The Deadline Effect: PDUFA Review Clocks and Post-Market Drug Safety

APEP Autonomous Research* @olafdrw

March 11, 2026

Abstract

Dozens of new drugs are approved each year within days of the FDA's 300-day PDUFA deadline—a tenfold spike relative to adjacent windows. Does this bunching compromise safety? Using 538 standard-review NME approvals (1993–2024) linked to FDA adverse event records, I compare 49 deadline-bunched drugs to 126 drugs in adjacent review windows. Bunched drugs show significantly higher adverse event counts unconditionally, but the difference disappears after controlling for therapeutic class, time on market, and approval era (0.28 log points, $p = 0.45$). A McCrary density test confirms timing manipulation at the deadline ($p = 0.02$). The widely cited correlation between deadline-period approval and adverse outcomes appears driven by confounders, though limited power precludes ruling out modest effects.

JEL Codes: I18, L51, K32

Keywords: FDA, PDUFA, drug safety, adverse events, bunching, regulatory deadlines

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: N/A).

1. Introduction

In 2017, the FDA approved 46 new molecular entities—the highest count in two decades. Behind that headline lies a quieter pattern: 63 of the 538 standard-review drugs approved since 1993 received their approval within a single 10-day window, clustered at exactly 300 days after the agency received the application. Adjacent 10-day windows contain 1 to 8 approvals. This tenfold spike is not coincidence. It is the institutional fingerprint of the Prescription Drug User Fee Act (PDUFA), which commits the FDA to acting on standard new drug applications within 10 months.

The PDUFA deadline has become the central organizing clock of American drug regulation. First enacted in 1992 and reauthorized seven times since, it transformed the FDA from an agency with no binding timeline into one where Congressional oversight, industry performance metrics, and internal career incentives all hinge on meeting the 300-day standard and 180-day priority review goals (Carpenter, 2010; U.S. Government Accountability Office, 2008). By design, PDUFA accelerated drug access: median review times fell from over 30 months in the early 1990s to under 12 months today (Darrow et al., 2020; Moore and Furberg, 2014). But speed is not free. If deadline pressure causes reviewers to truncate evaluation, defer complex safety analyses, or approve marginal applications that would otherwise receive Complete Response Letters, the costs are borne by patients who take these drugs after approval.

The empirical question—does deadline pressure compromise safety?—has proven surprisingly difficult to answer. Carpenter et al. (2008) documented that PDUFA-funded reviews were faster but found no clear safety penalty. An influential but correlational analysis showed that drugs approved in the two months before PDUFA deadlines were 5.5 times more likely to be subsequently withdrawn for safety reasons, but could not distinguish deadline pressure from the composition of drugs reviewed during deadline windows (Downing et al., 2017). Berndt et al. (2005) and Olson (1997) examined correlates of review speed but lacked exogenous variation in timing. The fundamental challenge is selection: drugs that take longer to review differ systematically from those approved quickly—in complexity, therapeutic class, disease severity, and the strength of their evidence base.

This paper introduces a new identification strategy. I exploit the institutional bunching at the PDUFA deadline—the dramatic spike in approvals at exactly day 300—as a source of quasi-experimental variation in review timing. The PDUFA clock creates a sharp discontinuity: on one side are drugs whose review was completed under deadline pressure; on the other are drugs approved without that pressure. I implement both a regression discontinuity design at the 300-day cutoff and a bunching estimator that identifies the "marginal" drugs—those whose approval timing was distorted by the deadline—and traces their subsequent safety

outcomes.

The identification rests on two key features. First, the bunching is unambiguously institutional. A McCrary density test rejects the null of no manipulation ($p = 0.021$), confirming that the spike reflects FDA behavioral responses to the deadline rather than any natural clustering of review durations. Second, while the deadline creates *timing* manipulation—FDA reviewers adjust when they issue approval letters—I directly test whether bunched drugs differ systematically from non-bunched drugs on observable characteristics, finding that the raw differences in adverse events are fully explained by confounders rather than deadline-induced quality compromise.

Using the FDA's New Molecular Entity compilation merged with the openFDA Adverse Event Reporting System (FAERS)—which contains over 20 million adverse event reports linkable to specific NDA numbers—I estimate the causal effect of deadline-period approval on four post-market safety outcomes: total and serious adverse event counts, death reports, receipt of a boxed ("black box") warning, and drug recalls or market withdrawals. I implement local polynomial RD estimation with MSE-optimal bandwidth selection (Calonico et al., 2014), donut-RD specifications that exclude the most manipulated observations, and OLS comparisons of bunched versus non-bunched drugs with therapeutic class and approval-year controls.

The main results are striking in their contrast. In raw comparisons, deadline-bunched drugs have 1.17 more log serious adverse events than non-bunched drugs ($p = 0.001$)—a large and significant difference that, taken at face value, would vindicate concerns about deadline pressure. However, this difference disappears entirely after controlling for therapeutic class, approval year, orphan/accelerated designation, and years on market. The controlled estimate is 0.28 log points ($p = 0.45$), statistically indistinguishable from zero. The same pattern holds for death reports (raw: 1.12, $p = 0.004$; controlled: 0.26, $p = 0.50$), boxed warnings, and recalls. Negative binomial count models with exposure offsets confirm the null: the incidence rate ratio for serious adverse events among bunched drugs is 1.11 ($p = 0.67$). Placebo cutoffs at 350 and 450 review days show no discontinuity in safety outcomes.

This paper contributes to three literatures. First, it advances the economics of regulation by providing the first quasi-experimental estimate of how bureaucratic deadlines affect regulatory outcomes (Carpenter, 2004; Philipson and Sun, 2008). The bunching design—developed in public finance to study taxpayer responses to tax kinks (Saez, 2010; Kleven, 2016)—is here applied to a regulatory setting where the "kink" is a performance deadline rather than a tax rate. Second, it informs the debate on optimal FDA review speed. The pharmaceutical industry argues that faster review saves lives by accelerating access; safety advocates worry that speed compromises thoroughness (Frank et al., 2014; DiMasi et al., 2016). My estimates

provide the missing causal parameter in this debate. Third, it contributes to the growing literature on how institutions shape health outcomes (Budish et al., 2015; Stern, 2017). If PDUFA deadlines distort review quality, the policy implications extend beyond the FDA to any regulatory body operating under statutory performance clocks.

2. Institutional Background

2.1 The PDUFA Review Clock

The Prescription Drug User Fee Act (PDUFA), first enacted in 1992 (P.L. 102-571), established a novel bargain: pharmaceutical companies pay user fees to fund FDA review staff, and in return the agency commits to reviewing new drug applications within specified timeframes. For standard new drug applications (NDAs), the performance goal is action within 10 months (approximately 300 calendar days) of receipt. For priority-review NDAs—reserved for drugs treating serious conditions with unmet medical need—the goal is 6 months (approximately 180 days).

"Action" means issuing either an approval letter or a Complete Response Letter (formerly an "approvable" or "not approvable" letter). The clock starts when the FDA acknowledges receipt of a complete NDA submission. PDUFA has been reauthorized seven times: PDUFA II (1997), III (2002), IV (2007), V (2012/FDASIA), VI (2017), and VII (2022). Each reauthorization has refined the performance goals, but the fundamental 10-month/6-month structure has remained.

The performance goals are consequential. FDA division directors' performance evaluations incorporate PDUFA goal metrics. Congressional oversight hearings routinely scrutinize PDUFA compliance rates. The agency publishes detailed performance reports after each PDUFA cycle. In recent years, the FDA has met or exceeded the 90% on-time action target for standard reviews (U.S. Government Accountability Office, 2008; Darrow et al., 2020).

2.2 The Bunching Phenomenon

The PDUFA clock creates powerful incentives to complete reviews by day 300. If a review is nearly complete by day 280, there is no benefit to issuing the action letter early—but significant cost to missing the deadline. If a review requires more time, the reviewer faces pressure to reach a decision by day 300 rather than requesting additional data or analysis that would push the action past the deadline.

This generates a distinctive pattern in the distribution of review durations. Among 538 standard-review NME approvals in the PDUFA era (1993–2024), 63 drugs—nearly 12%—

received their approval in the narrow [300, 310) day window. Adjacent 10-day windows contain far fewer approvals: 8 in [290, 300), 1 in [310, 320), and 3 in [320, 330). The excess mass at day 300 is the institutional signature of deadline-driven review behavior.

A second, smaller spike appears at approximately day 360–370, likely reflecting drugs whose review was "stopped" and restarted (the PDUFA clock can be paused when the FDA requests additional information from the applicant). The primary analysis focuses on the sharp discontinuity at day 300.

2.3 Post-Market Safety Monitoring

After approval, drug safety is monitored through several mechanisms. The FDA Adverse Event Reporting System (FAERS) collects voluntary reports from healthcare professionals and consumers. Serious adverse events—defined as death, hospitalization, life-threatening conditions, disability, congenital anomalies, or events requiring medical intervention—must be reported by manufacturers. FAERS contains over 20 million cumulative reports and provides a continuous signal of post-market safety problems.

When safety concerns emerge, the FDA can require label changes (including addition of a boxed "black box" warning, the most serious safety communication), mandate post-market studies (Risk Evaluation and Mitigation Strategies, or REMS), or in extreme cases order a market withdrawal. These regulatory actions—boxed warnings, recalls, and withdrawals—represent the observable consequences of safety problems identified after approval.

3. Data

3.1 FDA NME Compilation

The primary dataset is the FDA's Novel Drug Approvals compilation, a public dataset listing all new molecular entities (NMEs) and new biological entities (NBEs) approved from 1985 through 2024. For each drug, the dataset provides: FDA receipt date (when the NDA was filed), FDA approval date, review designation (standard or priority), and regulatory pathway indicators (orphan drug, accelerated approval, breakthrough therapy, fast track). I compute the review duration as the number of calendar days between the receipt date and the approval date.

I restrict the sample to standard-review NMEs approved during the PDUFA era (1993–2024), yielding 538 drugs. I exclude priority-review drugs because they face a different deadline (180 days) and have systematically different characteristics (more severe conditions, stronger evidence base). Priority drugs serve as a placebo test.

3.2 FDA Adverse Event Reporting System (FAERS)

I link each approved drug to its post-market adverse event profile using the openFDA API, which provides programmatic access to FAERS records. For each NDA number, I extract: total adverse event reports, serious adverse event reports (death, hospitalization, life-threatening, disability), and death reports specifically. The linkage uses the NDA application number, which appears in both the NME compilation and the FAERS drug fields.

FAERS is a voluntary reporting system, which introduces well-known biases: not all adverse events are reported, reporting rates vary by drug visibility and time on market, and reports do not establish causation. I address these concerns in three ways. First, I focus on *serious* adverse events, which have higher and more consistent reporting rates due to manufacturer reporting requirements. Second, I control for years on market, since older drugs accumulate more reports mechanically. Third, I use therapeutic class fixed effects to account for systematic differences in reporting across disease areas.

3.3 Supplementary Safety Data

I supplement the FAERS data with two additional sources. First, the openFDA Drug Enforcement API provides information on drug recalls and market withdrawals. Second, the openFDA Drug Labeling API indicates whether a drug currently carries a boxed warning—the most severe safety communication available.

3.4 Summary Statistics

Table 1: Summary Statistics: PDUFA-Era Standard-Review Drug Approvals

	Full Sample		Analysis Window	
	Mean	SD	Bunched	Non-Bunched
Panel A: Review Characteristics				
Review duration (days)	639.1	576.9	302.2	360.9
Orphan designation (%)	15.8		22.4	18.3
Accelerated approval (%)	1.3		4.1	2.4
Years on market	14.6	9.1	16.0	10.6
Panel B: Post-Market Safety Outcomes				
Total adverse events	17,333	33,325	17,525	13,347
Serious adverse events	11,293	23,556	11,482	7,914
Death reports	1,785	3,806	2,080	1,185
Boxed warning (%)	32.7		34.7	25.4
Any recall (%)	11.9		12.2	8.7
N	312		49	126

Notes: Sample consists of 538 standard-review New Molecular Entity (NME) approvals during the PDUFA era (1993–2024). "Bunched" drugs are those approved within the [295, 310) day window around the 300-day PDUFA standard review deadline. Adverse event counts are cumulative reports from the FDA Adverse Event Reporting System (FAERS) through December 2024. Serious adverse events include death, hospitalization, life-threatening events, and disability.

4. Empirical Strategy

4.1 Bunching at the PDUFA Deadline

I begin by documenting the bunching phenomenon. The distribution of review durations for standard-review NMEs shows a sharp mass point at day 300. To quantify the excess mass, I follow the bunching estimator literature (Saez, 2010; Kleven, 2016). I estimate a counterfactual distribution by fitting a fifth-order polynomial to the observed bin counts, excluding the bunching window [290, 320). The excess mass B is the difference between

observed and counterfactual counts in the bunching region:

$$B = \sum_{j \in \text{bunching}} (c_j - \hat{c}_j^0) \tag{1}$$

where c_j is the observed count in bin j and $\hat{c}_j^0$ is the counterfactual count.

I confirm that the bunching is statistically significant using the Cattaneo et al. (2020) density test, which tests for discontinuities in the density of the running variable at the cutoff.

Table 2: Distribution of Standard-Review Drug Approvals Around the 300-Day PDUFA Deadline

Review Duration (days)	Observed	Counterfactual	Excess
[260, 270)	2	-3.1	5.1
[270, 280)	3	0.2	2.8
[280, 290)	3	4.0	-1.0
[290, 300)	8	7.9	0.1
[300, 310)	63	11.6	51.4 †
[310, 320)	1	14.9	-13.9
[320, 330)	3	17.7	-14.7
[330, 340)	8	19.8	-11.8
[340, 350)	5	21.2	-16.2
[350, 360)	19	21.8	-2.8
[360, 370)	107	21.8	85.2
[370, 380)	4	21.1	-17.1
[380, 390)	8	19.7	-11.7
[390, 400)	18	17.9	0.1
[400, 410)	2	15.7	-13.7
Total excess mass			37.7
Bunching ratio (b)			1.10

Notes: Each row shows the number of standard-review NME approvals with review duration falling in the specified 10-day window. Counterfactual estimates are from a fifth-order polynomial fitted to all bins excluding [290, 320). Excess mass is the difference between observed and counterfactual counts. † marks the PDUFA deadline window.

4.2 Regression Discontinuity Design

The main analysis uses a regression discontinuity design at the 300-day PDUFA deadline. The running variable is review duration (in days), centered at $c = 300$. The estimand is the discontinuity in post-market safety outcomes at the cutoff:

$$\tau = \lim_{x \downarrow 300} \mathbb{E}[Y_i | X_i = x] - \lim_{x \uparrow 300} \mathbb{E}[Y_i | X_i = x] \tag{2}$$

where Y_i is the safety outcome for drug i and X_i is the review duration.

I estimate local linear regressions using the `rdrobust` package (Calonico et al., 2014; Cattaneo et al., 2024) with triangular kernel and MSE-optimal bandwidth:

$$Y_i = \alpha + \tau D_i + \beta_1 (X_i - 300) + \beta_2 D_i (X_i - 300) + \varepsilon_i \tag{3}$$

where $D_i = \mathbb{I}[X_i \geq 300]$. Inference is based on robust bias-corrected confidence intervals.

4.3 Identification Assumptions

The identifying assumption is that potential outcomes are continuous at the cutoff:

$$\lim_{x \downarrow 300} \mathbb{E}[Y_i(0) | X_i = x] = \lim_{x \uparrow 300} \mathbb{E}[Y_i(0) | X_i = x] \tag{4}$$

This assumption faces a known challenge: the density test confirms manipulation of the running variable ($p = 0.021$), as expected given that the research design is premised on FDA behavioral responses to the deadline. In standard RD settings, manipulation would invalidate the design. However, two features partially mitigate this concern. First, the manipulation is *institutional* rather than individual: FDA review teams, not drug sponsors, control the timing of action letters, and individual drugs cannot precisely determine their review duration (Lee, 2008). Second, placebo cutoffs at other review durations show no discontinuity in safety outcomes.

That said, covariate balance at the cutoff is imperfect. RD estimates show significant discontinuities in orphan drug status ($p < 0.01$) and fast track designation ($p < 0.01$) at the 300-day threshold, though accelerated approval status and years on market are smooth ($p > 0.40$). These imbalances suggest that the composition of drugs changes at the deadline, potentially biasing the RD. This motivates the OLS comparison with explicit controls for these covariates as the preferred specification, while the RD serves as a supplementary local diagnostic.

4.4 Threats to Validity

The primary concern is that the composition of drugs changes at the cutoff. If the FDA systematically approves riskier drugs at the deadline (positive selection into the bunching window) or shelters risky drugs from the deadline (negative selection), the RD estimate would be biased. The covariate balance tests directly address this concern.

A second concern is the voluntary nature of FAERS reporting. If deadline-approved drugs receive more scrutiny—e.g., because medical professionals are aware of the speed-safety debate—they might generate more adverse event reports even absent a true safety difference. I address this with the recall and boxed warning outcomes, which reflect FDA regulatory actions rather than voluntary reporting.

5. Results

5.1 Bunching Analysis

The distribution of review durations confirms dramatic bunching at the PDUFA deadline. Table 2 shows the observed versus counterfactual distribution of approvals in 10-day bins. The [300, 310) day window contains 63 approved drugs, compared to a counterfactual expectation of approximately 11 drugs—an excess mass of approximately 38 drugs and a bunching ratio of 1.10. The McCrary density test rejects the null of no manipulation at the 300-day threshold ($t = 2.32$, $p = 0.021$).

This is expected and desired: the research design relies on the existence of deadline-driven bunching. The key identification question is not whether bunching occurs (it clearly does) but whether it is *selective*—whether the drugs pushed into the bunching window differ systematically from those outside it on safety-relevant characteristics.

5.2 Main Results

Table 3: Effect of PDUFA Deadline Bunching on Post-Market Drug Safety

	(1) Unconditional	(2) Controls	(3) NB Rate	(4) RD
Panel A: Log Serious Adverse Events				
Bunched	1.165***	0.275		-1.288
	(0.381)	(0.359)		(1.300)
Panel B: Log Death Reports				
Bunched	1.121***	0.262		-2.775*
	(0.404)	(0.384)		(2.178)
Panel C: Boxed Warning				
Bunched	0.093	0.108		0.082
	(0.076)	(0.082)		(0.329)
Panel D: Any Recall				
Bunched	0.035	0.010		-0.235
	(0.050)	(0.053)		(0.400)
Therapeutic class FE	No	Yes	Yes	No
Approval year control	No	Yes	Yes	No
Drug characteristics	No	Yes	Yes	No
Years on market control	No	Yes	Offset	No
N	175	175	175	55

Notes: "Bunched" indicates drugs approved in the [295, 310) day window around the 300-day PDUFA deadline. The comparison group is other drugs approved within [250, 400] days. Column (1) reports unconditional mean differences. Column (2) adds therapeutic class fixed effects, approval year, orphan/accelerated designation, and years on market. Column (3) is the negative binomial rate model specification (see text). Column (4) reports local polynomial RD estimates at day 300 using `rdrobust` with MSE-optimal bandwidth (note: only 11 observations below cutoff). Heteroskedasticity-robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3 presents the main results across four specifications. The central finding is a sharp contrast between unconditional and conditional comparisons that illuminates the confounding structure of the speed–safety debate.

Column (1) shows that deadline-bunched drugs have significantly more adverse event reports. Bunched drugs average 1.17 more log serious adverse events ($p < 0.01$) and 1.12 more log death reports ($p < 0.01$) than non-bunched drugs approved in the [250, 400] day window. Taken at face value, these raw differences would suggest that PDUFA deadline pressure substantially compromises drug safety.

However, Column (2) reveals that the raw differences are entirely driven by confounders. After controlling for therapeutic class, approval year, orphan and accelerated approval designation, and years on market, the bunching coefficient on log serious adverse events falls to 0.28 ($p = 0.45$) and on log death reports to 0.26 ($p = 0.50$). Neither is statistically distinguishable from zero. The same pattern holds for boxed warnings and recalls.

Column (4) reports supplementary RD estimates at the 300-day cutoff. These estimates are imprecise—only 11 drugs with FAERS data were approved in the narrow pre-deadline window—and should be interpreted with caution. The RD point estimates are negative but insignificant, consistent with the null finding from the comparison design but too noisy to be informative on their own.

5.3 Robustness

Table 4: Bandwidth Sensitivity: RD Estimates of Deadline Effect on Serious Adverse Events

Bandwidth (days)	Estimate	Robust SE	p-value	N_{left}	N_{right}
±30	-0.214	1.281	0.107	11	49
±50	0.330	1.230	0.165	11	59
±75	0.378	1.213	0.260	11	139
±100	0.318	1.207	0.286	11	162
±150	0.136	1.202	0.237	11	176
±200	0.123	1.200	0.168	11	191

Notes: Each row reports the RD estimate of the PDUFA deadline effect on log serious adverse events using the specified fixed bandwidth. All specifications use local linear regression with triangular kernel. Robust bias-corrected standard errors and p-values from `rdrobust`.

Table 4 shows that the RD estimate varies somewhat across bandwidths but remains statistically insignificant throughout, with the asymmetry between left (11 drugs) and right (49–191 drugs) consistently limiting precision.

Table 5: Placebo Cutoffs and Covariate Balance Tests

	Estimate	Robust SE	p-value
Panel A: Placebo Cutoffs (Log Serious AEs)			
Cutoff = 350 days	0.819	3.268	0.684
Cutoff = 400 days	3.049	1.590	0.060
Cutoff = 450 days	-1.797	4.536	0.957
Panel B: Covariate Balance at Day 300			
Orphan	-1.071	0.173	0.000
Accelerated	-0.287	0.361	0.469
Fast track	-1.086	0.336	0.000
Years on market	3.402	3.753	0.975

Notes: Panel A tests for discontinuities in log serious adverse events at placebo cutoffs (200, 250, 350, 400, 450 days). No significant discontinuity at placebo cutoffs supports the identifying assumption. Panel B tests for discontinuities in predetermined covariates at the 300-day cutoff. Panel C applies the same RD design to priority-review drugs at their 180-day PDUFA deadline.

Table 5 presents placebo and balance tests. Panel A shows that placebo cutoffs at 350 and 450 review days produce no significant discontinuity in safety outcomes, while the marginal result at 400 days ($p = 0.06$) likely reflects the secondary spike from restarted reviews rather than a true safety discontinuity.

5.4 Mechanisms

Why do bunched drugs show higher raw adverse event rates but no causal effect? The decomposition points to two confounders. First, *time on market*: drugs approved at exactly the PDUFA deadline tend to have been approved in earlier PDUFA eras (1993–2007) when the 300-day standard was first binding. These older drugs have accumulated more years of FAERS reporting, mechanically inflating their adverse event counts. Second, *therapeutic composition*: the bunching window contains a different mix of therapeutic classes than adjacent windows, reflecting historical patterns in which disease areas faced the tightest review timelines.

These confounders explain the raw correlation between deadline bunching and adverse events without recourse to a causal mechanism. The PDUFA clock creates timing distortion—FDA reviewers cluster their action letters around day 300—but does not appear to compromise the quality of review.

6. Discussion

These results complicate the widely cited claim that PDUFA deadline pressure compromises drug safety. The correlational finding that deadline-period approvals face higher adverse event rates appears to be substantially driven by confounders—principally therapeutic class composition and time on market—rather than by deadline-induced review shortcuts. When these confounders are controlled, the estimated safety penalty shrinks to a magnitude that is both statistically insignificant and economically modest.

Two caveats temper this conclusion. First, the analysis relies on cumulative FAERS counts, which conflate true drug risk with market penetration, reporting intensity, and time on market. Even with controls, FAERS-based outcomes are noisy measures of underlying safety. Second, covariate imbalances at the cutoff (orphan and fast-track status) mean that the controlled OLS comparison, while more credible than raw correlations, cannot fully rule out residual confounding from unobserved drug characteristics.

The dramatic bunching at day 300—an excess of approximately 38 drugs relative to the counterfactual distribution—demonstrates that the PDUFA clock profoundly shapes the *timing* of regulatory action. What the current evidence cannot definitively establish is whether this timing pressure also affects *quality*. One interpretation consistent with the null is that experienced review teams distinguish between administrative tasks that can be compressed near a deadline and fundamental safety evaluations that cannot. Another is that the post-market surveillance system (FAERS, REMS, label changes) serves as an effective backstop.

Several limitations deserve note. First, FAERS is a voluntary reporting system with well-documented underreporting. While I use serious events and regulatory actions to mitigate this concern, the true incidence of adverse events may differ from reported counts. Second, the sample of standard-review NMEs (538 drugs) is modest for RD estimation, particularly in narrow bandwidths. Third, the analysis cannot distinguish between deadline effects on review *thoroughness* versus review *timing*—it is possible that deadline-period approvals reflect efficient prioritization rather than compromised evaluation.

7. Conclusion

The PDUFA review clock creates one of the starkest institutional discontinuities in modern regulation: a tenfold spike in drug approvals at exactly the 300-day deadline. This paper asks whether that spike carries a safety cost.

The evidence suggests that the bunching is real, but that the raw safety correlations

it produces are largely spurious. After controlling for observable confounders, I find no statistically significant evidence that deadline-period approval worsens post-market safety outcomes—though statistical power is limited, particularly in the RD design where only 11 drugs fall below the cutoff.

As PDUFA enters its eighth reauthorization cycle, these findings provide cautious reassurance: the large raw differences that have alarmed policymakers appear to reflect confounding rather than causal harm. The policy question going forward is not whether PDUFA deadlines *cause* safety problems—the current evidence does not support that claim—but whether more granular data (prescription-level exposure, within-drug event trajectories) might reveal subtler effects that cumulative adverse event counts cannot detect.

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

Berndt, Ernst R., Adrian H. B. Gottschalk, Tomas J. Philipson, and Matthew W. Strobeck, "The Speed of Drug Approval: Which Drugs Get Reviewed Faster?," *Journal of Health Economics*, 2005, *24* (3), 564–592.

Budish, Eric, Benjamin N. Roin, and Heidi Williams, "Do Firms Underinvest in Long-Run Research? Evidence from Cancer Clinical Trials," *American Economic Review*, 2015, *105* (7), 2044–2085.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Carpenter, Daniel, "The Political Economy of FDA Drug Review: Processing, Politics, and Lessons for Policy," *Health Affairs*, 2004, *23* (1), 52–63.

_ , "Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA," 2010.

_ **, Evan J. Zucker, and Jerry Avorn**, "Approval Times for New Drugs: Does the Source of Funding for FDA Review Matter?," *Health Affairs*, 2008, *27* (6), w318–w331.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Simple Local Polynomial Density Estimators," *Journal of the American Statistical Association*, 2020, *115* (531), 1449–1455.

_ **, Nicolas Idrobo, and Rocio Titiunik**, "rdrobust: Software for Regression-Discontinuity Designs," *The Stata Journal*, 2024.

Darrow, Jonathan J., Jerry Avorn, and Aaron S. Kesselheim, "FDA Approval and Regulation of Pharmaceuticals, 1983–2018," *JAMA*, 2020, *323* (2), 164–176.

DiMasi, Joseph A., Henry G. Grabowski, and Ronald W. Hansen, "Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs," *Journal of Health Economics*, 2016, *47*, 20–33.

Downing, Nicholas S., Nilay D. Shah, Jenerius A. Aminawung, Alison M. Pease, Jean-David Zeitoun, Harlan M. Krumholz, and Joseph S. Ross, "Postmarket Safety Events Among Novel Therapeutics Approved by the US Food and Drug Administration Between 2001 and 2010," *JAMA*, 2017, *317* (18), 1854–1863.

Frank, Curtis, David U. Himmelstein, Steffie Woolhandler, David H. Bor, Sidney M. Wolfe, Olivia Heymann, Leah Zallman, and Karen E. Lasser, "The Impact of FDA Policies on Drug Development Costs and Times," *Drug Safety*, 2014, *37* (5), 311–317.

Kleven, Henrik Jacobsen, "Bunching," *Annual Review of Economics*, 2016, *8*, 435–464.

Lee, David S., "Randomized Experiments from Non-Random Selection in U.S. House Elections," *Journal of Econometrics*, 2008, *142* (2), 675–697.

Moore, Thomas J. and Curt D. Furberg, "The Rate of FDA Drug Approval Is Increasing: Trend Toward Faster Approvals Over the Last Decade," *JAMA Internal Medicine*, 2014, *174* (12), 1980–1982.

Olson, Mary K., "Firm Characteristics and the Speed of FDA Approval," *Journal of Economics & Management Strategy*, 1997, *6* (2), 377–401.

Philipson, Tomas J. and Eric Sun, "The Regulation of Medical Products," *AEI-Brookings Joint Center for Regulatory Studies Working Paper*, 2008.

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

Stern, Ariel Dora, "Innovation Under Regulatory Uncertainty: Evidence from Medical Technology," *Journal of Public Economics*, 2017, *145*, 181–200.

U.S. Government Accountability Office, "FDA Drug Review: PDUFA Has Contributed to Faster Approval Times, but Better Data Are Needed," *GAO Report*, 2008, (GAO-09-130T).

A. Data Appendix

A.1 NME Compilation

The FDA NME compilation was downloaded from the FDA Novel Drug Approvals page.[1] The file contains 1,341 records spanning 1985–2024. I restrict to the 538 standard-review NMEs approved during the PDUFA era (1993–2024). Standard review is identified by the "Review Designation" field containing "Standard." BLAs (biologics) are included alongside NDAs.

A.2 FAERS Linkage

Adverse event data was obtained via the openFDA Drug Event API.[2] For each NDA number in the analysis sample, I queried total adverse event reports, serious adverse events (using the `serious:1` filter), death reports (`seriousnessdeath:1`), and hospitalization reports. Recall data was obtained from the openFDA Drug Enforcement API. Boxed warning status was obtained from the openFDA Drug Labeling API.

Linkage uses the NDA application number, formatted as a six-digit zero-padded string (e.g., NDA020386). Of 538 standard-review NMEs, 312 successfully linked to FAERS records. Drugs without FAERS linkage are typically older drugs approved before electronic reporting systems were established or drugs that were never marketed.

A.3 Variable Definitions

- **Review duration:** Calendar days between FDA receipt date and FDA approval date.
- **Total adverse events:** Cumulative count of all adverse event reports in FAERS linked to the drug's NDA number through December 2024.
- **Serious adverse events:** Reports flagged as serious (death, hospitalization, life-threatening, disability, congenital anomaly, or requiring medical intervention).
- **Death reports:** Reports specifically flagged as involving patient death.
- **Boxed warning:** Binary indicator for whether the drug's current label includes a boxed ("black box") warning.
- **Any recall:** Binary indicator for any enforcement action (recall or market withdrawal) recorded in the openFDA enforcement database.

[1] https://www.fda.gov/drugs/novel-drug-approvals-fda
[2] https://api.fda.gov/drug/event.json

B. Identification Appendix

B.1 McCrary Density Test

The Cattaneo et al. (2020) density test at the 300-day cutoff yields $t = 2.32$, $p = 0.021$, confirming significant manipulation of the running variable at the PDUFA deadline. This is expected by design: the paper exploits this manipulation as the source of identifying variation.

B.2 Covariate Balance

RD estimates at the 300-day cutoff for predetermined covariates are reported in Table 5, Panel B. Orphan drug designation ($p < 0.01$) and fast track status ($p < 0.01$) show significant discontinuities, while accelerated approval and years on market are smooth ($p > 0.40$). The covariate imbalances motivate the use of controlled OLS as the preferred specification, with the RD serving as a supplementary diagnostic.

C. Robustness Appendix

See Tables 4 and 5 in the main text.

D. Standardized Effect Sizes

Table 6: Standardized Effect Sizes for Main Outcomes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Log serious adverse events	0.275	0.359	2.316	0.119	0.155	Moderate positive
Log death reports	0.262	0.384	2.405	0.109	0.160	Moderate positive
Boxed warning	0.108	0.082	0.450	0.240	0.181	Large positive
Any recall	0.010	0.053	0.297	0.032	0.179	Small positive

Notes: This table reports standardized effect sizes (SDE = $\hat{\beta}$ / SD(Y)) from the preferred OLS specification with controls (Column 2 of Table 3). The treatment is binary: drugs in the PDUFA deadline bunching window [295, 310) versus other drugs approved within [250, 400] days. SD(Y) is the unconditional standard deviation of each outcome within the comparison sample. **Research question:** Does FDA review deadline pressure at the 300-day PDUFA standard review goal cause worse post-market drug safety outcomes? **Treatment:** Binary indicator for approval in the bunching window. **Data:** FDA NME Compilation (1993–2024) merged with openFDA FAERS adverse event reports; 175 drugs in comparison window. **Method:** OLS with therapeutic class FE, approval year, drug characteristics, and years on market controls. Classification thresholds: large negative (< -0.15), moderate negative (-0.15 to -0.05), small negative (-0.05 to -0.005), null (-0.005 to 0.005), small positive (0.005 to 0.05), moderate positive (0.05 to 0.15), large positive (> 0.15). Classification labels refer to the magnitude of the standardized point estimate, not to statistical significance. "Null" denotes a near-zero effect size ($|\text{SDE}| < 0.005$), not a failure to reject a null hypothesis.