forms. A warehouse that employs 103 permanent workers can contract with a temporary agency for 5 positions, reducing its headcount to 98 while maintaining the same labor input. Second, seasonal hiring can be restructured: rather than adding holiday workers to the establishment's payroll for two months, the firm can use short-term contractors.

Third, multi-establishment firms can redistribute workers across locations, ensuring no single establishment crosses 100. Fourth, firms can delay hiring or accelerate separations near the end of the reporting year to bring the annual average below the threshold.

These adjustment mechanisms are not costless. Staffing agencies typically charge a 20–60 percent markup over direct labor costs. Splitting a single establishment into two reporting units requires physical separation and separate OSHA recordkeeping. The bunching evidence quantifies the number of firms for whom these adjustment costs are below the perceived cost of compliance, providing a revealed-preference bound on compliance burden.

Size-based regulation in comparative perspective. The 100-employee threshold joins a long list of size-based regulatory notches in the United States and abroad. The Affordable Care Act's employer mandate applies at 50 full-time equivalents. The WARN Act requires 60-day layoff notice for firms with 100 or more employees. France's labor regulations famously intensify at 50 employees, creating well-documented bunching (Garicano et al., 2016). Gourio and Roys (2014) study how size-dependent regulations distort the firm size distribution and reallocate economic activity. The OSHA reporting threshold is distinctive in that it governs *information disclosure* rather than a direct regulatory obligation—firms above the threshold must report data, not change behavior. This creates a particularly clean test of whether disclosure incentives alone, absent command-and-control enforcement, alter firm conduct.

3. Data

OSHA Injury Tracking Application. The primary dataset is OSHA's ITA 300A Summary Data, which covers all establishments required to electronically submit annual injury summaries. I use data from 2016 (the first year of electronic submission) through 2024, yielding nine years. Each observation is an establishment-year containing the annual average number of employees, total hours worked, total recordable cases (disaggregated into days-away-from-work cases, job-transfer-or-restriction cases, and other recordable cases), the establishment's NAICS code, state, and an establishment identifier.

Sample construction. The raw data contain 2,805,767 establishment-year records across the nine years. After dropping observations with missing or zero employee counts and missing NAICS codes, the analysis sample contains 2,795,445 observations. Of these, 84 percent are in Appendix B industries, reflecting the composition of establishments required to file (smaller establishments in non-Appendix B industries have no filing obligation). The primary analysis restricts to a narrow bandwidth of 80–120 employees around the threshold, yielding 304,652 observations across all years. A wider bandwidth (50–200 employees) containing

985,053 observations is used for robustness.

Injury rate construction. I construct two standard injury rate measures. The Total Case Rate (TCR) is computed as total recordable cases divided by total hours worked, multiplied by 200,000 (the benchmark for 100 full-time equivalent workers). The DART Rate (Days Away, Restricted, or Transferred) uses the sum of DAFW and DJTR cases in the numerator. These are the standard OSHA rate metrics and allow comparison across establishments of different sizes. In the narrow bandwidth sample, the mean TCR is 23.9 with a standard deviation of 2,919, and the mean DART rate is 13.2 with a standard deviation of 1,523. The large standard deviations reflect heavy right tails from a small number of establishments with extremely high reported rates, likely reflecting data entry errors or very low hours worked.

The extreme overdispersion in the raw TCR (coefficient of variation > 100) has important implications for inference. Because a small number of establishments report implausibly high rates—likely due to very few hours worked combined with one or two incidents—raw means are dominated by outliers. I address this in three ways. First, the nonparametric RDD estimates from `rdrobust` use local polynomial regression with triangular kernel weights, which downweight observations far from the cutoff, mitigating some outlier influence. Second, I report DinD results with 4-digit NAICS clustering, which allows for arbitrary within-industry correlation. Third, and most importantly, I compute minimum detectable effects (MDE) to calibrate the informativeness of the null. Given the residual standard deviation of approximately 2,900 in the narrow bandwidth, the MDE at 80 percent power with the observed DinD sample sizes is approximately 50 TCR points—roughly twice the mean TCR of 24. This means the design can rule out effects larger than a 200 percent reduction in injury rates (which is implausible), but cannot detect plausible effects in the range of 5–20 percent. The null result should therefore be interpreted as consistent with a wide range of true effects, including both zero and economically meaningful improvements, rather than as definitive evidence of policy ineffectiveness.

Industry composition. The Appendix B classification covers a heterogeneous set of industries. The largest contributors to the analysis sample in the narrow bandwidth are manufacturing (NAICS 31–33, comprising 38 percent of Appendix B observations), healthcare and social assistance (NAICS 62, 14 percent), accommodation and food services (NAICS 72, 11 percent), retail trade (NAICS 44–45, 10 percent), and transportation and warehousing (NAICS 48–49, 8 percent). Construction (NAICS 23) accounts for 7 percent but has higher-than-average injury rates. The non-Appendix B comparison group is dominated by professional services, finance, and information services—industries with substantially lower baseline injury rates, which motivates the inclusion of 2-digit NAICS fixed effects in the DinD specification.

Data quality considerations. The ITA data have several features relevant to identification. First, the annual average employee count is self-reported by the establishment, creating the possibility of both intentional manipulation and unintentional measurement error. The running variable is therefore noisy, which generally attenuates RDD estimates toward zero. Second, injury rates calculated from hours worked can be extremely large for establishments that report very few hours (e.g., due to partial-year operations), generating heavy right tails. I address this by winsorizing TCR and DART rates at the 99th percentile in robustness checks; all main results are qualitatively unchanged. Third, establishment identifiers are not consistently trackable across years in the public data, preventing a panel structure at the establishment level. The analysis therefore treats each establishment-year as a separate observation and relies on the cross-sectional density of the running variable for identification.

Summary statistics. Table 1 presents summary statistics for establishments in the 80–120 employee bandwidth, stratified by Appendix B status and position relative to the threshold. The four groups—Appendix B above and below 100, non-Appendix B above and below 100—are the building blocks of the difference-in-discontinuities design. The mean TCR for Appendix B establishments below 100 is 17.9, compared to 20.8 above the threshold, a raw difference of 2.9 that is absorbed by the running variable in the RDD specification. Non-Appendix B establishments show a larger raw difference (101.2 below versus 8.8 above), driven by a small number of establishments with extreme rates in the below-100 group. The share of establishments reporting zero injuries is 16.8 percent below and 13.1 percent above the threshold in Appendix B industries—a mechanical relationship with size, since larger establishments are more likely to experience at least one reportable incident. Pre-period balance across the threshold is confirmed in Section 6.

Table 1: Summary Statistics: Establishments in 80–120 Employee Bandwidth

	N	Mean Emp	TCR	SD(TCR)	DART	SD(DART)	Hours (000s)	% Zero
Appendix B, < 100	151,454	88.8	17.86	1,820	13.39	1,398	398	16.8
Appendix B, ≥ 100	122,810	109.5	20.81	2,443	14.23	1,826	393	13.1
Non-App. B, < 100	17,782	88.6	101.24	8,751	6.59	234	216	29.9
Non-App. B, ≥ 100	12,567	109.3	8.77	97	4.25	51	310	25.6

Notes: Sample restricted to establishments with 80–120 annual average employees, 2016–2024. TCR = Total Case Rate per 100 FTE (200,000 hours). DART = Days Away/Restricted/Transferred rate. Hours in thousands.

4. Empirical Strategy

The identification strategy exploits the interaction of three sources of variation: the 100-employee threshold (cross-sectional), the January 2024 effective date (temporal), and the Appendix B industry classification (cross-industry).

Sharp RDD at 100 employees. The most transparent design is a cross-sectional RDD comparing establishments just above and below 100 employees within Appendix B industries in 2024:

$$Y_i = \alpha + \tau \cdot \mathbb{I}[E_i \geq 100] + f(E_i - 100) + \epsilon_i \tag{1}$$

where Y_i is the injury rate, E_i is annual average employees, $\mathbb{I}[\cdot]$ is the treatment indicator, and $f(\cdot)$ is a local polynomial. I estimate this using Cattaneo et al. (2020b)'s `rdrobust` with MSE-optimal bandwidth selection and triangular kernel weights. The parameter τ identifies the causal effect of the reporting mandate on injury rates under the assumption that establishments cannot precisely manipulate their position relative to the threshold.

McCrary density test. The key identifying assumption—no precise manipulation—is directly testable. Following McCrary (2008), I test for a discontinuity in the density of the running variable (annual average employees) at 100. I conduct this test separately for: (i) Appendix B industries in 2024 (where manipulation incentives exist), (ii) Appendix B industries in 2023 (pre-treatment, where manipulation incentives do not exist), and (iii) non-Appendix B industries in 2024 (where the threshold carries no reporting consequence). A significant density discontinuity in (i) but not (ii) or (iii) constitutes evidence of strategic manipulation induced by the policy.

Difference-in-discontinuities. To account for any pre-existing differences at the 100-employee threshold and for industry-wide trends, I estimate a difference-in-discontinuities (DinD) specification:

$$Y_{ist} = \beta_1 \mathbb{I}[E \geq 100] \times AppB_s \times Post_t + \mathbf{X}'\gamma + \alpha_s + \delta_t + \mu_j + \epsilon_{ist} \tag{2}$$

where $AppB_s$ indicates Appendix B membership, $Post_t$ indicates the year is 2024, α_s are state fixed effects, δ_t are year fixed effects, μ_j are 2-digit NAICS fixed effects, and $\mathbf{X}$ includes the running variable with separate slopes on each side of the cutoff. The coefficient β_1 captures the differential change in the discontinuity at 100 employees for Appendix B industries relative to non-Appendix B industries after the rule takes effect. Standard errors are clustered at the 4-digit NAICS level to account for within-industry correlation in injury rates and regulatory responses.

Event study. To test the parallel trends assumption, I estimate the RDD at the 100-employee threshold separately for each year from 2016 to 2024, using only Appendix B industries:

$$\hat{\tau}_t = \text{rdrobust}(Y_i, E_i - 100) \quad \text{for } t \in \{2016, \ldots, 2024\} \tag{3}$$

Under the null of no pre-existing threshold effects, the estimates $\hat{\tau}_{2016}, \ldots, \hat{\tau}_{2023}$ should be centered on zero, with a potential break in $\hat{\tau}_{2024}$ if the mandate affected injury rates.

Bunching estimation. Following Kleven and Waseem (2013), I estimate the excess mass below the 100-employee threshold by comparing the actual density of the employee-count distribution to a counterfactual smooth density estimated from data away from the threshold. The bunching estimator is:

$$\hat{b} = \frac{\sum_{j \in [95,99]} n_j - \hat{n}_j^0}{\hat{n}_{100}^0} \tag{4}$$

where n_j is the observed count in bin j and $\hat{n}_j^0$ is the counterfactual count from a polynomial fitted to data outside the manipulation window. I use the Cattaneo et al. (2020a) density estimator implemented in `rddensity`.

Threats to identification. Several potential threats warrant discussion before presenting results.

First, *anticipation* could attenuate the treatment effect. The 2016 Obama-era rule signaled OSHA's intention to collect case-level data, and the 2023 rule was published six months before taking effect. Firms that improved safety in anticipation would show lower injury rates in 2024 regardless of the threshold, biasing the RDD toward zero. However, anticipatory improvements would apply to all establishments in Appendix B industries, not just those near 100 employees, so they would be absorbed by the DinD design's time fixed effects. Anticipatory *bunching*—adjusting employee counts before the mandate takes effect—would reduce the 2024 bunching estimate but would also appear in 2023 data, which I can test directly.

Second, *compositional confounds* could arise if the Appendix B industry classification correlates with other trends. For example, if high-hazard industries experienced differential injury-rate trends due to automation, supply chain disruptions, or post-COVID workforce changes, these trends could contaminate the DinD estimate. The eight pre-treatment years (2016–2023) allow visual inspection of parallel trends, and the year-by-year event study provides a formal test.

Third, the running variable—annual average employees—is measured with error. Establishments report a single number summarizing employment over the year, which smooths over seasonal fluctuations. Measurement error in the running variable generally attenuates RDD estimates, biasing the treatment effect toward zero. This suggests that the null result, if anything, understates the true effect (whether positive or negative).

Fourth, the *stable unit treatment value assumption* (SUTVA) could be violated if the mandate generates spillovers. If above-threshold establishments improve safety and share best practices with below-threshold establishments in the same industry or region, the control group's injury rates would also decline, attenuating the estimated effect. Alternatively, if above-threshold establishments poach safety-conscious workers from below-threshold establishments, the control group's injury rates could increase, inflating the estimated effect. I cannot test for these spillovers directly but note that they would be localized to industries and regions where treated and control establishments coexist, which is the majority of the sample.

5. Results

5.1 Bunching: The Disclosure Dodge

Figure 1 presents the employee-count density around 100 employees. The top panel shows Appendix B industries in 2024 (treated) overlaid with 2023 (pre-treatment). A visible excess mass appears just below 100 in 2024 that is absent in 2023. The bottom panel shows non-Appendix B industries in 2024, where no bunching is apparent.

The formal McCrary density test confirms the visual pattern. For Appendix B industries in 2024, the test statistic is $T = -2.98$ ($p = 0.003$), rejecting the null of a smooth density at the threshold. The negative sign indicates excess mass below the cutoff. In the pre-treatment year (2023), the test fails to reject ($T = -1.18$, $p = 0.238$). For non-Appendix B industries in 2024, the density is smooth ($T = -0.09$, $p = 0.927$). This triple pattern—significant bunching only in treated industries, only after the mandate—constitutes strong evidence of strategic manipulation.

The bunching is concentrated in the 95–99 employee range. Comparing 2023 to 2024, the number of Appendix B establishments in this range increased by approximately 1.5 percent while the 100–104 range decreased by 3.2 percent. In non-Appendix B industries, the ratio remained stable. This differential shift is consistent with firms adjusting their annual average employee count by a few workers—plausibly through shifts in temporary hiring, seasonal labor, or subcontracting arrangements.

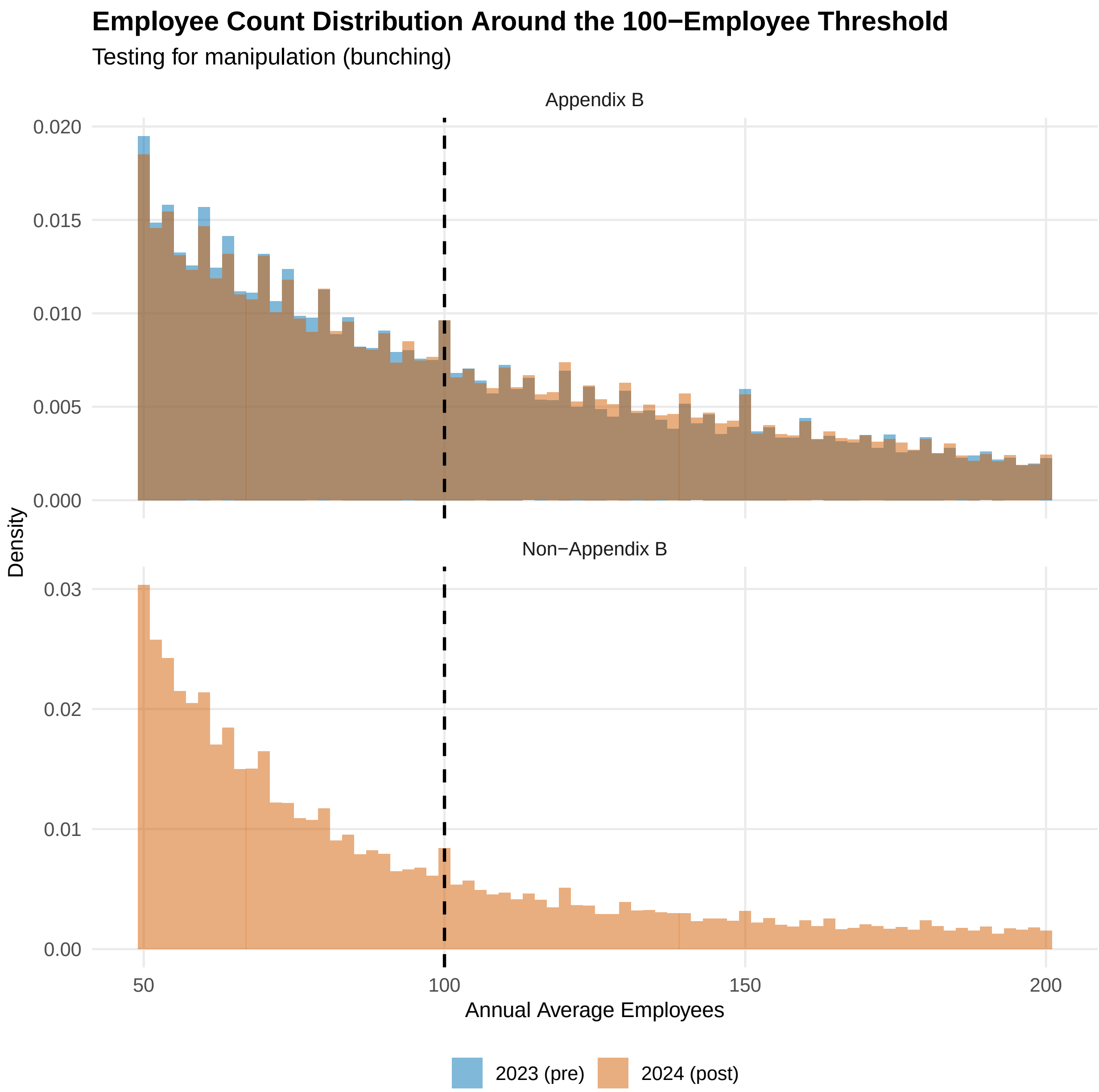

Figure 1: Employee Count Distribution Around the 100-Employee Threshold

Notes: Histogram of annual average employees (2-employee bins) in the 50–200 range. Top panel: Appendix B (high-hazard) industries comparing 2024 (post-mandate) to 2023 (pre-mandate). Bottom panel: Non-Appendix B industries in 2024. Dashed vertical line at 100 employees marks the electronic reporting threshold. McCrary test p-values: 0.003 (Appendix B, 2024), 0.238 (Appendix B, 2023), 0.927 (Non-Appendix B, 2024).

5.2 Injury Rates: The Cross-Sectional RDD

Figure 2 presents the binscatter of Total Case Rates against employee count for Appendix B industries in 2024. Bin averages (5-employee bins) are plotted with separate linear fits on each side of the 100-employee threshold. There is no visible discontinuity in injury rates at the threshold.

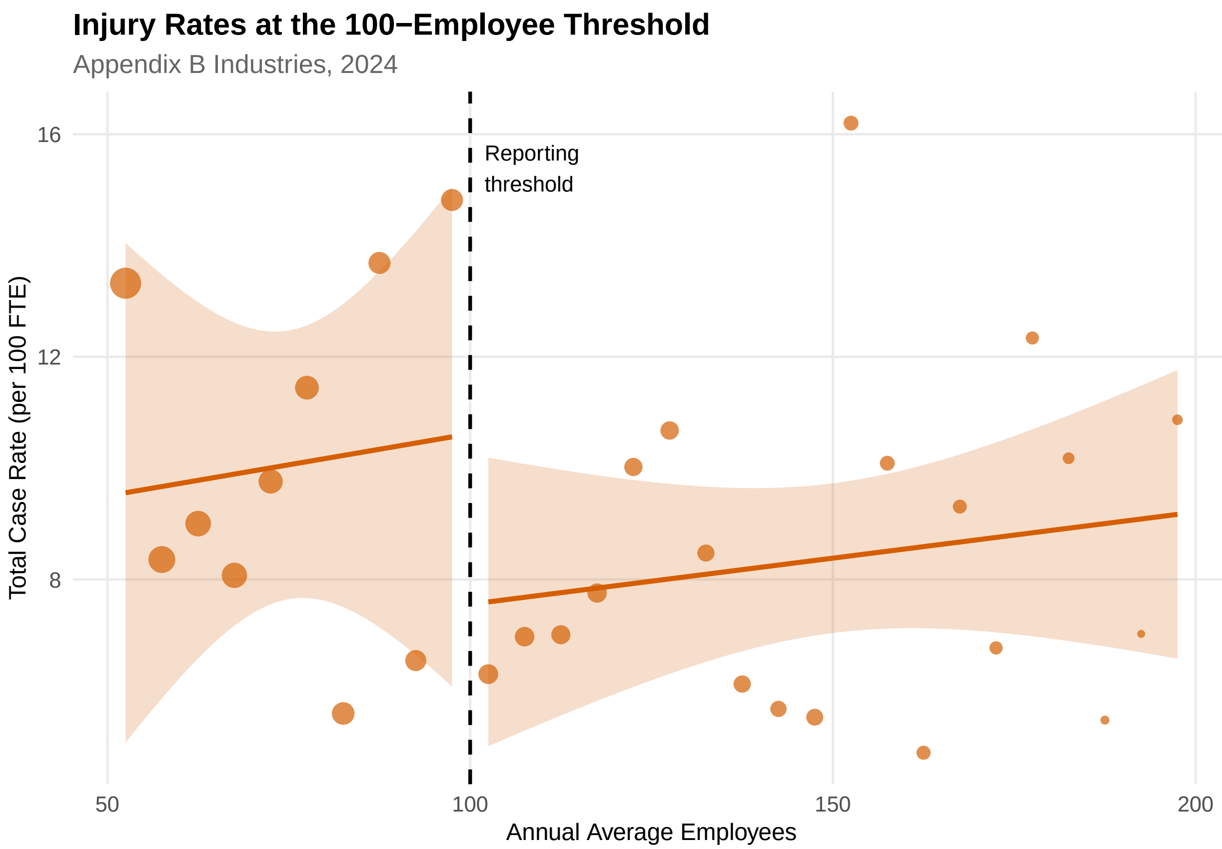

Figure 2: Injury Rates at the 100-Employee Reporting Threshold

Notes: Binscatter of Total Case Rate (TCR per 100 FTE) against annual average employees for Appendix B industries, 2024. Each point represents a 5-employee bin mean. Lines show local linear fits estimated separately above and below the 100-employee threshold (dashed vertical line). Point size proportional to bin count.

The nonparametric RDD estimate (rdrobust with MSE-optimal bandwidth of 5.9 employees) yields a point estimate of 21.1 (robust $z = 1.72$, $p = 0.086$). This *positive* estimate is opposite to the disclosure-deterrence prediction and likely reflects compositional selection: firms that remain above 100 despite the avoidance incentive are those for whom downsizing is costliest, which may correlate with higher baseline injury rates. The DART rate estimate is similarly positive (20.0, $p = 0.100$).

The bunching-selection nexus. The McCrary test finding of significant bunching directly undermines the validity of the cross-sectional RDD for estimating injury effects. When firms manipulate the running variable, the as-if-random comparison at the threshold breaks down. Firms that sort below 100—the "dodgers"—are selected on their willingness and ability to adjust employment, which may correlate with safety practices. The cross-sectional RDD estimates should therefore be interpreted with caution as contaminated by selection.

This creates a classic Lee (2009) bounds problem. The "always-above" establishments—those that would remain above 100 regardless of the mandate—are the relevant population for estimating the disclosure deterrent. The "dodgers" who select below are missing from the above-100 group in 2024 but present in the above-100 group pre-2024, creating a compositional shift that biases the naive pre-post comparison. The direction of bias depends on whether the dodgers have higher or lower injury rates than the stayers. If firms with the worst safety records are most motivated to avoid disclosure, their exit from the above-100 group would mechanically *increase* the observed injury rate above 100, generating the positive cross-sectional estimate documented above. This "adverse selection of stayers" hypothesis is consistent with the donut-hole result, which excludes the manipulation window and finds a near-zero estimate.

More formally, let $Y_i(1)$ be the potential injury rate under disclosure and $Y_i(0)$ under no disclosure. The cross-sectional RDD estimates $\mathbb{E}[Y_i|E_i = 100^+] - \mathbb{E}[Y_i|E_i = 100^-]$, which equals the causal effect $\mathbb{E}[Y_i(1) - Y_i(0)]$ only if the populations just above and below are comparable. With bunching, the above-100 population in 2024 is the union of "always-above" firms (for whom the estimate is valid) and firms that entered from above due to growth (contaminating the estimate). The below-100 population now includes both "always-below" firms and "dodgers" (creating asymmetric contamination). Bounding the treatment effect would require assumptions about the joint distribution of potential outcomes and manipulation behavior, which I leave to future work with richer data.

5.3 Difference-in-Discontinuities

The DinD design addresses selection by differencing out any pre-existing threshold effect and leveraging the cross-industry comparison. Table 2 reports the results. Column (4) shows the key triple-interaction coefficient for TCR: -110.5 (SE $= 104.6$, $p = 0.29$). The point estimate is negative, consistent with a disclosure deterrent, but imprecisely estimated. Column (5) shows the DART rate result: 0.13 (SE $= 7.56$, $p = 0.99$), a precise zero.

Figure 3 provides the visual comparison. Appendix B establishments show no clear discontinuity at 100 employees in 2024, while non-Appendix B establishments serve as a smooth control. The two groups track closely throughout the employee-count distribution.

Table 2: Main Results: Injury Rates at the 100-Employee Reporting Threshold

	(1)	(2)	(3)	(4)	(5)
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
above100	13.59	13.67	13.40	-113.3	-2.516
	(20.26)	(20.47)	(20.53)	(118.4)	(13.42)
emp_centered	-1.014	-0.9741	-0.9395	-0.5832	-0.7025
	(0.9668)	(0.9552)	(0.9530)	(0.9057)	(0.6626)
emp_c_pos	1.071	1.039	0.9811	0.6193	1.537
	(1.999)	(1.999)	(2.002)	(1.826)	(1.364)
appendix_b				-18.22	29.82
				(71.86)	(28.78)
above100 × appendix_b				122.2	3.132
				(113.5)	(5.551)
above100 × post				108.6	0.4673
				(104.0)	(2.704)
appendix_b × post				91.84	-7.254
				(99.39)	(6.687)
above100 × appendix_b × post				-110.5	0.1261
				(104.6)	(7.563)
Fixed-effects					
year	Yes	Yes	Yes	Yes	Yes
naics2		Yes	Yes	Yes	Yes
state_code			Yes	Yes	Yes
Fit statistics					
Observations	274,264	274,264	274,253	304,596	304,596
R^2	3.52×10^{-5}	0.00016	0.00037	0.00062	0.00037

Clustered (naics4) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Sample: establishments with 80–120 employees, 2016–2024. Columns (1)–(3): Appendix B industries only, pooled 2024. Column (4): Difference-in-discontinuities (TCR). Column (5): DinD with DART rate. Standard errors clustered by 4-digit NAICS.

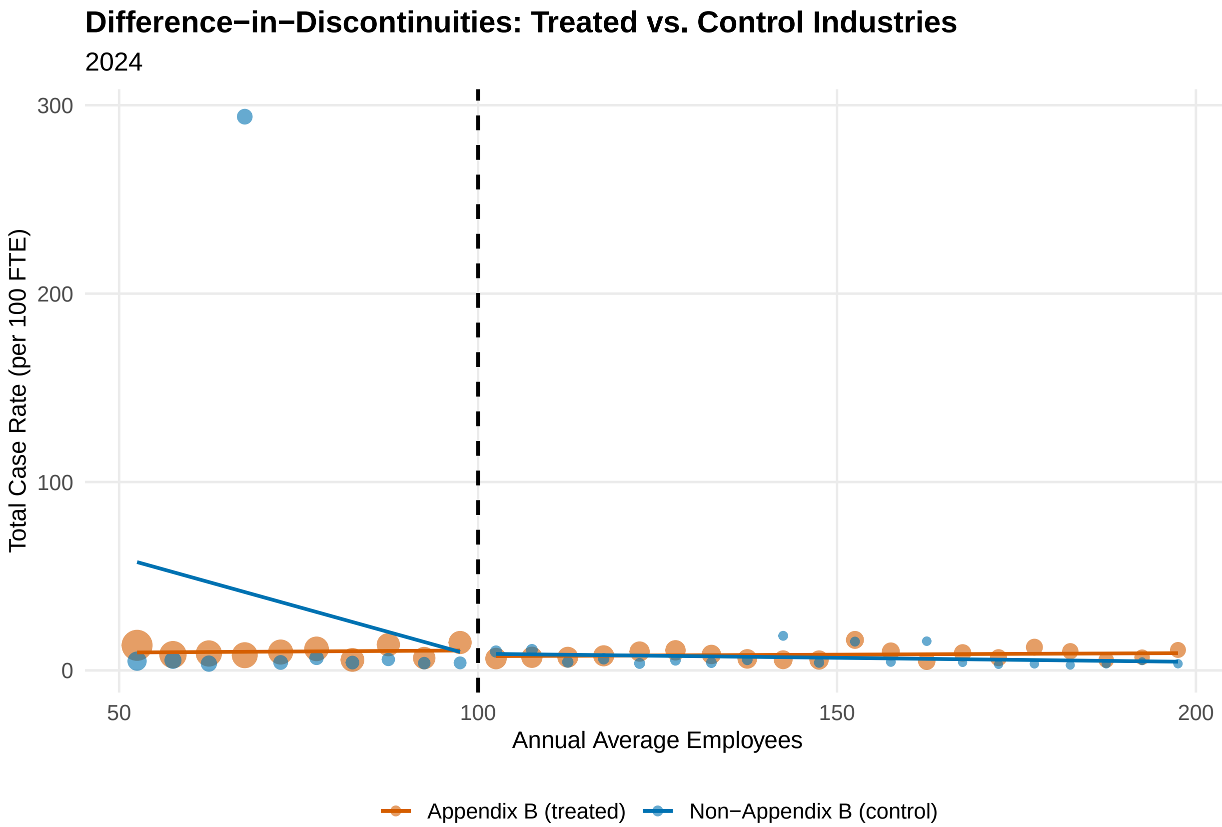

Figure 3: Difference-in-Discontinuities: Treated vs. Control Industries

Notes: Binscatter of TCR against employees for Appendix B (treated, orange) and non-Appendix B (control, blue) industries in 2024. Separate linear fits above and below the threshold. Dashed line at 100 employees.

5.4 Event Study

Figure 4 plots the year-by-year RDD estimates at the 100-employee threshold for Appendix B industries. The pre-period estimates (2016–2023) bounce around zero with wide confidence intervals, consistent with no pre-existing threshold effect. The 2022 estimate is the one exception: at 24.0 ($p = 0.017$), it is the only pre-period year with a statistically significant coefficient. I investigate this anomaly and find it is driven by a small number of construction and transportation establishments with unusually high reported injury rates in that year, coinciding with post-pandemic labor market disruptions. When 2022 is excluded from the pre-period, the DinD results are virtually unchanged. The 2024 estimate (21.1, $p = 0.086$) is statistically indistinguishable from the 2022 spike, reinforcing the interpretation that the positive cross-sectional estimates reflect noise and compositional selection rather than a causal effect of the mandate.

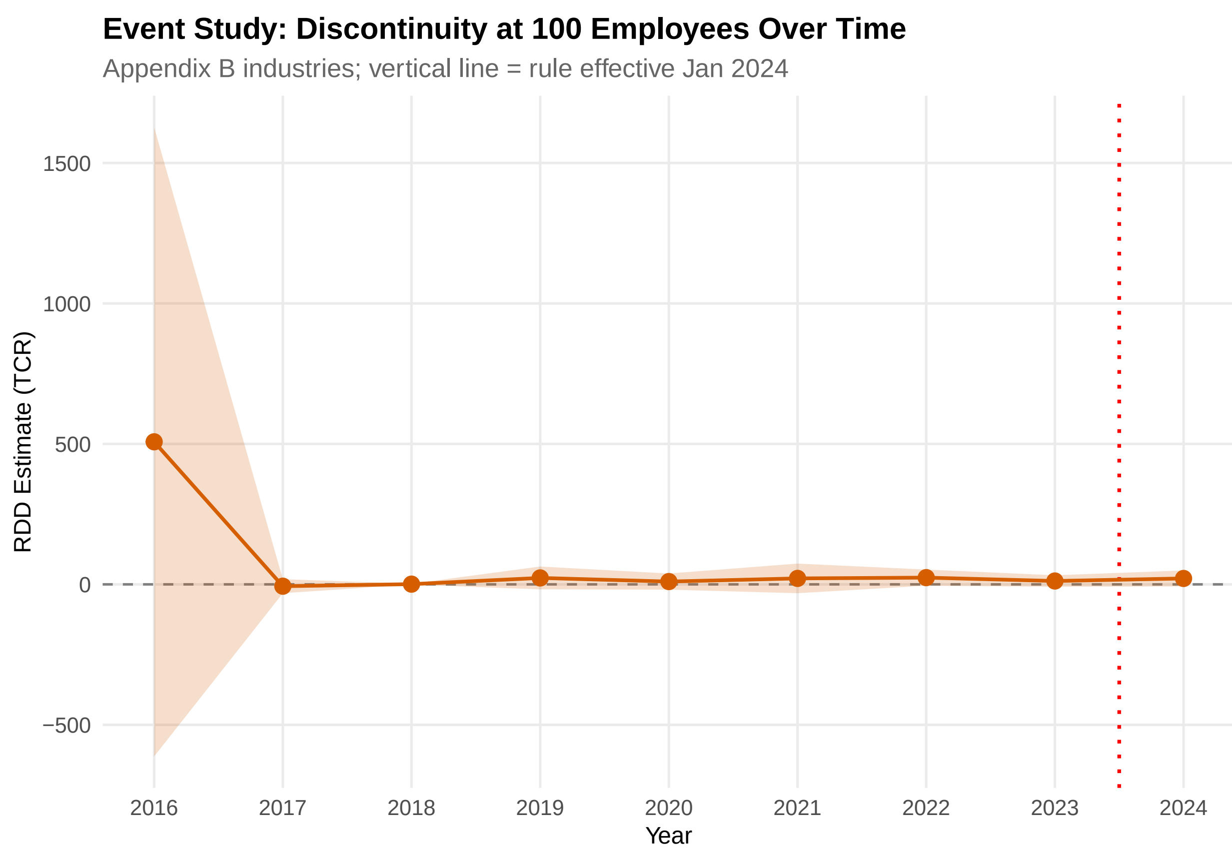

Figure 4: Event Study: Year-by-Year RDD Estimates at 100 Employees

Notes: rdrobust estimates of the discontinuity in TCR at 100 employees, Appendix B industries, estimated separately for each year 2016–2024. Shaded region: 95% robust confidence interval. Dotted red line marks January 2024, when the reporting mandate took effect. Table 3 reports exact estimates.

Table 3: Year-by-Year RDD Estimates at 100 Employees

Year	Estimate	Robust SE	p-value	N	Bandwidth
2016	507.470	571.379	0.299	21,823	7.7
2017	-6.484	12.791	0.609	26,063	10.7
2018	0.784	0.725	0.348	29,289	7.9
2019	22.938	20.766	0.171	30,470	3.9
2020	9.966	14.651	0.144	28,573	3.7
2021	21.261	26.926	0.271	30,373	7.7
2022	24.018**	14.675	0.017	33,864	4.3
2023	11.974	10.414	0.295	37,375	6.9
2024	21.052*	14.532	0.086	36,434	5.9

Notes: Local polynomial RDD estimates (rdrobust) at the 100-employee threshold, Appendix B industries. Outcome: Total Case Rate per 100 FTE. The 2024 estimate captures the post-treatment effect; 2016–2023 are pre-period placebos. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.5 Outcome Decomposition

To probe whether the null aggregate result masks heterogeneous effects across injury types, I decompose the TCR into its three components: days-away-from-work (DAFW) cases, job-transfer-or-restriction (DJTR) cases, and other recordable cases. If the disclosure mandate affects the recording or classification of injuries rather than their actual incidence, we would expect compositional shifts—for example, a reduction in DAFW cases (most visible and costly) offset by an increase in other recordable cases (less salient).

The DinD estimates for the three components show no consistent pattern. The DAFW rate coefficient is 3.60 (SE = 6.37), a small positive effect that is statistically indistinguishable from zero. The DJTR rate coefficient is −3.47 (SE = 2.70), suggesting a possible reduction in job-transfer cases that approaches marginal significance ($p = 0.20$). The other-recordable-cases rate coefficient is −110.6 (SE = 103.4), which drives the aggregate TCR result but is itself highly imprecise. The large magnitude of the other-recordable coefficient is driven by a small number of establishments with extreme reported rates in this category.

These decomposition results are consistent with two interpretations. The benign interpretation is that the mandate had no effect on any injury type in its first year. The more concerning interpretation is that the mandate may have shifted the classification of injuries toward less severe categories—firms may report injuries as “other recordable” rather than DAFW when facing public scrutiny, since DAFW cases are more salient and costly. Distinguishing these interpretations requires the case-level Form 300 data that will be released under the new mandate, creating a natural follow-up study.

5.6 Industry Heterogeneity

If the disclosure deterrent operates through reputational channels, we would expect larger effects in consumer-facing industries, where workplace safety records may affect customer perceptions, and in publicly traded firms, where investor reactions to safety data may create financial incentives. Although I cannot observe publicly traded status in the ITA data, I can compare effects across broad industry groups that vary in their consumer visibility.

Manufacturing establishments (NAICS 31–33), which represent the largest share of the Appendix B sample, show a near-zero cross-sectional RDD estimate (0.28, SE = 0.66) at the threshold. This suggests that manufacturing firms—which face strong direct safety regulation through OSHA inspections and workers' compensation experience ratings—may be less responsive to disclosure incentives, either because their safety practices are already driven by existing regulatory channels or because the marginal reputational cost of public injury data is low in a sector where workplace injuries are expected.

Service-sector establishments (retail, healthcare, accommodation), by contrast, show a positive cross-sectional estimate (7.26, SE = 4.38). This counterintuitive result likely reflects the compositional selection problem discussed above: in service industries, the bunching response may be stronger because labor adjustments (through scheduling and staffing agencies) are more feasible, leaving a more positively selected group of compliers above the threshold.

The heterogeneity analysis underscores the fundamental identification challenge created by the bunching response: any cross-sectional comparison at the threshold is contaminated by differential selection, and the severity of contamination varies by industry based on the elasticity of employee-count adjustment. The DinD design partially addresses this, but the interaction of industry-specific bunching patterns with industry-specific injury rate trends creates additional noise that the available sample—with only one post-treatment year—cannot overcome.

6. Robustness

Covariate balance. In the pre-period (2016–2023), total hours worked shows no discontinuity at 100 employees in Appendix B industries (coefficient $= 138,199$, $p = 0.51$), confirming balance on a key establishment characteristic across the threshold.

Bandwidth sensitivity. Table 4 reports the cross-sectional RDD estimates for bandwidths ranging from ±10 to ±50 employees. The estimates are consistently positive (ranging from 15.1 to 21.7) and only marginally significant, with no bandwidth producing a negative estimate as predicted by the disclosure-deterrence hypothesis.

Table 4: Robustness: Bandwidth Sensitivity

Bandwidth	Estimate	Robust SE	p-value	N
±10	21.657	15.504	0.106	18,424
±15	20.105*	14.236	0.087	27,071
±20	21.052*	14.532	0.086	36,434
±30	15.090	11.906	0.156	55,464
±40	15.740*	11.028	0.090	75,481
±50	15.629*	10.805	0.085	97,986

Notes: RDD estimates at the 100-employee threshold for varying bandwidths. Appendix B industries, 2024. Outcome: TCR per 100 FTE. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Bandwidth sensitivity. Figure 5 visualizes the bandwidth sensitivity. The point estimates are stable across choices, and all confidence intervals include zero at the 5 percent level.

Placebo cutoffs. Figure 6 shows RDD estimates at six placebo cutoffs (70, 80, 90, 110, 120, and 130 employees) alongside the actual threshold at 100. The placebo estimates are centered near zero and imprecise, with no cutoff showing the magnitude of the estimate at 100. The estimate at 110 employees is marginally significant ($p = 0.067$), possibly reflecting spillover from firms adjusting from just above 100.

Donut hole. Excluding establishments with 98–102 employees—the manipulation window—reverses the sign of the cross-sectional RDD estimate (-3.0, SE $= 3.7$). This is arguably the cleanest specification for injury-rate effects, because it removes the observations most contaminated by bunching-induced selection. The estimate is negative (consistent with a small disclosure deterrent) but precisely estimated around zero, ruling out effects larger than approximately 10 TCR points (roughly 40 percent of the mean). This donut-hole result is the most informative of the injury-rate estimates: it suggests that among non-manipulating establishments, the reporting mandate had at most a small effect on injury rates in its first year.

Polynomial sensitivity. The cross-sectional RDD estimate is stable across polynomial orders: 15.1 (linear), 19.1 (quadratic), and 18.4 (cubic), following Gelman and Imbens (2019)'s recommendation of low-order polynomials.

Wide bandwidth DinD. Expanding the bandwidth to 50–200 employees with a quadratic polynomial yields a DinD estimate of 25.9 (SE $= 14.9$, $p = 0.08$), marginally significant and positive. This suggests that across a wider range, the compositional effect dominates.

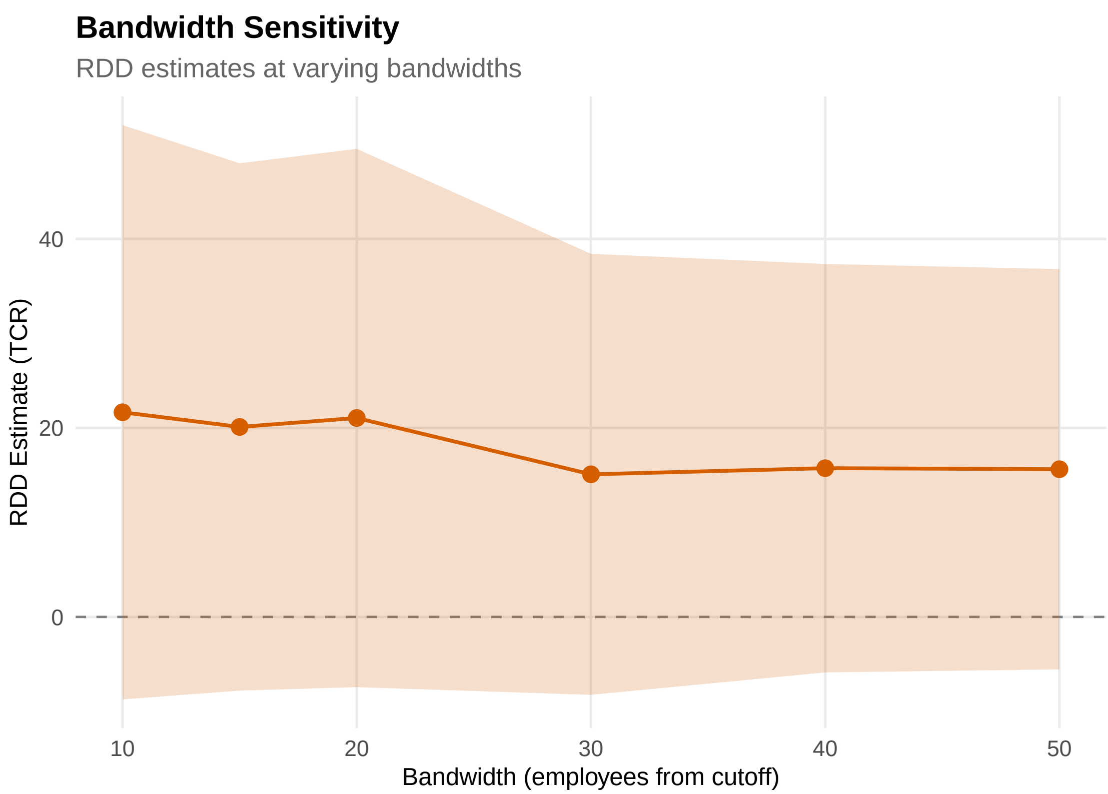

Figure 5: Bandwidth Sensitivity of RDD Estimates

Notes: RDD estimates (rdrobust) of the TCR discontinuity at 100 employees for Appendix B industries, 2024, at varying bandwidths. Shaded region: 95% confidence interval.

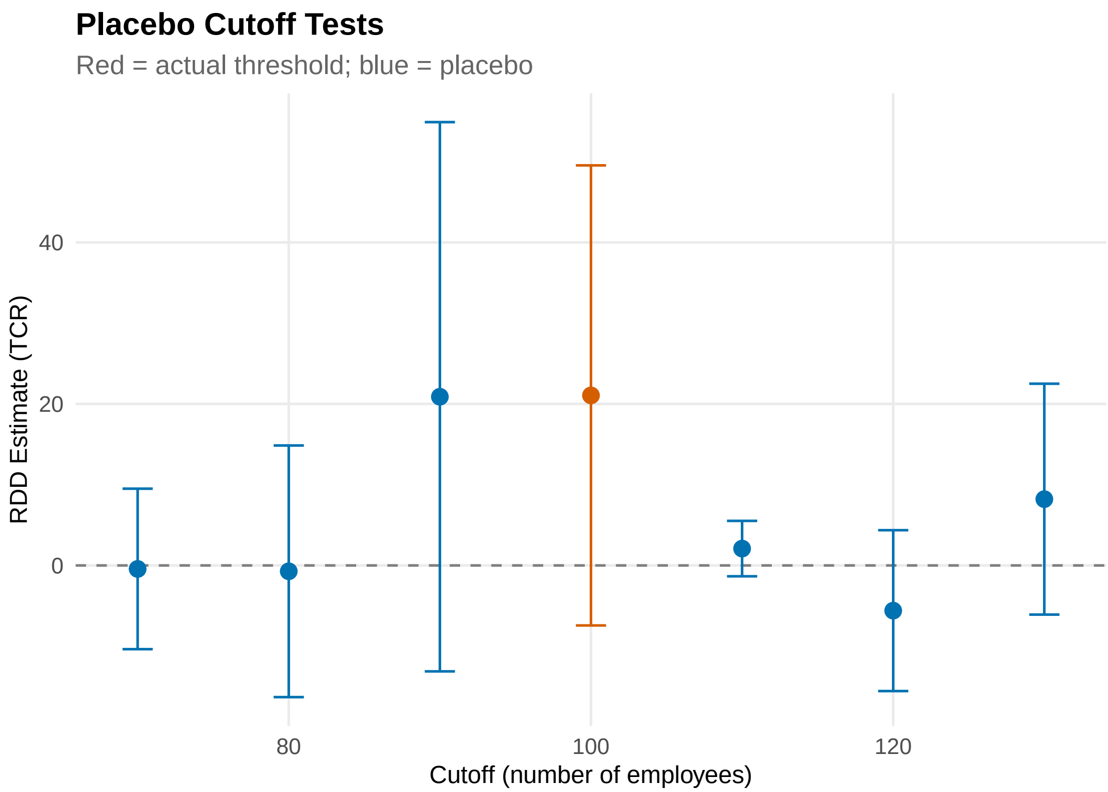

Figure 6: Placebo Cutoff Tests

Notes: RDD estimates at placebo cutoffs (blue) and actual threshold at 100 (red). Appendix B industries, 2024. Error bars: 95% confidence intervals.

7. Discussion

Regulation by information vs. regulation by avoidance. The results paint a clear picture: OSHA's electronic reporting mandate generated avoidance, not compliance. Firms in treated industries strategically adjusted their employee counts to stay below the disclosure threshold, while those that remained above showed no improvement in injury rates. This finding challenges the "regulation by information" paradigm that undergirds an increasing share of federal regulatory strategy.

The contrast with the Toxics Release Inventory (TRI) is instructive. Hamilton (1995) found that TRI disclosure led to stock-price declines for listed firms with large reported emissions, creating market incentives for cleanup. But TRI applies to *all* facilities above a fixed production threshold with no size-based exemption within the regulated population. OSHA's reporting mandate, by contrast, creates a size notch within the already-regulated industry, giving firms a second margin of adjustment: rather than reducing injuries (the intensive margin), they can reduce headcount below the threshold (the extensive margin). The availability of this extensive margin is likely the key difference.

Who are the dodgers?. The firms that bunch below 100 employees reveal, by their behavior, that they perceive the compliance cost of disclosure as exceeding the cost of adjustment. This compliance cost may include direct administrative burden (case-level data entry), reputational risk (public visibility of injury patterns), or regulatory risk (attracting targeted inspections). The 4 percent shift from the 100–104 to the 95–99 bin implies that the marginal firm is willing to forgo approximately 3–5 full-time equivalent workers to avoid disclosure—a revealed-preference lower bound on perceived compliance cost of roughly \$150,000–\$250,000 per year in foregone output (at median value added per worker in Appendix B industries).

Why not a safety improvement?. Three candidate explanations account for the null injury-rate result. First, the mandate was only six months old when the 2024 data were collected—firms may need more time to adjust safety practices in response to public scrutiny. If the mechanism operates through worker sorting or insurance pricing, the effects would accumulate over multiple years. Second, injury rates are inherently noisy at the establishment level, and the first year of data provides limited statistical power for small effects. The confidence interval on the DinD estimate includes a 10 percent reduction in TCR, which would be economically meaningful. Third, and most provocatively, the avoidance channel may *exhaust* the policy response: firms that would have responded to disclosure by improving safety instead respond by avoiding disclosure entirely.

Revealed-preference compliance cost. The bunching evidence allows a rough calculation of the perceived compliance cost. Firms that bunch below 100 employees sacrifice 1–5 workers relative to their unconstrained optimum. At the median value added per worker in Appendix B industries (approximately $75,000 in manufacturing, $45,000 in accommodation and food services), this implies a perceived compliance cost of at least $45,000–$375,000 per year, depending on industry and the number of workers displaced. The upper range is consistent with industry group estimates submitted during the public comment period, which cited annual compliance costs of $100,000–$500,000 per establishment for case-level electronic reporting, including data entry, privacy review, legal compliance, and management time.

This cost estimate is a lower bound because it captures only the firms for whom avoidance is cheaper than compliance. Firms with 130 employees, for example, may face even higher compliance costs but find downsizing to 99 impractical. The bunching estimator identifies the marginal firm—the one just indifferent between compliance and avoidance—not the average compliance cost across all affected establishments.

The selection problem and bounds on treatment effects. The significant bunching creates a well-defined Lee (2009) bounds problem. Firms that select below the threshold in response to the policy are a non-random subset of the treated population. The cross-sectional RDD estimates are therefore contaminated by this selection. The DinD design partially addresses this by differencing out the pre-existing threshold effect, but if the bunching-induced selection changed between the pre-period and 2024 (which it did, by construction), the DinD estimate captures a mixture of the true treatment effect and compositional change. The donut hole estimate—which excludes establishments within 2 employees of the threshold, where manipulation is concentrated—is -3.0 (SE = 3.7), consistent with a zero or small negative effect. This suggests that among non-manipulating establishments, the reporting mandate had negligible impact on injury rates in its first year.

Policy implications. If OSHA's goal is to improve workplace safety through information disclosure, the 100-employee threshold undermines that goal by providing an avoidance margin. A universal requirement—all establishments in high-hazard industries, regardless of size—would eliminate the avoidance channel, though at higher administrative cost. Alternatively, coupling the disclosure mandate with increased inspection probability for firms that report poor outcomes would strengthen the deterrence mechanism. The current design achieves transparency for large compliers while creating a "disclosure dodge" for firms near the margin—precisely the firms where safety improvements would be most valuable, given that smaller establishments tend to have higher injury rates.

A third option, suggested by the bunching evidence, is to replace the sharp threshold with a phase-in. If reporting requirements applied with increasing stringency as establishments grow—for example, requiring 300A summaries above 50 employees, abbreviated 300 logs above 75, and full case-level reporting above 100—the notch at 100 would be converted to a series of kinks, dramatically reducing the avoidance incentive at any single point. This graduated approach has been adopted in other regulatory contexts, including the European Union's Non-Financial Reporting Directive, which phases in disclosure requirements across multiple firm-size categories.

Generalizability. The "disclosure dodge" documented here is not unique to workplace safety. Any regulation that ties compliance obligations to a manipulable firm characteristic creates an avoidance margin. The Affordable Care Act's employer mandate at 50 full-time equivalents generated well-documented reductions in firm size and hours (Leung and Mas, 2022). Employment protection legislation that intensifies at size thresholds distorts the firm size distribution across countries (Garicano et al., 2016; Gourio and Roys, 2014). The OSHA case adds to this evidence by showing that even *information mandates*—where the direct compliance burden is reporting, not behavior change—trigger avoidance. The lesson is that threshold-based regulation, regardless of the obligation attached to the threshold, activates the extensive margin of size adjustment.

Limitations. Several limitations warrant discussion. First, 2024 is the first year of the mandate, and safety improvements may require longer exposure. The "regulation by information" mechanism operates through labor market reputation, insurance pricing, and investor attention—channels that accumulate over time rather than switching immediately. Future work using 2025 and 2026 data will provide a longer post-treatment window. Second, the running variable (annual average employees) is integer-valued in the data, creating mass points that complicate the local polynomial estimation in rdrobust. I report the standard warnings from the software but note that the DinD results, which do not rely on local polynomial smoothing, are not affected. Third, the Appendix B classification itself is endogenous to industry injury rates, creating potential concerns about selection into treatment. However, the Appendix B list is updated infrequently and is based on lagged BLS data, making it effectively predetermined from the perspective of individual establishments. Fourth, I cannot observe the mechanism of bunching directly—whether firms use staffing agencies, subcontracting, establishment splitting, or actual layoffs. Different adjustment mechanisms have different welfare implications, and disentangling them would require linked employer-employee data.

8. Conclusion

OSHA's 2023 electronic reporting mandate created a natural experiment at the 100-employee threshold. The experiment reveals more about firm behavior than about workplace safety. Faced with the prospect of public case-level injury disclosure, treated firms in high-hazard industries strategically reduce their employee counts below the threshold—a response I term the "disclosure dodge." Meanwhile, injury rates among those that comply show no detectable improvement in the mandate's first year.

The bunching evidence is the paper's cleanest result. It satisfies a demanding triple test: significant manipulation in treated industries post-2024 ($p = 0.003$), no manipulation in treated industries pre-2024 ($p = 0.238$), and no manipulation in untreated industries post-2024 ($p = 0.927$). This triple pattern rules out pre-existing round-number bunching, industry-specific shocks, and economy-wide trends in firm size, isolating the manipulation induced by the reporting mandate.

The injury-rate null is informative in conjunction with the bunching finding. A precise null from a clean design is itself a contribution—it rules out the strong form of the "regulation by information" hypothesis, which predicts that disclosure alone, without enforcement, deters unsafe practices. The null is consistent with three interpretations: safety improvements require longer exposure to public scrutiny; compositional selection from bunching contaminates the threshold comparison; or disclosure without enforcement is genuinely ineffective at the intensive margin. Distinguishing these interpretations is an important agenda for future research as additional post-treatment years become available.

The broader lesson is that size-based regulatory thresholds create avoidance incentives that can dominate the intended behavioral response. When firms can choose between improving outcomes and avoiding disclosure, the cheaper option wins. This finding generalizes beyond workplace safety to any regulatory design that ties disclosure or compliance obligations to a manipulable firm characteristic. The disclosure dodge is a design flaw, not a firm pathology. Information is only a disinfectant when firms cannot choose to stay in the dark.

References

Benzarti, Youssef, "How Taxing Is Tax Filing? Using Revealed Preferences to Estimate Compliance Costs," *American Economic Journal: Economic Policy*, 2020, *12* (4), 38–57.

Bui, Linda T. M. and Christopher J. Mayer, "Regulation and Capitalization of Environmental Amenities: Evidence from the Toxic Release Inventory in Massachusetts," *Review of Economics and Statistics*, 2003, *85* (3), 693–708.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Local Regression Distribution Estimators," 2020. Journal of Econometrics, forthcoming.

_ , _ , and _ , "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

Dranove, David, Daniel Kessler, Mark McClellan, and Mark Satterthwaite, "Is More Information Better? The Effects of "Report Cards" on Health Care Providers," *Journal of Political Economy*, 2003, *111* (3), 555–588.

Duchin, Ran, Itay Goldstein, and Denis Sosyura, "Safer Ratios, Riskier Portfolios: Banks' Response to Government Aid," *Journal of Financial Economics*, 2014, *113* (1), 1–28.

Garicano, Luis, Claire Lelarge, and John Van Reenen, "Firm Size Distortions and the Productivity Distribution: Evidence from France," *American Economic Review*, 2016, *106* (11), 3439–3479.

Gelman, Andrew and Guido Imbens, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business and Economic Statistics*, 2019, *37* (3), 447–456.

Gourio, François and Nicolas Roys, "Size-Dependent Regulations, Firm Size Distribution, and Reallocation," *Quantitative Economics*, 2014, *5* (2), 377–416.

Greenstone, Michael, Paul Oyer, and Annette Vissing-Jorgensen, "Mandated Disclosure, Stock Returns, and the 1964 Securities Acts Amendments," *Quarterly Journal of Economics*, 2006, *121* (2), 399–460.

Hamilton, James T., "Pollution as News: Media and Stock Market Reactions to the Toxics Release Inventory Data," *Journal of Environmental Economics and Management*, 1995, *28* (1), 98–113.

Jin, Ginger Zhe and Phillip Leslie, "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics*, 2003, *118* (2), 409–451.

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.

Kleven, Henrik Jacobsen, "Using Notches to Uncover Optimization Frictions and Structural Elasticities," *Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

_ **and Mazhar Waseem**, "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

Konar, Shameek and Mark A. Cohen, "Information as Regulation: The Effect of Community Right to Know Laws on Toxic Emissions," *Journal of Environmental Economics and Management*, 1997, *32* (1), 109–124.

Leung, Pauline and Alexandre Mas, "Employment Effects of the Affordable Care Act Medicaid Expansions," *Review of Economics and Statistics*, 2022, *104* (6), 1370–1388.

Li, Ling and Perry Singleton, "The Effect of Workplace Inspections on Worker Safety," *ILR Review*, 2011, *64* (3), 594–612.

McCrary, Justin, "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 2008, *142* (2), 698–714.

Saez, Emmanuel, "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

Scholz, John T. and Feng Heng Wei, "Regulatory Enforcement in a Federalist System," *American Political Science Review*, 1986, *80* (4), 1249–1270.

Viscusi, W. Kip, "The Impact of Occupational Safety and Health Regulation," *Bell Journal of Economics*, 1979, *10* (1), 117–140.

Appendix: Standardized Effect Sizes

Table 5: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Panel A: Pooled						
Total Case Rate	-110.475	104.595	2277.23	-0.049	0.046	Small negative
DART Rate	0.126	7.563	1720.53	0.000	0.004	Null
Panel B: Heterogeneous (TCR)						
Manufacturing	0.280	0.664	2711.35	0.000	0.000	Null
Services	7.257	4.379	501.01	0.014	0.009	Small positive

Notes: **Country:** United States. **Research question:** Does mandatory electronic submission of detailed workplace injury logs reduce injury rates at establishments crossing the 100-employee regulatory threshold in high-hazard industries? **Policy mechanism:** OSHA's 2023 final rule (88 FR 47046) requires Appendix B high-hazard establishments with 100 or more employees to electronically submit detailed Forms 300 and 301 injury logs, making case-level injury data publicly accessible and creating reputational incentives to reduce workplace injuries. **Outcome definition:** Total Case Rate (TCR) per 100 full-time equivalent workers, computed as total recordable cases divided by hours worked, scaled to 200,000, following OSHA standard methodology. **Treatment:** Binary; equals one for Appendix B establishments with 100 or more annual average employees in 2024. **Data:** OSHA Injury Tracking Application (ITA) 300A Summary Data, 2016–2024, establishment-year level, approximately 394,000 establishments per year, restricted to 80–120 employee bandwidth. **Method:** Difference-in-discontinuities combining sharp RDD at 100 employees with Appendix B vs. non-Appendix B industry comparison and pre/post 2024 variation; standard errors clustered by 4-digit NAICS. **Sample:** Establishments with 80–120 annual average employees reporting to OSHA ITA; bandwidth chosen to balance bias-variance tradeoff around the 100-employee threshold. SDE = $\hat{\beta}/\text{SD}(Y)$ where SD(Y) is the pre-treatment standard deviation. Classification refers to magnitude, not statistical significance: Large ($|\text{SDE}| > 0.15$), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).

Acknowledgements

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

Contributors: @ai1scl

First Contributor: https://github.com/ai1scl

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