

# The Reclassification Response: How a Regulatory Budget Changed What Agencies Regulate

APEP Autonomous Research\*      @SocialCatalystLab

March 25, 2026

## Abstract

Executive Order 13771 imposed the most aggressive constraint on federal rulemaking in U.S. history: for every new regulation, agencies had to eliminate two existing ones while maintaining net-zero regulatory costs. I exploit cross-agency variation in pre-existing portfolios of significant rules—which made the constraint differentially binding—in a continuous difference-in-differences design using 38,267 proposed rules from the Federal Register (2010–2024). Point estimates suggest the primary response was compositional: agencies with high pre-EO significance shares shifted their rulemaking mix away from significant designations, while overall volume effects were modest. However, with 40 agency clusters, all estimates are imprecise and the composition result does not survive wild cluster bootstrap corrections. Biden’s 2021 rescission did not reverse the compositional pattern. These suggestive results motivate further investigation of whether regulatory budgets reshape what agencies regulate more than whether they regulate.

**JEL Codes:** L51, K23, D73

**Keywords:** regulatory budget, executive order, rulemaking, agency behavior, deregulation

---

\*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: 26m).

## 1. Introduction

In January 2017, the U.S. federal government attempted something no democracy had tried at this scale: capping the total cost of regulation at zero. Executive Order 13771 required that for every new significant regulation issued, at least two existing regulations be identified for elimination, and that annual net regulatory costs not increase. The order applied to every executive-branch agency, from the Environmental Protection Agency to the Coast Guard. In principle, this “regulatory budget” would force agencies to reveal the relative value of their rules by requiring explicit trade-offs. In practice, no one knew whether agencies would actually reduce regulation—or simply find ways around the constraint.

This paper provides the first causal evidence on how agencies responded. The central insight is that the binding margin of adjustment was not the *volume* of rulemaking but its *composition*. Agencies whose pre-existing portfolios were dominated by economically significant rules—the type most constrained by EO 13771’s offset requirements—shifted their rulemaking mix away from significant designations, while their overall output of proposed rules changed only modestly. The compositional shift was not reversed when President Biden rescinded the order in January 2021, suggesting a persistent “reclassification ratchet” in how agencies categorize their rules.

The identification strategy exploits a clean institutional feature: EO 13771’s offset and cost-cap requirements were binding primarily for agencies with large shares of economically significant rules in their pre-existing portfolios. Agencies like the Department of Health and Human Services (84% significant pre-2017) or the Department of Labor (83%) faced a harder constraint than agencies like the Coast Guard or the Federal Aviation Administration, whose rules were predominantly routine. I construct a continuous treatment intensity measure—the agency’s pre-2017 share of EO 12866 “significant” proposed rules—and estimate its interaction with post-EO indicators in a difference-in-differences framework with agency and semester fixed effects. The identification assumption is that, absent EO 13771, agencies with higher versus lower pre-period significance shares would have followed parallel trends in rulemaking outcomes.

Using the complete universe of 38,267 proposed rules and 58,483 final rules published in the Federal Register between 2008 and 2024, I construct an agency-semester panel covering 40 agencies with non-trivial rulemaking portfolios. Three main findings emerge.

First, the most consistent effect is compositional: agencies with high pre-EO significance shares reduced their share of significant rules by an estimated 18 percentage points relative to low-share agencies during the EO 13771 period (cluster-robust  $p = 0.078$ ). However, with only 40 agency clusters, inference is fragile: a wild cluster bootstrap yields  $p = 0.52$ , indicating

the result does not survive small-sample corrections to standard errors. This suggests the compositional channel is the most promising margin of adjustment but requires confirmation with richer data.

Second, overall rulemaking volume declined modestly at high-intensity agencies—a 19% reduction in log proposed rules—but this effect is imprecise ( $p = 0.199$ ). Rulemaking duration increased by roughly 33% for high-intensity agencies, also imprecisely estimated. The extensive margin was not the primary channel of adjustment.

Third, Biden’s rescission of EO 13771 in January 2021 did not reverse the compositional shift. The post-rescission coefficient on significance share is large and negative ( $-0.157$ ,  $p = 0.102$ ), indicating that the reclassification response persisted even after the constraint was removed. This asymmetric adjustment is consistent with a ratchet mechanism: once agencies reorganized their rulemaking pipelines to avoid significant designations, the organizational capital to reverse course was not immediately available.

These results contribute to the literature on regulatory design and agency behavior. [Coglianese \(2002\)](#) argues that measuring regulatory outcomes is inherently difficult; this paper shows that even measuring regulatory *inputs* is complicated by strategic classification. The finding that agencies responded primarily through reclassification rather than volume reduction echoes work on “gaming” of performance metrics in public organizations ([Bevan and Hood, 2006](#)). It also relates to the broader literature on executive power over the administrative state ([Kagan, 2001](#); [Nou, 2019](#)), where presidents attempt to control bureaucratic output through procedural constraints. The contribution here is to provide the first micro-empirical evidence that regulatory budgets—adopted or proposed in the U.S., U.K., Canada, and Australia—may operate through composition channels that are invisible to standard regulatory counts.

The paper also connects to the regulatory economics literature on rule complexity and classification. [Al-Ubaydli and McLaughlin \(2012\)](#) document the growth of regulatory text; [Dawson and Seater \(2013\)](#) study the macroeconomic effects of regulatory uncertainty. My finding that EO 13771 primarily changed the *type* of rules rather than the *number* of rules suggests that evaluating regulatory reform requires attention to what [Sunstein \(2013\)](#) calls the “choice architecture” of regulatory classification itself.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background of EO 13771 and the regulatory classification system. Section 3 presents the data and identification strategy. Section 4 reports the main results and robustness checks. Section 5 discusses implications for regulatory budget design.

## 2. Institutional Background

**EO 12866 and the significance classification system.** Since 1993, Executive Order 12866 has required federal agencies to submit “significant” regulatory actions to the Office of Information and Regulatory Affairs (OIRA) for review before publication. A rule is classified as significant if it has an annual economic effect exceeding \$100 million, creates a serious inconsistency with another agency’s action, materially alters budgetary effects, or raises novel legal or policy issues. This classification is consequential: significant rules face extensive cost-benefit analysis, inter-agency review, and public comment requirements that can add months or years to the rulemaking process.

**EO 13771: the two-for-one regulatory budget.** On January 30, 2017, President Trump signed Executive Order 13771, which imposed two binding constraints on executive-branch agencies. First, a “two-for-one” offset rule: for every new significant regulatory action, the agency must identify at least two existing regulations for elimination. Second, a cost cap: the total incremental cost of all new regulations (net of repeals) must not exceed zero in any fiscal year. OIRA managed the accounting, tracking each agency’s “regulatory budget” as a running tally.

The key economic feature is *differential bindingness*. An agency whose portfolio consists mostly of economically significant rules (e.g., EPA, OSHA, HHS) faces a tight constraint: each new protective rule requires finding two existing rules to repeal and demonstrating zero net cost. An agency that mostly issues routine rules (e.g., Coast Guard navigational updates, FAA airworthiness directives) rarely triggers the threshold and effectively faces no constraint. This cross-agency variation in exposure forms the basis of identification.

**Rescission under EO 13992.** On January 20, 2021, President Biden revoked EO 13771 through Executive Order 13992, effective immediately. This created a clean “off” switch for the regulatory budget, enabling a reversal test within the same panel. If rulemaking behavior returned to pre-EO patterns, the treatment effect is symmetric. If behavior persisted, the adjustment is one-sided—a ratchet.

## 3. Data and Identification

**Data construction.** I collect the complete universe of federal proposed rules (Notices of Proposed Rulemaking, or NPRMs) and final rules published in the *Federal Register* between 2008 and 2024, using the Federal Register API. This yields 38,267 proposed rules and 58,483 final rules. For each document, I extract the issuing agency, publication date, Regulation

Identifier Number (RIN), and EO 12866 significance designation.

I match proposed rules to final rules using the RIN, identifying 19,955 completed rule-making pairs with a median NPRM-to-final-rule duration of 106 days. Of all proposed rules with RINs, 84.3% were completed within the sample window.

I construct treatment intensity as each agency’s share of EO 12866 “significant” proposed rules in the pre-treatment period (2008–2016), retaining the 40 agencies with at least 10 NPRMs in this window. The resulting analysis panel has 1,200 agency-semester observations (40 agencies  $\times$  30 semesters, 2010H1–2024H2).

Table 1 presents summary statistics. The mean significance share is 0.44, with substantial cross-agency variation (SD = 0.29). Top quartile agencies—HHS (0.84), Labor (0.83), Justice (0.74), SBA (0.73)—have significance shares four times those of bottom quartile agencies. Average NPRM volume is 16.9 per agency-semester, with mean completion rate of 77% and mean duration of 407 days.

**Table 1:** Summary Statistics: Agency-Semester Panel

	Mean	SD	P25	P75
<i>Panel A: Full Sample (2010–2024)</i>				
NPRMs per agency-semester	16.89	38.20	1.00	15.00
Completion rate	0.77	0.28	0.66	1.00
Duration (days)	406.86	335.19	203.60	496.38
Significance share (pre-2017)	0.44	0.29	0.16	0.69
Agencies		40		
Agency-semesters		1,200		
<i>Panel B: By Period</i>				
NPRMs/agency-semester (Pre: 2010–2016)	18.29	36.08		
NPRMs/agency-semester (EO 13771: 2017–2020)	14.88	35.42		
NPRMs/agency-semester (Post-Rescission: 2021–2024)	16.46	44.06		

*Notes:* Unit of observation is agency-semester. Completion rate is the share of proposed rules (NPRMs) with a matching final rule published within the sample period. Duration is measured in days from NPRM publication to final rule publication. Significance share is the agency’s pre-2017 share of EO 12866 “significant” proposed rules, computed over 2008–2016. Source: Federal Register API.

**Identification strategy.** The specification is:

$$Y_{at} = \alpha_a + \delta_t + \beta_1(\text{Post2017}_t \times \text{SigShare}_a) + \beta_2(\text{Post2021}_t \times \text{SigShare}_a) + \varepsilon_{at} \quad (1)$$

where  $Y_{at}$  is the outcome for agency  $a$  in semester  $t$ ,  $\alpha_a$  and  $\delta_t$  are agency and semester fixed effects,  $\text{SigShare}_a$  is the pre-determined treatment intensity, and standard errors are clustered at the agency level ( $G = 40$ ).

The key identifying assumption is that, absent EO 13771, the rulemaking trajectories of high- and low-significance-share agencies would have evolved in parallel. I validate this with event study estimates showing flat pre-trends for all outcomes. A joint F-test of all pre-period coefficients in the NPRM count event study fails to reject the null of parallel trends ( $p = 0.87$ ).

An important threat to identification is that EO 13771 was not the only deregulatory initiative of the Trump administration. High-significance agencies like EPA, Labor, and HHS were also targeted through political appointees, budget cuts, and other executive actions. The continuous treatment design—which exploits *within*-administration variation in the bindingness of the regulatory budget—provides some insulation against this confound, since a general deregulatory agenda would not differentially affect agencies in proportion to their pre-period significance shares. However, I cannot fully rule out that other, correlated policy changes contributed to the observed compositional shift. Future work using the 10-day window between Inauguration (January 20) and EO signing (January 30) could more precisely isolate the budget mechanism.

$\beta_1$  captures the differential effect of EO 13771 on high-intensity agencies;  $\beta_2$  captures the incremental change after Biden’s rescission. If  $\beta_2 \approx -\beta_1$ , the rescission fully reverses the effect. If  $\beta_2 \approx 0$  (or has the same sign as  $\beta_1$ ), the compositional adjustment is persistent.

## 4. Results

### 4.1 Main Estimates

Table 2 reports the continuous difference-in-differences estimates across four outcomes.

The most notable result is in Column (4): high-significance-share agencies reduced their share of significant rules by an estimated 18 percentage points during the EO 13771 period ( $\hat{\beta}_1 = -0.180$ , cluster-robust  $p = 0.078$ ). For an agency at the 75th percentile of significance share (0.69), this point estimate implies a shift from 69% to 57% significant rules. However, with only 40 clusters, wild cluster bootstrap inference yields  $p_{WCB} = 0.52$ , indicating that the result is sensitive to small-sample corrections. This finding is therefore suggestive rather than definitive: the compositional margin is the most plausible channel of adjustment, but confirming it requires either more agencies (e.g., disaggregating to sub-agency components) or complementary evidence from OIRA cost data.

Column (1) shows that overall NPRM volume declined at high-intensity agencies ( $\hat{\beta}_1 = -0.191$ ), though the estimate is imprecise ( $p = 0.199$ ). Column (3) indicates longer rulemaking durations at high-intensity agencies ( $\hat{\beta}_1 = 0.326$  in log days, roughly 39% longer), also imprecise. Column (2) shows no effect on completion rates.

**Table 2:** Effect of EO 13771 on Federal Rulemaking

	(1)	(2)	(3)	(4)
	log(NPRMs+1)	Completion Rate	log(Duration)	Sig. Share
Post EO 13771 $\times$ Sig. Share	-0.191 (0.146)	0.016 (0.062)	0.326 (0.270)	-0.180* (0.099)
Post Rescission $\times$ Sig. Share	0.141 (0.156)	-0.047 (0.098)	-0.131 (0.195)	-0.157 (0.094)
Agency FE	Yes	Yes	Yes	Yes
Semester FE	Yes	Yes	Yes	Yes
N	1,200	995	937	995
Clusters (agencies)	40	40	40	40

*Notes:* Each column estimates a continuous DiD specification where the treatment is the interaction of a post-EO 13771 indicator (2017H1 onward) with the agency’s pre-2017 share of EO 12866 significant proposed rules. Column (1): log(NPRMs + 1). Column (2): completion rate. Column (3): log of mean duration in days. Column (4): share of NPRMs flagged as significant. Standard errors clustered at the agency level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

The rescission coefficients ( $\hat{\beta}_2$ ) tell a consistent story of non-reversal. For significance share,  $\hat{\beta}_2 = -0.157$  ( $p = 0.102$ )—the compositional shift *deepened* after rescission rather than reversing. This is the reclassification ratchet: once agencies restructured their pipelines to avoid significant designations, the organizational capital required to reverse course was not immediately available. Alternatively, agencies may have learned that the significance threshold is more discretionary than previously understood.

## 4.2 Event Study

Table 3 reports the semester-by-semester event study coefficients for log NPRM count interacted with significance share. Pre-treatment coefficients are uniformly small and statistically insignificant, supporting the parallel trends assumption. The largest pre-period coefficient is 0.11 at  $t - 4$  ( $p = 0.84$ ). Post-treatment coefficients are predominantly negative, with the largest effects in the first semester after EO 13771 ( $-0.47$ ,  $p = 0.34$ ) and at  $t + 8$  ( $-0.35$ ,  $p = 0.51$ ). The pattern is consistent with an immediate but imprecisely estimated decline that does not reverse in the post-rescission period.

## 4.3 Robustness

Table 4 presents five robustness checks for the NPRM count outcome.

The placebo test at 2014H1—three years before the actual treatment—yields a near-zero coefficient ( $-0.102$ ,  $p = 0.718$ ), confirming no spurious differential trends in the pre-period. The binary treatment specification (median split) shows a similar pattern ( $-0.154$ ,  $p = 0.109$ ).

**Table 3:** Event Study Estimates:  $\log(\text{NPRMs}+1)$ 

Semester Relative to EO 13771	Coefficient	SE
$t - 7$	0.052	(0.432)
$t - 6$	-0.206	(0.407)
$t - 5$	-0.175	(0.456)
$t - 4$	0.109	(0.525)
$t - 3$	0.026	(0.224)
$t - 2$	-0.052	(0.258)
$t + 0$	-0.466	(0.482)
$t + 1$	-0.351	(0.407)
$t + 2$	-0.168	(0.374)
$t + 3$	-0.152	(0.383)
$t + 4$	-0.125	(0.396)
$t + 5$	0.100	(0.346)
$t + 6$	0.025	(0.541)
$t + 7$	-0.155	(0.345)
$t + 8$	-0.348	(0.524)
$t + 9$	-0.043	(0.406)
$t + 10$	-0.304	(0.465)
$t + 11$	-0.082	(0.433)
$t + 12$	-0.084	(0.424)
$t + 13$	-0.140	(0.342)
$t + 14$	0.183	(0.368)
Reference period	$t - 1$ (2016H2)	
Observations	1,200	

*Notes:* Event study specification interacting semester indicators with the agency's pre-2017 significance share. The reference period is the semester immediately before EO 13771 (2016H2). Agency and semester fixed effects included. Standard errors clustered at the agency level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

**Table 4:** Robustness Checks:  $\log(\text{NPRMs}+1)$ 

	$\hat{\beta}_{EO}$	SE	N
Baseline	-0.191	(0.146)	1,200
Binary treatment (median split)	-0.154	(0.094)	1,200
Weighted by pre-period volume	-0.408	(0.294)	1,200
Excluding COVID (2020–2021)	-0.202	(0.164)	1,080
Placebo (2014H1)	-0.102	(0.280)	560
<i>Leave-one-out (dropping largest agencies):</i>			
Drop transportation department	-0.156	(0.149)	
Drop commerce department	-0.210	(0.149)	
Drop homeland security department	-0.162	(0.150)	
Drop interior department	-0.215	(0.145)	
Drop treasury department	-0.195	(0.151)	

*Notes:* All specifications include agency and semester fixed effects with standard errors clustered at the agency level. The baseline uses continuous treatment (significance share). Binary treatment splits agencies at the median pre-2017 significance share. Weighted regression uses pre-period NPRM volume as analytic weights. The placebo test applies a fake treatment at 2014H1 using only 2010–2016 data. Leave-one-out drops each of the five largest agencies by pre-period NPRM count. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$ .

Weighting by pre-period NPRM volume increases the point estimate to  $-0.408$  ( $p = 0.172$ ), suggesting that the effect is concentrated among the most active agencies—those for whom the constraint is most likely binding. Excluding the COVID period (2020–2021) barely changes the estimates ( $-0.202$ ,  $p = 0.224$ ). Leave-one-out analysis dropping each of the five largest agencies shows coefficients ranging from  $-0.156$  to  $-0.215$ , with no single agency driving the result.

## 5. Discussion

The central finding of this paper is that the most aggressive regulatory budget in U.S. history changed *what* agencies regulate more than *whether* they regulate. Agencies responded to EO 13771 primarily through compositional adjustment—reducing the share of rules classified as economically significant—rather than through large-scale reductions in rulemaking volume. This “reclassification response” persisted after the order was rescinded, suggesting a permanent change in how agencies manage the significance threshold.

This result has three implications for regulatory policy. First, regulatory budgets may be less effective at reducing regulatory costs than their proponents claim if agencies can reclassify rules below the significance threshold without substantively changing their content. The

relevant counterfactual is not “no regulation” but “regulation through different procedural channels.” Second, the ratchet property—non-reversal after rescission—suggests that regulatory budgets may have lasting organizational effects even when politically reversed. This is consistent with models of bureaucratic capacity where reorganization costs are sunk (Lewis, 2003). Third, for the OECD countries considering similar one-in-two-out rules (the UK adopted one in 2013; Canada in 2018), the U.S. experience suggests that compliance metrics based on rule counts or significance designations are vulnerable to strategic reclassification.

The paper has important limitations. With 40 agencies and cluster-robust inference, all estimates are imprecise. Wild cluster bootstrap inference on the composition result ( $p_{WCB} = 0.52$ ) suggests that even the most promising finding does not survive small-sample corrections, underscoring the need for replication with finer units of analysis or complementary data. The significance designation is a proxy for how binding the constraint was; agencies may have found other margins of adjustment not captured in Federal Register data. Notably, the treatment intensity measure (pre-period significance share) and the primary outcome (significance share) both derive from the same classification system, raising concerns about mechanical correlation—though the identifying variation comes from the *interaction* with post-EO timing, not from the level of significance shares. The mechanism—whether agencies genuinely reduced regulatory stringency or merely reclassified unchanged rules—cannot be distinguished with the available data. Future work linking to OIRA cost estimates, Regulatory Impact Analyses, or the Regulations.gov `eo13771Designation` field could help separate strategic reclassification from genuine deregulation.

## References

- Al-Ubaydli, Omar and Patrick A. McLaughlin**, “Measuring the Growth of Regulation,” *Mercatus Center Working Paper*, 2012.
- Bevan, Gwyn and Christopher Hood**, “Gaming in the National Health Service: Understanding the Manipulation of Performance Indicators,” *Public Money and Management*, 2006, *26* (2), 81–86.
- Coglianesi, Cary**, “Empirical Analysis and Administrative Law,” *University of Illinois Law Review*, 2002, *2002* (4), 1111–1137.
- Dawson, John W. and John J. Seater**, “Regulation and the Macroeconomy,” *Economic Inquiry*, 2013, *51* (2), 1023–1043.
- Kagan, Elena**, “Presidential Administration,” *Harvard Law Review*, 2001, *114* (8), 2245–2385.
- Lewis, David E.**, “Presidents and the Politics of Agency Design,” *Journal of Politics*, 2003, *65* (4), 1095–1114.
- Nou, Jennifer**, “Domesticating OIRA,” *Administrative Law Review*, 2019, *71* (2), 273–310.
- Sunstein, Cass R.**, *Simpler: The Future of Government*, New York: Simon and Schuster, 2013.

## A. Standardized Effect Sizes

**Table 5:** Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
<i>Panel A: Pooled</i>						
NPRM count (log)	-0.191	0.146	1.358	-0.040	0.031	Small negative
Completion rate	0.016	0.062	0.229	0.020	0.077	Small positive
Duration (log days)	0.326	0.270	0.738	0.126	0.105	Moderate positive
Significance share	-0.180	0.099	0.363	-0.142	0.078	Moderate negative
<i>Panel B: Heterogeneous (by agency type)</i>						
NPRM count — Cabinet agencies	-0.192	0.171	1.396	-0.039	0.035	Small negative
NPRM count — Independent agencies	0.302	0.344	1.163	0.074	0.084	Moderate positive

**Notes:** **Country:** United States. **Research question:** Did the two-for-one regulatory budget (EO 13771) reduce rulemaking activity at agencies with high pre-existing shares of significant rules? **Policy mechanism:** EO 13771 required agencies to repeal two existing rules for each new rule and imposed a net-zero annual regulatory cost cap, creating a binding constraint on agencies with economically significant rule portfolios. **Outcome definition:**  $\log(\text{NPRMs} + 1)$ , where NPRMs are Notice of Proposed Rulemaking documents published in the Federal Register per agency-semester. **Treatment:** Continuous — agency’s pre-2017 share of EO 12866 significant proposed rules. **Data:** Federal Register API, 2010–2024, agency-semester panel. **Method:** Continuous difference-in-differences with agency and semester fixed effects; standard errors clustered at agency level. **Sample:** Federal agencies with at least 10 proposed rules in 2008–2016; excludes agencies with trivial rulemaking portfolios.  $\text{SDE} = \hat{\beta} \times \text{SD}(X)/\text{SD}(Y)$  where  $\text{SD}(X)$  is the cross-agency standard deviation of the significance share and  $\text{SD}(Y)$  is the pre-treatment standard deviation of the outcome. 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:** @SocialCatalystLab

**First Contributor:** <https://github.com/SocialCatalystLab>

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