

Contagious Deregulation*

Jakub Hajda

Joseph Kalmenovitz

Billy Y. Xu

September 3, 2026

Abstract

We show that deregulation propagates across firms through a shared intermediary. A 2015 reform raised the audit threshold for federal award recipients, generating a negative demand shock for auditors serving clients below the threshold. These auditors relaxed scrutiny of the larger clients they retained, issuing more favorable opinions without any improvement in documented compliance. The effect increases with the number of clients lost, concentrates in previously stable auditor-client relationships, and weakens under external monitoring. Post-reform, each year federal awards worth \$6.2 billion receive audits that would otherwise have been negative. A structural model quantifies deregulation reforms that preserve monitoring quality.

Keywords: Economics of Regulation, Auditor Incentives, Federal Awards, Nonprofit Finance

JEL Codes: H83, L51, M48, K2

*Joseph Kalmenovitz is with the University of Rochester (jkalmeno@simon.rochester.edu, corresponding author); Jakub Hajda is with HEC Montréal (jakub.hajda@hec.ca); Billy Xu is with the University of Rochester (bxu22@simon.rochester.edu). For helpful comments we thank Marco Ceccarelli, Qingkai Dong, Joseph Gerakos, Zongbo Huang, Jonathan Karpoff, Narayana Kocherlakota, Max Miller, Vivek Pandey, Thomas Wollmann, Yufeng Wu, David Yermack, and Ben (Miao) Zhang; and conference and seminar participants at the American Law and Economics Association annual meeting, Chinese University of Hong Kong, City University of Hong Kong Finance conference, Conference on Auditing and Capital Markets, EDHEC Business School, ESSEC Business School, European Finance Association annual meeting, Federal Reserve, Haskell & White Corporate Reporting and Governance conference, HKUST Finance Symposium, National University of Singapore, NBER Economic Analysis of Regulation conference, New York University, Rotman School of Business, UNC-Duke Corporate Finance conference, University of Connecticut Finance conference, University of Florida, and University of Texas Austin Forensic Finance conference.

Introduction

Financial regulation often hinges on thresholds to determine which activities are subject to regulatory oversight. For example, prudential supervision tightens as a bank’s assets grow, and disclosure requirements expand as a firm’s public float increases. When policymakers raise such a threshold, they exempt some firms from regulatory requirements, and a large literature studies how these newly deregulated firms adjust (Iliev, 2010; Alvero, Ando, and Xiao, 2023; Ewens, Xiao, and Xu, 2024). We ask a different question: what happens to the firms that remain above the threshold? In our setting, the answer runs through the compliance professionals who serve entities on both sides of the cutoff. We show that raising a threshold contracts demand for compliance services, motivating these professionals to weaken oversight of the still-regulated population. We refer to this mechanism as the contagion effect of deregulation: a targeted deregulatory act propagating through the intermediary.

We document the contagion effect in the market for federal award auditing. Federal awards are a major source of capital for nonprofit organizations, including universities and hospitals. Over the 25 years in our sample, annual disbursements exceed \$390 billion – roughly 2% of U.S. GDP. To ensure that these funds are spent properly, recipients above a threshold must hire an independent auditor to conduct a comprehensive review, known as “Single Audit.” A 2015 reform raised the mandatory audit threshold, exempting roughly 14% of previously audited recipients. We show that auditors exposed to this demand contraction become more lenient toward their remaining, still-regulated clients. A structural model isolates the cost components driving the auditor’s decision and identifies deregulation designs that preserve monitoring quality.

Our results are relevant beyond the setting of federal award auditing. The contagion effect can arise whenever a deregulatory reform contracts demand for the services of a private intermediary that depends financially on the entities it serves and faces little consequence for softening its assessments. Such a demand shock can follow a threshold increase, a common deregulatory tool given that threshold-based regulation is ubiquitous in the Code of Federal

Regulations. But more broadly, any deregulation can reduce the demand for compliance services – through exemptions, scope reductions, or other means – and similar incentive distortions may arise. Thus, the standard evaluation of deregulatory reforms, which weighs the direct costs and benefits for deregulated entities, may miss the indirect contagion effect that operates through the market for compliance services.

Our analysis rests on a comprehensive new dataset covering the universe of Single Audit reports from 1998 to 2022. We construct this dataset from the Federal Audit Clearinghouse, which collects all Single Audit submissions to the federal government. The raw data are released in two vintages with distinct file structures, a historical archive (1998–2015) and a redesigned format (2016 onward), that we harmonize into a single panel.¹ We link auditors and award recipients across years by standardizing firm names, recovering missing identifiers with a fuzzy string-matching algorithm, and manually resolving ambiguous cases. The resulting dataset covers 993,000 auditor–recipient annual observations spanning \$9.8 trillion in awards over 25 years, with 95,000 unique recipients monitored by 27,000 unique auditors. The key variables are the identities of the auditor and the client, the size of the federal award, and the audit outcome – whether the auditor issued a positive (“unqualified”) or negative (“qualified” or “adverse”) opinion. For additional analyses we use more granular documentation of concrete deficiencies.

The reform triggered a large contraction in demand for Single Audits. To see that, we analyzed all 664 public comments filed on the reform. The audit profession expected the higher threshold to sharply reduce the number of Single Audits. The government’s own estimates agreed on the scale, projecting that roughly 5,000 entities would lose the requirement and recording some 3,400 fewer Single Audits per year afterward. Crucially, this contraction did not fall evenly across auditors: exposure depended on how many of an auditor’s clients happened to sit just below the newly raised threshold. We exploit this key feature in a difference-in-differences framework. Treated auditors are those who had at least

¹A data construction kit with code and documentation is available on [the authors’ website](#), as well as a list of regulatory thresholds to facilitate future research.

one client below \$750,000 before the reform; when those clients were exempted, they ceased to demand auditing services. Control auditors had no client below \$750,000 but at least one between \$1,000,000 and \$1,500,000, ensuring comparability.²

We find that treated auditors became more reluctant to issue negative audits after the reform. The treatment effect is economically large: in our tightest specification, we estimate a 1.9 percentage point decline in negative audits, corresponding to 23% of the pre-reform mean. Aggregating outcomes to the auditor level, we find that treated auditors are 11% less likely to issue any negative audit and reduce their overall negative audit rate by 15% following the reform, relative to the corresponding pre-reform means. To gauge the magnitude, we apply the average treatment effect to federal awards audited by treated auditors in the post-reform period. This calculation implies that \$6.2 billion per year in awards receive audits that otherwise would have been negative.

Our core specification saturates the model with time and pair fixed effects, absorbing common shocks and persistent differences across clients, auditors, and auditor-client relationships.³ We further show that treatment and control groups demonstrate parallel trends in audit outcomes before the reform, breaking sharply only at the reform date. Moreover, interacting a rich set of pre-reform auditor and client characteristics with the post-reform indicator yields similar estimates, indicating that the results are not driven by differential trends in other observables. Finally, a placebo test using only an early period that precedes our estimation window finds no difference in audit outcomes between clients above and below \$750,000, further tying the effect to the reform itself rather than to underlying differences between these clients.

The decline in negative audits is consistent with two broad channels: a change in auditor behavior and a change in client compliance. Under the first, the reform contracts demand for audit services and raises the value of each surviving engagement, so auditors soften

²We obtain similar results using a range of upper bounds for the control group.

³In addition to raising the threshold, the reform revised several elements of the auditing process. Analyzing these changes, we show that they apply equally to treated and control auditors and are thus absorbed by time fixed effects.

their assessments to retain clients who might otherwise exit or switch. Under the second, clients of treated auditors become more compliant relative to clients of control auditors – for instance, because the reform simplified regulatory requirements in ways that disproportionately benefited entities served by treated auditors.

The evidence consistently supports the auditor-leniency channel, and it comes in three groups. We show that leniency pays for the auditor; that the decline in negative audits is concentrated where auditor leniency would predict; and that treated clients did not simply become more compliant. A cost-benefit calculation shows why leniency is a rational response to the demand shock: client relationships in the Single Audit market are persistent and command a substantial premium over other auditing services, yet a negative opinion raises the probability that the client relationship ends. At the same time, accountability for a deficient audit is minimal. Searching a decade of enforcement records from the Department of Justice, state accountancy boards, and American Institute of Certified Public Accountants (AICPA), we find a single federal quality-control review that ended in a professional referral, a handful of state-board disciplinary notices, and no criminal proceeding against any Single Audit auditor. The cost-benefit tradeoff is sharper for auditors serving clients around the new threshold, which we find to be small, local, specialized practices: the median auditor holds just five Single Audit clients and, on average, draws more than a third of its fee revenue from clients below the new threshold. Losing even one or two engagements is therefore material, and the clients that remain become all the more valuable to retain. The auditor thus has much to gain from a lax opinion and little to fear from it.

The aggregate stakes are large. Before the reform, treated auditors audited only about \$4.4 billion in the small, at-risk awards that deregulation would exempt, but more than seventy times as much in awards of the larger clients that remained subject to audit. The leniency the reform induced falls on that far larger retained pool: as mentioned, applying our tightest estimate implies that \$6.2 billion in federal awards each year receive a favorable opinion they would not have received absent the reform. The auditors who perform Single

Audits hold large private stakes in retaining clients yet face little accountability for leniency; in such an environment, small and idiosyncratic demand shocks aggregate across thousands of exposed auditors into a broad erosion of oversight over public funds.

In addition to the evidence above, we document three patterns that are collectively more consistent with auditor leniency than with improved client compliance. First, the effect scales with auditor exposure: auditors with a larger share of pre-reform engagements below \$750,000 exhibit a larger decline in negative opinions. Second, it is largest in the most stable relationships and least competitive markets (measured in the pre-reform period) – those the reform most disrupted. Third, it is muted where clients are externally monitored: merging the Single Audit data with IRS Form 990 filings, we find no decline in negative audits among tax-exempt organizations that file the form, while non-filers exhibit a large and significant decline.⁴ Improved compliance should not vary with an auditor’s exposure to the reform, the stability of its relationships, or whether its clients file Form 990.

A final set of results rules out the competing explanation that treated clients simply became more compliant. Most directly, treated auditors document more internal-control deficiencies after the reform even as their opinions grow more favorable – a divergence improved compliance cannot produce. We also find no evidence that treated clients have a stronger incentive or ability to comply. The post-reform compliance manual is longer and more densely cross-referenced, rather than simpler, and it applies identically to treated and control auditors. Moreover, the data we collected from federal, state, and professional enforcement records show no post-reform increase in enforcement intensity. Finally, the effect is not due to compositional changes in the pool of auditors or clients: there is essentially no differential sorting of clients around the threshold, and the estimate holds in a sample that fixes the set of auditors and clients across the reform.

In sum, the evidence points to a demand-driven shift in auditor incentives. The decline in negative audits is concentrated among the most exposed auditors, strongest where the reform

⁴Donors to tax-exempt organizations pay particular attention to Form 990 disclosures, which must be publicly available per Section 6104 of the Internal Revenue Code (Harris, Petrovits, and Yetman, 2015).

was most disruptive, and absent where external monitoring constrains auditor behavior. The shift is not accompanied by any reduction in underlying documented deficiencies, nor by differential simplification of compliance requirements, more intense enforcement, or compositional changes in the underlying sample. These patterns collectively indicate a shift in auditor incentives induced by a contraction in demand for compliance services.

The reduced-form analysis does not isolate the underlying cost components that shape the auditor’s decision. Those are central for evaluating policy interventions that could mitigate the contagion effect of deregulation. To do so, we develop a parsimonious model where the auditor weighs the cost of a negative opinion against a private information shock – an unobserved signal about the client’s true compliance. The model’s parameters are motivated by three empirical patterns: the client-loss risk associated with negative opinions, the high persistence of audit outcomes, and the higher rate at which large auditors initiate negative opinions. In the model these patterns map, respectively, to a *recurring cost* paid whenever a negative opinion is issued, a *switching cost* incurred when transitioning from positive to negative, and a *scale discount* that lowers the cost of strictness for larger auditors.

Comparing estimates before and after the reform, we find that the recurring cost roughly doubled and the scale discount shrank by more than half, more than offsetting a smaller decline in the switching cost. Importantly, the cost structure moves significantly only for auditors exposed to the demand shock, mirroring the difference-in-differences design. Counterfactual exercises show that the contagion effect is not an inevitable consequence of a deregulation reform. In particular, coupling deregulation with targeted cost subsidies – for instance, compensating auditors for the client-loss risk associated with negative opinions – can preserve pre-reform negative audit rates in the still-regulated portion of the market. Our calculations put the cost of such a subsidy on the order of \$25 million per year, less than 0.01% of the awards those audits cover.⁵

Overall, our results uncover a channel that standard policy analysis may understate: the

⁵The experimental setting closest to this counterfactual policy is [Duflo et al. \(2013\)](#), where pollution inspectors in India were paid accuracy-based bonuses from a central pool.

contagion effect of deregulation. Whenever regulatory oversight relies on private compliance professionals, a deregulatory action that contracts demand for their services – by reducing the pool of regulated entities or the scope of services required – can weaken scrutiny of the still-regulated population. This contagion operates through shifts in intermediary incentives and, in principle, can be mitigated by policies that offset the rising cost of rigorous oversight. The margin of adjustment in our setting is only one form the contagion can take. For instance, when compliance professionals provide ex-ante services such as tax advisory or financial planning, a deregulatory demand shock could lead to lower prices or market consolidation. How demand contraction reshapes responses across different settings, and whether successive deregulatory actions produce cumulative impact on still-regulated firms, are promising directions for future research.⁶

Contributions. We make three main contributions to the literatures on threshold-based regulation, the costs of deregulation, and compliance professionals.⁷ First, we uncover a spillover of threshold-based regulation that operates through intermediaries. Prior work focuses on how regulated entities respond to thresholds, for example, by bunching below it. We show that threshold changes also reshape the incentives of the intermediaries who serve still-regulated entities, implying that the standard cost-benefit analysis of threshold policies is incomplete: it misses the weakened oversight for entities that remain above the cutoff.

Second, we identify an overlooked cost of deregulation. A large literature documents the economic gains of reducing regulatory burdens, including cash flow improvements and reduced legal risk (Ewens and Farre-Mensa, 2020). When translated to policy, these gains are often viewed as a justification for threshold increases and other forms of deregulation. Our findings show that deregulation can contract the demand for compliance services, resulting in weaker scrutiny of entities that were never deregulated. This mechanism suggests that the net benefits of deregulation depend not only on the direct relief afforded to deregulated entities,

⁶Indeed, the threshold we study has since risen further, to \$1 million, exempting an additional cohort of recipients and deepening the demand contraction that drives the contagion effect.

⁷See extended literature review in [Appendix A.4](#).

but also on the disruption to the compliance industry and the externalities this disruption imposes on the still-regulated population and the taxpayers who fund it.

Third, we identify a novel link between deregulation and the quality of external monitoring, and show that deregulation propagates across firms due to the market structure of the compliance industry. Existing papers study how competition among private monitors degrades the information they produce and how incentive distortions arise when gatekeepers are paid by the entities they monitor (Becker and Milbourn, 2011). We extend this literature by documenting a new channel affecting monitoring quality: a demand contraction caused by deregulation. Such contraction raises the value of all retained clients at once, resulting in a loss of rigor. The same deregulatory act can therefore yield more or less oversight, depending on the competitive environment in which monitoring is delegated. Because compliance professionals both advise entities on meeting regulatory standards and certify their adherence, future research can explore how this deterioration affects the full scope of the monitoring relationship – not only the ex-post opinion but also the ex-ante guidance that shapes compliance behavior.

Taken together, our results point to a broader lesson: the effectiveness of threshold-based regulation depends not only on where the threshold is set, but also on the market environment in which enforcement is delegated. By documenting how a threshold increase propagates through intermediary incentives, we provide a new perspective on the equilibrium effects of regulation and highlight the importance of incorporating intermediary responses into the design of optimal deregulation.

1 Institutional Setting

Our paper is focused on Single Audit: the mandatory annual audit of entities receiving federal awards.⁸ This is the primary tool for monitoring the use of annual federal award

⁸We use the term “award” as defined in the Code of Federal Regulations, to describe the type of federal funds that are subject to the Single Audit requirement (2 CFR §200.1). This includes federal financial

expenditures, holding recipients accountable and ensuring that taxpayers’ money is properly distributed and utilized. Crucially, entities are required to undergo the Single Audit only if their annual federal expenditures exceed a specified threshold – a key feature motivating our empirical strategy (Section 3.1).

The Single Audit is governed by Title 31 of the U.S. Code, Chapter 75, and Title 2 of the Code of Federal Regulations (CFR), Part 200, Subpart F. It has two components. The first is a *financial-statement audit*: the auditor opines on whether the entity’s financial statements are presented fairly in accordance with generally accepted accounting principles and standards (GAAP and GAAS). Additional auditing procedures – unique to the Single Audit setting – are detailed in the Generally Accepted Government Auditing Standards, commonly known as GAGAS or the **Yellow Book**. The second component is a *compliance audit*. Here, the auditor tests whether the entity complied with the specific requirements governing each federal program, following procedures prescribed in the Compliance Supplement: a detailed manual that specifies the compliance requirements and audit procedures for each federal program. The two components are typically conducted jointly by the same auditor, who must satisfy independence, competence, and licensing requirements and is selected through a competitive process.

The audit produces two types of output (2 CFR §200.515–516). The first is *opinions*. The auditor issues one opinion on the entity’s financial statements and a separate opinion on compliance for each major federal program.⁹ Each opinion falls into one of four categories: unqualified (clean), qualified, adverse, or disclaimer (where the auditor is unable to form an opinion). The second output is the *schedule of findings and questioned costs*, which documents specific deficiencies identified during the audit. Findings fall into two broad categories. *Internal-control deficiencies* assess the entity’s control environment: a *material*

assistance such as grants, loans, and subsidies. The term “Single Audit” dates to the Single Audit Act of 1984, which for the first time allowed recipients to conduct a single audit covering all the federal funds it has received during the year (Pub. L. 98-502, October 19, 1984, 98 Stat. 2327).

⁹Major federal programs are determined through a risk-based approach prescribed in federal regulations. Generally, classification depends on the program’s prominence within the client’s overall federal expenditures and a risk assessment conducted by the auditor.

weakness indicates a serious deficiency such that a material misstatement may not be prevented or detected on a timely basis, and a *significant deficiency* is less severe but important enough to merit attention. *Compliance findings* assess adherence to federal program requirements: *material noncompliance* indicates failure to comply with federal statutes, regulations, or grant terms, and *questioned costs* arise when expenditures may not comply with program requirements. Findings are reported at the financial-statement level and at the individual award level for each federal program. The full audit package is submitted to the **Federal Audit Clearinghouse** (FAC), a centralized and publicly accessible repository.

Audit outcomes have direct consequences. Recipients must respond with a corrective action plan. Moreover, federal agencies may impose sanctions such as withholding payments, disallowing costs, or suspending future funding (2 CFR §200.339). These make audit outcomes economically meaningful for recipients and, indirectly, for auditors whose clients may exit following negative audit outcomes. From the auditors’ perspective, a broader system of oversight aims to impose professional discipline. They are potentially subject to review by the American Institute of Certified Public Accountants and state licensing boards, and, in severe cases, investigations by the Department of Justice. The public availability of audit reports through the FAC further creates reputational incentives.

2 Data and Descriptive Statistics

2.1 Data Sources

We source our primary dataset from the Federal Audit Clearinghouse (FAC),¹⁰ the central repository that collects and disseminates all Single Audit reports. [Appendix A.3](#) provides a detailed description of the data construction process, as well as ancillary datasets, and a

¹⁰In 2023, the FAC migrated from the U.S. Census Bureau to the Government Accountability Office (GAO). An earlier draft of this paper used the Census version; the current draft relies exclusively on the GAO data. The two datasets cover the same underlying audit reports but differ in file structure and field definitions. Our main estimates are qualitatively unchanged across the two versions.

replication kit with code and documentation is available on [the authors' website](#).

Our sample spans audit years 1998–2022. The raw FAC data are released in two vintages with distinct file structures: a historical archive covering 1998–2015 and a redesigned format from 2016 onward. Each vintage consists of several interlinked files. We rely primarily on the report-level files which include the identities of the auditor and auditee, total federal expenditures, and audit opinions. We identify recipients and auditors by their Employer Identification Number (EIN), which provides consistent tracking across years. We recover missing auditor EINs, primarily in the earlier portion of the sample, by standardizing firm names and matching them to known EINs. The opinions are reported separately for financial statements and major programs, each classified as unqualified, qualified, adverse, or disclaimer. Other data files record concrete findings at a more granular level; we use this information in supplementary analyses.

2.2 Descriptive Statistics

The raw dataset covers all Single Audits conducted from 1998 to 2022. It includes 993,414 auditor×recipient annual observations covering \$9.8 trillion over 25 years, expressed in constant 2023 USD. Over the entire sample period, the awards were disbursed among 95,182 unique clients and monitored by 27,264 unique auditors. In an average year, \$390.6 billion in federal awards are expended by 39,494 recipients and monitored by 5,740 auditors. Note that our identification strategy utilizes a subset of all federal awards, essentially those around the \$750,000 audit threshold between 2013–2018. Within this narrower sample and time frame, in an average year, \$351.6 billion in federal awards are disbursed among 33,497 recipients and monitored by 4,268 auditors.

[Table 1](#), Panel A, presents key statistics for our main panel of 993,414 auditor×recipient annual observations.¹¹ The average recipient spends \$9.8 million of federal awards a year (expressed in constant 2023 USD). The distribution is heavily skewed, with the median client

¹¹Since a client undergoes only one annual audit, the auditor×recipient panel is effectively a recipient-level panel. A comprehensive variable dictionary is in [Appendix A.1](#).

spending only \$2.7 million. Pair relations are sticky: conditional on having a relation at time t , the survival rate at time $t + 1$ is 71.9% and the switching rate (where both auditor and client continue to time $t + 1$ but are no longer connected) is 11.0%.

Our key outcome variable, *NegativeAudit*, equals 100 if the auditor issued a qualified or adverse opinion on either the financial statements or any major program. The vast majority of audits (90.5%) resulted in a clean opinion while 9.4% resulted in a negative opinion.¹² Decomposing by component, negative financial-statement opinions occur in 6.3% of audits and negative major-program opinions in 3.6%; decomposing by type, qualified opinions occur in 8.0% and adverse opinions in 1.7%.¹³ Regarding the schedule of findings, 26.9% of all audits reported an internal-control deficiency (material weakness or significant deficiency), 4.9% reported questioned costs, and 4.5% reported material noncompliance with federal program requirements. At the award level, 7.5% of audits had at least one finding and the average share of awards receiving findings is 3.1%. [Table A.3](#) reports pairwise correlations among these various audit outcomes, which are all positively correlated with each other.

To provide a benchmark, we turn to Audit Analytics. We focus on the corresponding sample period (1998–2022) and limit the sample to include only small auditees, defined as firms in the bottom quartile of total assets (below \$28 million). In this sample, pair survival rate from time t to time $t + 1$ – conditional on both auditor and auditee continuing to time $t + 1$ – is 78.2%. As for audit results, note that Audit Analytics does not report adverse or qualified audit opinions, as these outcomes are absent in practice.¹⁴ Instead, we note that the average rates of restatements and adverse restatements are 9.5% and 8.2%, respectively.

In Panel B of [Table 1](#) we reshape the data at the auditor level, with 216,758 auditor×year pairs. The average auditor audits \$45 million in annual federal expenditures, whereas the median auditor audits \$4.4 million. The average auditor has 5 clients in a given year,

¹²The residual is either undisclosed or undetermined.

¹³Neither decomposition sums to the 9.4% negative audit rate because an audit can receive negative opinions on multiple components (financial statements and major programs) and of different types (qualified and adverse).

¹⁴Firms confronted with a qualified opinion almost never release their financial statements, opting to delay publication until an unqualified opinion is obtained.

compared to 1 client for the median auditor. Conditional on having any relation at time t , the survival rate of the average auditor at time $t + 1$ is 57.2%. For comparison, in the Audit Analytics universe, the median auditor in the corresponding period oversees 2 clients and the average survival rate is 83.7%.

3 The Contagion Effect

3.1 Hypotheses and empirical strategy

Our identification strategy is centered on the reform titled the Uniform Administrative Requirements, Cost Principles, and Audit Requirements for Federal Awards (78 FR 78590). The reform raised the threshold from \$500,000 to \$750,000 of annual federal expenditures, exempting entities below the new threshold from the Single Audit process. In particular, 6,056 recipients were below the \$750,000 threshold in the year preceding the reform, connected to 2,276 auditors. Our focus is on the *still-regulated* clients, that is, clients that remained subject to the Single Audit requirement after the reform. The question is whether audit outcomes changed for those clients if their auditors lost a subset of small clients because of the threshold increase. The null hypothesis is that audit outcomes were unaffected by the reform. Our alternative hypothesis is that the reform changed the incentives or constraints faced by exposed auditors, which in turn affected outcomes for their remaining clients. We refer to this dynamic as the contagion effect.¹⁵ In subsequent analysis we show why such a response is economically rational: having lost part of its client base, an exposed auditor places greater value on retaining each remaining engagement, while facing little accountability for a lenient opinion. We develop both sides of this trade-off – the low cost of leniency in [Section 4.4](#) and its high benefit in [Section 4.5](#) – after establishing the baseline effect.

To examine the existence of this potential contagion effect, we estimate the following

¹⁵We discuss concrete channels in [Section 4](#).

difference-in-differences specification:

$$NegativeAudit_{a,c,t} = \alpha + \beta \cdot Treated_a \times Post_t + \lambda + \varepsilon_{a,c,t}, \quad (1)$$

where the outcome equals 100 if auditor a issued a negative opinion (qualified or adverse) for client c at time t , and equals 0 for a positive opinion.¹⁶ The reform applied to fiscal years beginning on or after December 26, 2014 (2 CFR §200.110). Because of that, the time unit t is a twelve-month period running from December 26 to December 25. For our main analysis we use a symmetric three-year window around the reform, yielding three pre-reform (December 26, 2011 – December 25, 2014) and three post-reform (December 26, 2014 – December 25, 2017) periods.

$Post = 1$ for the three post-reform periods. The treatment indicator, $Treated_a$, is defined at the auditor level using the last pre-reform period. An auditor is treated if it had at least one client with expenditures below \$750K in that period. Those clients became exempt from the Single Audit requirement after the reform, so their auditors were directly exposed to a contraction in demand. This contraction was substantial and, as the rulemaking record shows, discussed extensively by the profession itself (Section 4.1). The control group consists of auditors that had no such below-threshold clients in that period, but had at least one client with expenditures between \$1 million and \$1.5 million. This second condition ensures that control auditors were active around the threshold while remaining untreated by the reform.¹⁷ The coefficient of interest, β , captures the change in audit outcomes for continuing clients of treated auditors, relative to continuing clients of control auditors, after the reform.

In all specifications we include period fixed effects, absorbing aggregate shocks common to all observations in the same period, including the standalone $Post_t$ indicator.¹⁸ We include

¹⁶In subsequent tests we distinguish between levels of negativity and also consider the audit’s findings.

¹⁷We use a binary treatment indicator in the baseline specification, given the recent literature highlighting issues with difference-in-differences designs with non-binary treatments and high-dimensional fixed effects (de Chaisemartin and D’Haultfoeuille, 2020; Kempf, Luo, and Tsoutsoura, 2025). In Section 4.1, we exploit continuous variation in the intensity of exposure to the reform – allowing the response to scale with the size of the shock – and obtain consistent results.

¹⁸In particular, those fixed effects remove common post-reform shifts, such as changes in enforcement

various combinations of auditor, client, and pair fixed effects; those absorb time-invariant differences across auditors, clients, and auditor-client relationships, respectively. Standard errors are robust to heteroskedasticity.¹⁹

3.2 Main result

The results from estimating [Equation \(1\)](#) are reported in [Table 2](#).²⁰ In column (1), we include client and period fixed effects. We find that treated auditors reduced negative audits by 1.36 percentage points after the reform, which corresponds to 16.7% relative to the pre-reform baseline of 8.14%. In column (2), we add auditor fixed effects, so the effect is identified from within-auditor variation over time. The estimate increases to 1.75 percentage points (21.5% of the pre-reform mean), indicating that unobserved auditor heterogeneity, if anything, attenuates the baseline result. In column (3), we replace the separate auditor and client fixed effects with auditor×client (pair) fixed effects. The estimate is virtually unchanged at 1.85 percentage points (22.9% of the pre-reform mean), confirming that the effect is identified within enduring auditor-client relationships: the same auditor becomes less likely to issue a negative opinion to the same client after the reform.²¹

The effect is economically meaningful in dollar terms. The estimate implies that roughly 540 audits each year receive a favorable opinion that they would not have received absent the reform, covering \$6.2 billion in annual federal awards (expressed in 2023 dollars). We obtain this figure by applying the estimated treatment effect (-1.85) to the federal awards audited by treated auditors’ continuing clients per each post-reform period (29,372 audits and \$335 billion).²²

The identifying assumption for [Equation \(1\)](#) is that, absent the reform, audit outcomes intensity or in the regulatory environment. We discuss such potential changes in [Section 4.4](#).

¹⁹The results are robust to clustering standard errors by auditor or by client; these results are available upon request.

²⁰In the raw data, negative audit rates decline from 8.75% to 5.19% (treated auditors) and from 6.15% to 4.70% (control auditors), yielding an unconditional difference-in-differences of -2.11 p.p.

²¹We also find that auditor-only and client-only fixed-effect specifications, each with period fixed effects, yield very similar estimates. Those results are available upon request.

²²Dollar amounts are winsorized at the 1st and 99th percentiles, consistent with the descriptive statistics.

would evolve similarly for treated and control auditors. To assess this assumption, we estimate a dynamic version of Equation (1) in event time:

$$NegativeAudit_{a,c,t} = \alpha + \sum_{k \neq -1} \beta^k \cdot Treated_a \times \mathbb{1}\{t = k\} + \lambda + \varepsilon_{a,c,t}, \quad (2)$$

where k indexes periods relative to the reform cutoff and $k = -1$ is the omitted category. The period fixed effects in λ absorb the uninteracted event-time indicators, so only the interactions enter as regressors. The coefficients β^k trace the differential response of treated clients relative to the control group in each period. We use pair and period fixed effects, corresponding to the tightest specification in Table 2, and plot the coefficients $\hat{\beta}^k$ in Figure 1. Prior to the reform ($k < -1$), the estimates are small and statistically indistinguishable from zero, supporting parallel trends. Following the reform ($k \geq 1$), we observe a sharp decline in negative audits among treated auditors relative to the control group. The effect emerges immediately at $k = 1$ and persists through $k = 3$.²³

Finally, we address the possibility that pre-reform differences between treated and control auditors confound the contagion effect. First, we apply the test proposed by Oster (2019), which measures how important selection on unobservables would need to be, relative to selection on observables, to drive the coefficient to zero. Under the recommended assumption that $R^2_{max} = 1.3 \times R^2_{long}$, the resulting δ^* is -34.99 when moving in Table 2 from column (1) to column (2) and -13.78 when moving from column (1) to column (3). This indicates that the coefficient grows in magnitude as controls are added, so that unobservables would need to push in the *opposite direction* from observables to explain the result away. This qualitative reversal is a stronger condition than the conventional $|\delta^*| > 1$ benchmark, which requires only that unobservables matter more than observables while pushing in the same direction.²⁴ Second, we augment the baseline specification with interactions between $Post_t$

²³Our analysis of the rulemaking process (Appendix A.2.2) provides additional support for the absence of pre-trends. In the lead-up to the final rule, the proposed threshold ranged from \$500,000 to \$1,000,000, implying that both treated and control auditors (as we define them) were exposed to the reform’s prospect. This dampens the scope for differential adjustment prior to implementation.

²⁴More generally, any confounding factors that could explain our results must differentially affect treated

and the pre-reform characteristics that differ across groups,²⁵ continuing to include period and pair fixed effects. [Table A.1](#) reports the results. The coefficient on $Treated \times Post$ is virtually unchanged across all columns, ranging from -1.934 to -1.972 (compared to the baseline of -1.956),²⁶ confirming that pre-reform imbalances in observables do not confound the treatment effect.

Overall, the results document a robust contagion effect of deregulation. Following the reform, treated auditors become less likely to issue negative audit opinions for their remaining clients. This shift toward more favorable audit outcomes is economically large, stable across specifications, and evident both at the client and auditor levels. We investigate the channels behind this shift in [Section 4](#), but before that, we provide additional results on the scope of the contagion effect.

3.3 Auditor-Level Outcomes

The primary analysis is conducted at the client $\times$ auditor level, allowing us to absorb client-specific heterogeneity. To examine the effect at the level at which treatment is assigned, we collapse the data to the auditor level and estimate:

$$y_{a,t} = \alpha + \beta \cdot Treated_a \times Post_t + \lambda + \varepsilon_{a,t}, \quad (3)$$

where $y_{a,t}$ captures outcomes for auditor a in period t , and λ includes auditor and period fixed effects. The time unit is the same as in the client-level regressions, aggregating across all

auditors, coincide with the reform’s implementation, operate independently of the demand contraction that defines treatment, and push in the opposite direction from the observables we control for. We are not aware of any candidate that satisfies these conditions simultaneously.

²⁵In particular, in the period immediately before the reform, treated auditors served more Single Audit clients (mean 15.5 versus 4.4) and audited larger federal portfolios (mean \$152 versus \$36 million); relationship tenure was similar across the two groups. To be clear, identification does not require treated and control auditors to be identical in levels; it requires only that their audit outcomes would have evolved in parallel absent the reform, which the event study confirms.

²⁶This baseline re-estimates [Equation \(1\)](#) on the common sample required across all columns of [Table A.1](#): pairs with non-missing pre-reform auditor characteristics for the interaction terms. This sample is smaller than the full sample in [Table 2](#), which accounts for the difference between -1.956 and -1.852 .

clients audited within each period. Standard errors are heteroskedasticity-robust. We focus on auditors who remain active before and after the reform, to isolate behavioral responses from compositional changes, and return to those changes in [Section 4.6](#).

Panel A of [Table 3](#) reports audit outcomes. Treated auditors are 3.1 percentage points less likely to issue any negative audit, a decline of 10.6% relative to the pre-reform mean. The number of negative audits falls by 0.6 per period, a 51.6% decline, and the average negative audit rate falls by 1.1 percentage points, or 14.9% of the pre-reform mean. Panel B examines how the reform reshaped treated auditors’ client portfolios. We define client outflow using a backward-looking measure: an auditor lost a client if an auditor-client pair observed in period $t - 1$ is absent in period t . Treated auditors are 13.7 percentage points more likely to have lost at least one client, a 23.3% increase over the pre-reform rate. They are also 19.8 percentage points more likely to have fewer clients overall, a 69.6% increase. The total number of clients declines by 1.0 per period, or 6.6% of the pre-reform mean of 15.3 clients; conditional on experiencing a net loss, the decline is 2.2 clients, or 12.6% of the conditional mean. These client losses reflect both mechanical attrition – below-threshold clients becoming exempt from the Single Audit – and non-mechanical attrition of above-threshold clients, indicating that the reform’s disruption extended beyond the directly affected engagements. We examine the implications of this disruption in [Section 4](#).

While these per-auditor magnitudes may appear modest, they understate the reform’s impact on federal oversight. The same auditor becomes less likely to issue a negative opinion to the *remaining* clients, who are larger than the deregulated ones, so the weakened scrutiny reaches well beyond the client pool targeted by the deregulation. In the pre-reform years, treated auditors audited roughly \$4.4 billion in at-risk awards – those the reform would exempt – while the still-regulated clients they retained accounted for more than seventy times that amount. The contagion thus propagates a demand shock concentrated in a small slice of federal spending onto a vastly larger base of retained awards; as [Section 3.2](#) shows, roughly \$6.2 billion of that spending each year receives audits that would otherwise have

been negative. The reform’s impact on the oversight of federal awards is therefore far larger than the deregulation that set it in motion.

3.4 Opinions and Deficiencies

The main result establishes that treated auditors’ clients receive fewer negative audit opinions. Before turning to the question of *why*, we provide additional results on *where* in the audit’s output space this effect is concentrated. Recall that our main outcome, *NegativeAudit*, aggregates distinct opinions into a single indicator. We now re-estimate Equation (1) while replacing the dependent variable with each opinion type separately. We construct two intermediate aggregates – *Negative FS*, equal to 100 if the auditor issued any qualified or adverse opinion on the financial statements, and *Negative MP*, defined analogously for major programs – and then further decompose each into its qualified and adverse components.

Figure 2, Panel A, displays the results as a percentage of each outcome’s pre-reform mean, with the full regression output reported in Panel A of Table A.2. We find that the contagion effect flows almost entirely through financial-statement opinions. The *Negative FS* composite declines by 41.1% of its pre-reform mean ($\beta = -1.9$), driven by a sharp drop in adverse FS opinions (-63.9%) and a smaller decline in qualified FS opinions (-13.7%). By contrast, the *Negative MP* composite is unaffected: the point estimate is essentially zero ($+1.3\%$), the qualified MP coefficient is approximately zero, and the adverse MP coefficient, if anything, moves in the opposite direction – none of which is statistically significant at the 5% level.

While *opinions* reflect the auditor’s summary assessment, *findings* document concrete deficiencies identified during the audit. We estimate Equation (1) for five outcomes capturing those deficiencies, in a sample requiring non-missing values for all five.²⁷ Panel B of Figure 2 displays the results, with the full regression output in Panel B of Table A.2. The rate of internal-control deficiencies – an indicator equal to 100 if the entity has any material weakness or significant deficiency – increases by 2.2 percentage points, or 8.3% of the pre-reform mean.

²⁷This ensures that differences across columns reflect heterogeneity in responses rather than variation in sample composition.

Compliance findings are unaffected: both questioned costs and material noncompliance have coefficients indistinguishable from zero. At the award level, neither the total count of findings nor the share of awards with findings shows a significant change.

Taken together, the contagion effect produces a striking divergence between the auditor’s summary judgment and the underlying record of deficiencies. While overall assessments improve, especially the financial-statement opinions, the documentation of concrete deficiencies either increases or remains unchanged. We will return to this result in [Section 4.7](#).

4 Channels

The baseline results establish that treated auditors issue fewer negative opinions to their remaining clients after the reform. Two non-mutually exclusive channels could account for this contagion effect. The first centers on the auditor. Having lost their deregulated clients, treated auditors may offer leniency to those that remain. Most directly, the reform reduces the auditor’s revenue base: with fewer clients generating fees, the marginal value of each remaining engagement rises, and a negative opinion jeopardizes that revenue by raising the risk that the client loses access to federal awards or switches auditors. Treated auditors thus have stronger incentives to avoid negative opinions after the reform, even if underlying client behavior is unchanged. The second channel centers on the client. Clients connected to treated auditors may genuinely become more compliant, so that fewer negative opinions reflect better behavior rather than (or in addition to) weaker oversight – as could occur if the reform differentially simplified reporting requirements, reducing the compliance burden on treated recipients more than comparable controls.

In a reduced-form setting we observe only the final audit outcome, not the auditor’s monitoring choice or the client’s underlying behavior. With that in mind, we conduct a series of tests and find evidence consistent with the auditor-lenieny channel, but inconsistent with the client-compliance one. We describe those tests individually in [Section 4.1](#) through [Section 4.6](#)

and summarize the evidence in [Section 4.7](#). [Section 5](#) complements these reduced-form tests with a structural model that separates auditor and client behavior.

4.1 Exposure to Demand Shock

The reform reduced the pool of small-award recipients, generating a contraction in demand for audit services. The rulemaking record confirms that this contraction was substantial. To see that, we collected and analyzed all 664 public comments; [Appendix A.2.2](#) reports the full record. Many of those letters were sent by the audit community, including CPA societies, individual firms, and public auditors. In particular, commenters described Single Audits as a meaningful source of revenue for firms that specialize in this market, and discussed the expected demand contraction, as the higher threshold would reduce the number of Single Audits performed. The government’s own estimates agreed, projecting that roughly 5,000 entities would lose the requirement and recording 3,400 fewer Single Audit submissions per year afterward. Interestingly, government audit offices – which perform Single Audits under a public mandate rather than for profit – expected no effect on their operations; we will exploit this distinction below.²⁸

If the contagion effect is driven by this contraction, we should observe stronger responses among auditors with greater ex-ante exposure. We capture exposure along two dimensions, both measured in the period immediately preceding the reform: the count of at-risk clients – those with expenditures between \$500,000 and \$750,000 who became exempt after the reform – and the share of at-risk clients relative to the auditor’s total client base. For each measure, we split treated auditors at the median, retain the full control group in both subsamples, and estimate [Equation \(1\)](#) separately on each half.

The first four columns of [Table 4](#) report the results, with the p -value row reporting a formal test of equality across the two subsamples. The contagion effect is concentrated

²⁸This anticipation does not compromise our identification strategy, since the threshold under consideration ranged from \$500,000 to \$1,000,000, so both treated and control auditors faced its prospect; see discussion around [Footnote 23](#).

among auditors with greater exposure. Auditors with an above-median count of at-risk clients show a decline of 27.8%, versus a statistically insignificant effect for those below the median (columns 1–2). The at-risk ratio yields an even sharper contrast: the coefficient is -4.3 for the high-ratio group, a 55.2% decline relative to the pre-reform mean, compared to -1.6 for the low-ratio group (columns 3–4). In both splits the effect is concentrated among the most exposed auditors and absent for the least exposed. This pattern is consistent with an auditor leniency channel, whereby the most exposed auditors demonstrate the strongest response, but it is inconsistent with a story in which client behavior improves.

We demonstrate the importance of exposure from a different perspective. Auditors with larger portfolios have far less at stake in any single engagement, and thus weaker reason to respond to the demand shock. Some of those large auditors are government audit offices, which do not perform Single Audits for profit. If the contagion effect is driven by the demand shock, then excluding these less-exposed auditors should leave it intact. [Table A.4](#) confirms this. We re-estimate the baseline specification while progressively capping the total federal awards an auditor audits at \$1 billion, \$500 million, \$100 million, and finally \$50 million. The estimate is stable throughout and significant at the 1% level in every column, even in the tightest sample which retains only a quarter of the baseline observations. Excluding the largest auditors thus leaves the contagion effect intact, concentrated among the small auditors for which losing a client is most consequential.

Together, the exposure splits and the size restrictions point in the same direction. The contagion effect is strongest where the demand contraction was most severe and the loss of a single client most consequential. This pattern is difficult to reconcile with a pure compliance channel, under which the response should not depend on how exposed an auditor was to the reform.

4.2 Pair Stability

We next examine whether the contagion effect varies with the pre-reform stability of the auditor-client relationship. If the decline in negative audits reflects improved client compliance, pair stability should be largely irrelevant: a client’s incentive to comply depends on the regulatory environment, not on the strength of its bond with the auditor. In contrast, if the effect reflects auditor behavior, it should be strongest in relationships most disrupted by the reform – those where the demand contraction upends previously stable arrangements and compels auditors to adjust.

Columns 5–8 of [Table 4](#) implement this test along two dimensions, using the same split-sample approach. The first exploits variation in state-level market competition, measured as the ratio of unique auditors to unique auditees in the entity’s state. Less competitive markets limit outside options and foster stickier relationships, so the demand shock is more disruptive in these settings. The second classifies pairs based on whether the client ever switched auditors during the pre-reform years. The results follow a consistent pattern. Under the state competition split, the treatment effect is -2.4 (28.6% of the pre-reform mean) in less competitive states, compared to a statistically insignificant 0.2 in more competitive states. Under the pre-reform switching classification, the effect is -2.8 (34.1%) for stable pairs, nearly triple the -1.0 (12.8%) estimated for pairs that switched at least once (columns 7–8).

These findings are difficult to reconcile with a pure compliance channel, under which pair stability should not shape the response to the reform. Instead, they point to disruption in the Single Audit market. In less competitive markets with fewer outside options, and in pairs with no history of switching, the reform upended the status quo and gave auditors the strongest reason to adjust their behavior. Where markets were already more competitive and relationships less entrenched, the reform was less disruptive and the contagion effect is weaker or absent.

4.3 External Monitoring

We next examine whether the contagion effect varies with the degree of external monitoring. To do so, we exploit variation in whether organizations file IRS Form 990. Filing Form 990 subjects organizations to additional scrutiny along two dimensions. First, the IRS reviews these filings. Second, Section 6104 of the tax code requires public disclosure, and most organizations make these forms readily accessible. As a result, Form 990 filers face reputational pressures that are absent for other entities, particularly given their reliance on donations and public trust.²⁹ This setting provides a sharp test. If the decline in negative audits reflects improved client compliance, we should expect stronger effects among Form 990 filers, who face greater external discipline. In contrast, if the effect reflects auditor leniency, it should be attenuated among these entities, as external monitoring constrains auditors' ability to soften assessments. To test these competing predictions, we estimate the following triple-difference specification:

$$\begin{aligned}
 \text{NegativeAudit}_{a,c,t} = & \alpha + \beta_1(\text{Treated}_a \times \text{Post}_t) + \beta_2(\text{Treated}_a \times \text{Post}_t \times \text{Form990}_{c,t}) \\
 & + (\text{Treated}_a \times \text{Form990}_{c,t}) + (\text{Post}_t \times \text{Form990}_{c,t}) + \text{Form990}_{c,t} \\
 & + \lambda + \varepsilon_{a,c,t}.
 \end{aligned}
 \tag{4}$$

Here, $\text{Form990}_{c,t}$ is an indicator equal to one if auditee c files a Form 990 in period t , and zero otherwise. The coefficient β_1 captures the baseline contagion effect for non-filers, while β_2 measures how this effect differs for Form 990 filers. We report several specifications with the same combinations of auditor, auditee, pair, and period fixed effects.

Table 5 presents the results. Across all specifications, the baseline effect (β_1) is negative

²⁹See Harris, Petrovits, and Yetman (2015) for evidence that donors reward better governance as revealed by Form 990 disclosures, and Harris, Petrovits, and Yetman (2024) for evidence that donors reduce contributions following fraud disclosed on Form 990. Indeed, in our baseline sample, Form 990 filers have a negative audit rate of 3.9%, compared to 9.2% for non-filers and 6.9% for the full population. Recognizing this difference, we control for filing status in Equation (4).

and economically large, ranging from -2.57 to -3.57 percentage points. In contrast, the triple interaction term (β_2) is positive, large in magnitude (between 2.49 and 3.95 percentage points), and statistically significant. Taken together, these coefficients imply that the contagion effect is concentrated among entities that do *not* file Form 990. For Form 990 filers, the net effect ($\beta_1 + \beta_2$) is close to zero – ranging from -0.08 to 0.38 percentage points – and statistically indistinguishable from zero in all three specifications, indicating no meaningful decline in negative audits.

These findings are difficult to reconcile with a compliance-based explanation. If deregulation improved compliance, externally monitored organizations should respond at least as strongly, if not more so. Instead, the absence of any decline among Form 990 filers indicates that external scrutiny constrains the underlying mechanism. This is consistent with an auditor leniency channel: reductions in negative audits arise where auditors face fewer external constraints and attenuate when their assessments are subject to independent monitoring.

4.4 Compliance Incentives

In this section we study whether the reform differentially increased the incentives to comply, motivating treated clients to become more compliant. Such a differential response would result in fewer negative audits in the treated group, even if auditor leniency remains unchanged. We focus on two broad compliance incentives: lower regulatory burden, and more intense enforcement.

Regulatory burden. As noted earlier, the reform modified parts of the auditing process beyond raising the threshold (see [Section 3.1](#) and [Appendix A.2](#)). If these changes differentially reduced compliance costs for clients of treated auditors, the improvement in their audit outcomes could reflect better compliance rather than greater auditor leniency. To study this possibility, we use textual tools to quantify the amount and direction of change in

audit requirements. Specifically, we compare the text of the Compliance Supplement – the document that specifies audit procedures – in its pre-reform (March 2014) and post-reform (June 2016) editions.³⁰ Figure 3 summarizes the results, with all metrics normalized to the pre-reform value of 100.

The post-reform regulatory text is substantially longer, as the number of words increases by 70% and the number of sentences nearly doubles. Sentence-level readability, represented by five standard indices that estimate the education level required to comprehend a text,³¹ barely change and remain within a few points of 100. Structural complexity increases sharply: the absolute count of cross-references – instances of terms such as “refer to” – nearly doubles, and their density per 1,000 words rises by 12%. This means that the post-reform Supplement relies heavily on external texts, and understanding any individual provision now requires consulting multiple other sources, increasing the effective complexity of compliance.

Taken together, the evidence is inconsistent with a simplified post-reform regulatory environment. The post-reform compliance manual is 70% longer, nearly twice as cross-referenced, and only marginally more readable at the sentence level. Most importantly, even if compliance burden had fallen, it is difficult to see how it could contribute to the contagion effect. Both treated and control auditors operate under the same post-reform regulatory framework, and the identifying variation in Equation (1) comes from differential exposure to the demand shock, rather than from common changes in audit standards.³² The treatment effect we estimate reflects the differential response of auditors who lost clients relative to

³⁰The Compliance Supplement is published annually by OMB as an appendix to the governing audit framework. Before the reform it was designated Appendix B to OMB Circular A-133; after the reform it became Appendix XI to 2 CFR Part 200. We compare Part 3 of each Supplement edition (“Compliance Requirements”), which contains the general audit objectives and suggested procedures applicable to all programs. See Appendix A.2.4 for additional details and a robustness check using the full Supplement text.

³¹The Flesch-Kincaid Grade Level, which combines average sentence length and average syllables per word; the Gunning Fog Index, which weights the proportion of polysyllabic words alongside sentence length; the SMOG Index, which estimates years of education from the count of polysyllabic words per sample of sentences; the Coleman-Liau Index, which relies on character counts rather than syllable counts; and the Automated Readability Index, which substitutes characters-per-word for syllables-per-word.

³²A reduction in compliance burden could, in principle, contribute to the negative coefficient on the *Post* indicator, which captures the common post-reform shift in audit outcomes. But we find little evidence that the post-reform regulatory text is substantially simpler.

those who did not, within the same regulatory environment.

Enforcement. A separate possibility is that the reform increased the enforcement risk for treated clients. Such an elevated risk would motivate them to comply, resulting in fewer negative audits. To study this possibility, we collect additional data on various enforcement mechanisms related to Single Audits. In principle, auditors face review by the AICPA and state licensing boards and, in severe cases, investigation by the Department of Justice; in practice these safeguards rarely bind, and more importantly, we find no differential increase in enforcement intensity after the reform.

First, we searched for enforcement actions relating to federal awards between 2010 and 2019 on the Department of Justice’s news database, as well as in the enforcement database of the Department of Health and Human Services, the largest grantmaking agency (excluding procurement deals). Our search yields 76 enforcement actions relating predominantly to award recipients, and we list several examples in [Table A.9](#). We find no evidence of greater enforcement activity in the post-reform period. More importantly, we identified only one instance involving an independent auditor over the entire sample period.³³ The matter resulted in a referral to the AICPA but to our knowledge there is no record of charges, penalties, or licensing consequences. Relative to the 5,740 auditors of an average year, a single referral over a decade corresponds to roughly 0.002% of auditor-years, and we are not aware of any criminal proceeding against a Single Audit auditor in any year we examine.

We also examine disciplinary actions published by state accountancy boards in Texas and California, the two most populous states. We download the quarterly reports of the Texas State Board of Public Accountancy ([Link](#)) and the California Board of Accountancy’s UPDATE newsletter ([Link](#)) between 2010 and 2020, and we flag enforcement actions related to Single Audits using the keywords “single audit,” “A-133,” “Uniform Guidance,” “GAGAS,” and “government auditing standards.” Our search covers 5,360 actions in Texas and 569

³³In a 2015 quality control review of PricewaterhouseCoopers LLP’s Single Audit of Big Brothers Big Sisters of America, the DOJ Office of Inspector General identified extensive deficiencies, leading to a subsequent withdrawal and replacement of the report.

actions in California.³⁴ We identify only five notices (0.08% of all actions) clearly related to Single Audits, involving four practitioners; three of the four retained the right to practice.³⁵

Finally, we examine AICPA notices involving disciplinary actions against auditors. We are only able to observe notices from 2019 to 2026 ([Link](#)). The AICPA publishes notices that identify the member and the date of the action but does not provide detailed accounts for the causes of these disciplinary actions, and investigations are conducted confidentially. We find no instances of published AICPA actions that are explicitly tied to deficiencies or misconduct relating to Single Audits.

4.5 Rational Leniency

[Section 4.4](#) showed that the reform did not raise the cost of leniency to auditors: quality reviews and sanctions are rare, and enforcement did not intensify after the reform. This section turns to the other side of the trade-off and shows that leniency carries tangible benefits, making it a rational response to the demand shock.

First, a negative opinion sharply raises the odds that a relationship ends ([Table A.5](#)). In the years just before the reform, a pair that received a negative opinion was 6.4 percentage points less likely to survive to the next year, reflecting both a higher probability that the client exits the Single Audit universe and that it switches to another auditor. The gap survives entity-state and period fixed effects and holds over the full sample. Importantly, the risk is not confined to a single year: audit outcomes are persistent, with a negative opinion followed by another the next year 68.3% of the time, against 2.5% after a positive one ([Table 1](#)). A negative opinion thus threatens the relationship for as long as the auditor keeps issuing it.

³⁴Published actions are a subset of board activity, since some cases likely close without a sanction and are never printed. For reference, while Texas reported 623 disciplinary cases closed in fiscal year 2013, we identified 467 actions printed in the four corresponding Board Reports. Similarly, California reported 58 cases closed in fiscal year 2012–13, whereas 55 actions were printed in the three corresponding report issues.

³⁵Ronald Eugene Kahanek and Kahanek, Franke & Asso., L.C. were reprimanded in Texas in May 2016 after three successive deficient peer reviews, paid \$953 in costs, and were barred from performing Single Audits. In California, Ronald Lee Rienks (June 2017) and Peter Junio Sinambal (April 2019) each received a stayed revocation with three years' probation and costs of \$7,514 and \$17,610, and Wilfredo A. Penoliar surrendered his license and paid \$38,113 (October 2018).

A simple calculation illustrates the value at stake. A relationship pays an annual fee f and survives to the next year with probability s , so its present value is $f/(1 - \rho s)$ for a discount factor ρ . In the years before the reform, a relationship in good standing survived with probability 78.5% and, at $\rho = 0.95$, was worth nearly four annual fees ($f/(1 - \rho s) = 3.93f$). A negative opinion cut the one-year survival probability to 72.1%, lowering that value to $3.17f$. The auditor thus forfeits roughly $0.76f$ each time it issues a negative opinion – a recurring loss on a relationship worth close to four years of fees.³⁶

Second, a Single Audit engagement is valuable and worth protecting. Practitioners place Single Audit fees 25% to 50% above comparable audits ([Giddings Consulting Group](#)), a premium that reflects expertise general practitioners rarely maintain – essentially a non-transferable upfront investment.³⁷ Our analysis of a limited sample of nonprofit Form 990 filers (see [Section 4.3](#)) is consistent with this account, showing that Single Audit nonprofits report substantially higher accounting fees than non-Single Audit nonprofits.

Third, the demand shock was large enough to matter for treated auditors, given their pre-reform profile. The median treated auditor held five Single Audit clients, of which the at-risk clients – those in the newly deregulated \$500,000–\$750,000 band – made up a median of 31.2%. At-risk clients supplied a median of 18.6% of a treated auditor’s fee revenue, a mean of 35.9%, and 64% or more for a quarter of auditors.³⁸ Deregulating even one or two clients thus removes a meaningful share of the practice. Many treated auditors operate locally and cannot easily backfill a lost engagement from an adjacent market: the median auditor operates in a single state, and 81.1% draw their entire Single Audit book from one state. Finally, treated auditors specialize. Matching the 2021 directory of the Governmental Audit Quality Center – a voluntary AICPA association whose members commit to quality requirements for governmental audits – to our pre-reform treated auditors, we find that a third (33.7%) belong to the Center.³⁹ For a specialized, local firm, a lost Single Audit

³⁶An extension of our structural model prices this separation risk directly; see [Appendix A.6.6](#).

³⁷See [The CPA Journal \(May 2019\)](#) and [Waymire, Webb, and West \(2018\)](#).

³⁸These numbers are based on treated auditors matched to the Form 990 data.

³⁹We were unable to retrieve earlier directories, including through the Wayback machine.

engagement is not readily replaced.⁴⁰

Taken together, these facts describe an environment in which the stakes of client loss are high and, as [Section 4.4](#) showed, the cost of leniency is low. The median treated auditor was poised to lose roughly a third of its Single Audit book – a substantial contraction for a small, local, specialized practice with few outside options. Having already ceded part of its client base to deregulation, such an auditor has strong reason to protect the profitable, recurring relationships that remain, and little to deter it from doing so. This is why the demand shock can propagate into leniency toward the clients an auditor keeps.

4.6 Compositional Changes

In this section we study whether the contagion effect reflects compositional changes rather than a shift in auditor behavior. We consider three concrete possibilities: that stringent auditors exited the market after the reform, that noncompliant clients strategically stayed below the new threshold, and that clients above \$750,000 are inherently more compliant. [Table 6](#) addresses each in turn.

Panel A imposes cumulative sample restrictions on the baseline specification, where each column is a strict subset of the previous one. Column (1) reproduces the baseline with pair and period fixed effects. To rule out auditor exits, column (2) restricts to auditors observed in at least one pre-reform and one post-reform period, dropping 1,997 observations; column (3) further requires that the auditor appear in all six baseline periods, dropping an additional 11,311 observations. In both cases, the point estimates remain virtually identical to the baseline (-1.852 and -1.864) and equality with the baseline cannot be rejected. To rule out strategic sorting, column (4) further restricts to clients observed both before and

⁴⁰We randomly selected 10 treated auditors and collected public information on their profiles; see [Table A.10](#). These appear to be local practices that advertise governmental and nonprofit Single Audits as a core service line. Note that these characteristics describe why the reform mattered for treated auditors, but they do not drive our estimates. The identification in our empirical design relies on within-pair comparisons of treated and control auditors around the reform. As mentioned, treated auditors show no differential pre-trends (see [Figure 1](#)) and interacting the *Post* indicator with pre-reform auditor characteristics leaves our estimates essentially unchanged (see [Table A.1](#)).

after the reform, dropping 8,531 additional observations; the estimated effect is -1.864 , also indistinguishable from the baseline. Column (5) retains only clients whose federal expenditures exceed \$750,000 in every period, including the pre-reform years, so that these entities were never below the deregulation threshold. This removes another 61,559 observations. The point estimate is somewhat smaller (-1.250) but remains highly significant and similar to the baseline.⁴¹

Finally, Panel B tests whether clients above \$750,000 are inherently more compliant. If so, treated auditors – who lost their small clients and were left with larger recipients – would mechanically show fewer negative audits. To test this, we estimate:

$$NegativeAudit_{c,t} = \alpha + \beta \cdot Above750K_{c,t} + \lambda + \varepsilon_{c,t}, \quad (5)$$

where $NegativeAudit_{c,t}$ equals 100 if client c received a negative audit in year t , and $Above750K_{c,t}$ equals one if the client’s federal expenditures exceed \$750,000. The sample is limited to 2004–2012. During this early period, the \$750,000 level had no legal significance. Across all three specifications – auditee fixed effects (column 1), auditee and auditor fixed effects (column 2), and pair fixed effects (column 3), each with period fixed effects – the coefficient on $Above750K$ is small and positive; in column (1) it is statistically significant, but with proper fixed effects in columns (2)–(3) it is insignificant. Thus, clients above and below the threshold do not differ in their audit outcomes.

In sum, the contagion effect does not seem to be an artifact of compositional changes. The baseline estimate survives sample restrictions (Panel A), and we find no pre-existing differences in audit outcomes around the \$750,000 threshold (Panel B). These results reinforce the interpretation that the reform altered auditor behavior rather than the composition of audited entities.

⁴¹In [Appendix A.5](#), we directly test for strategic exit behavior. We find that some organizations respond to the reform by exiting the Single Audit system, a behavior that is concentrated in the bins closest to the threshold. A marginally significant differential between treated and control auditors appears only in the narrowest bin (\$750K–\$800K) and disappears in all wider bins.

4.7 Summary

The evidence in this section points toward a change in auditor incentives over other explanations. The rulemaking record shows how the audit profession anticipated a material contraction in Single Audit demand, and the government projected thousands of newly exempt entities. Leniency is a rational response to such a demand shock: the relationships a treated auditor protects are economically valuable, and the accountability an auditor faces for a deficient audit is minimal. The contagion effect is concentrated among the auditors most exposed to the demand shock and in previously stable relationships, the settings most disrupted by the reform. It attenuates under external monitoring, where independent scrutiny leaves auditors little room to soften their assessments, and it persists when we hold auditor and client composition fixed, ruling out compositional shifts. Furthermore, we find no evidence of a differential increase in enforcement, nor evidence of a differential reduction in compliance costs; if anything, the post-reform regulatory environment became somewhat more complex for all auditors.

The results in [Section 3.4](#) sharpen the distinction between auditor judgment and underlying compliance. If the reform had improved compliance, both the auditor’s overall assessment and the underlying record of deficiencies should improve. Instead, the opposite occurs: internal-control deficiencies increase, compliance-related outcomes (e.g., award-level findings) are unaffected, and yet negative opinions decline. This divergence between the auditor’s summary judgment and the factual record is difficult to reconcile with genuine improvement in client behavior.⁴²

Finally, if the contagion effect reflects a permanent improvement in compliance among treated auditors’ clients, it should persist rather than dissipate as markets adjust. We test this by extending the event study in [Figure 1](#) to a longer post-reform horizon, estimating [Equation \(2\)](#) with three pre-reform and seven post-reform periods. The effect is largest

⁴²The increase in internal-control deficiencies could reflect a genuine deterioration in compliance, which would suggest that the decline in negative auditor opinions reflects even greater leniency by auditors. It is also possible that auditors document deficiencies strategically to create a paper trail for lenient assessments.

in the immediate post-reform years and attenuates thereafter, consistent with a one-time demand shock whose competitive pressure weakens as treated auditors reconstitute their client portfolios.⁴³ A permanent compliance improvement, by contrast, would predict a persistent effect.

Taken together, treated auditors facing a negative demand shock soften their assessments for remaining clients.⁴⁴ The reduced-form evidence points consistently to auditor incentives, but it cannot separate the specific cost components driving the shift. We turn to a structural model in [Section 5](#) to decompose the overall effect into interpretable cost components and evaluate counterfactual reforms that would neutralize the contagion effect.

5 Quantifying Auditor Incentives

To decompose the contagion effect into interpretable forces, we develop a parsimonious structural model of the auditor’s decision to issue a negative opinion. The full model, estimation, and counterfactual exercises are in [Appendix A.6](#); here, we summarize the key ingredients and findings.

Three empirical patterns in the data point to distinct forces shaping the auditor’s decision. First, negative opinions are costly and strongly predict relationship dissolution. Second, audit outcomes are highly persistent. Both facts (discussed in [Section 4.5](#)) motivate a *recurring cost* paid whenever a negative opinion is issued, and a *switching cost* incurred once when transitioning from positive to negative, respectively. Third, large auditors are more likely to initiate a negative opinion than small auditors,⁴⁵ suggesting that portfolio size lowers the cost

⁴³As a complementary approach, we symmetrically expand both the pre- and post-reform windows around the baseline, starting from ± 2 event-time periods and extending to ± 7 . The estimated effect is largest in the narrowest windows and attenuates as the window expands.

⁴⁴We do not take a stand on the socially optimal rate of negative audits. The demand shock may have corrected pre-reform over-strictness, bringing opinions closer to their efficient level, or pushed auditors below it. Either way, the contagion mechanism is the same: deregulation of some clients altered auditor incentives for their remaining clients.

⁴⁵Among continuing relationships whose prior opinion was positive, auditors with an above-median number of clients initiate a negative opinion in 2.6% of cases, versus 1.9% for smaller auditors, a statistically significant gap.

of strictness and motivating a *scale discount* that decreases with the auditor’s client base. We parameterize these three forces in a dynamic discrete choice model. Each period, the auditor weighs the cost of a negative opinion against a private information shock – an unobserved signal about the client’s true compliance – and chooses the opinion that maximizes the expected present value of future payoffs.

We estimate the three cost parameters from the observed sequence of audit decisions and compare estimates across the pre- and post-reform periods. The three parameters are separately identified because each shapes a different feature of the audit sequence: the recurring cost governs how often negative opinions appear, the switching cost how persistent they are, and the scale discount how their initiation varies with auditor size. The results, reported in [Table A.7](#), reveal substantial shifts. The recurring cost θ nearly doubled, meaning the baseline penalty for issuing a negative opinion rose sharply – consistent with the increased value of retaining remaining clients. The switching cost η fell by roughly one-quarter, which on its own would make negative opinions easier to initiate; but this is more than offset by the other shifts. The scale discount δ shrank by more than half, meaning large auditors lost much of their cost advantage. The net effect is a decline in the model-implied rate of negative audits whose direction and order of magnitude are consistent with the reduced-form estimate. Importantly, when estimating the model separately for treated and control auditors, we find that the parameters move significantly only for the treated group. This suggests that the shift in the cost structure is a treatment effect of the reform on treated auditors, rather than a market-wide trend.

We develop and estimate three versions of the baseline model, each addressing a distinct concern. In the first, the client’s decision to terminate the relationship depends on the auditor’s opinion. This lets us interpret the recurring cost as the present value of the client relationship an auditor risks losing when it issues a negative opinion. In the second, the cost of a negative opinion depends on observable client characteristics. This version reveals that post-reform leniency concentrates on small clients near the deregulated threshold. In

the third, we allow unobserved heterogeneity across auditor-client pairs. This confirms that the persistence of audit outcomes reflects a genuine switching cost rather than permanent differences in client compliance. Across all three versions, the central conclusion remains: the cost of a negative opinion rose sharply for treated auditors after the reform.

The model also enables counterfactual analysis. Shutting down the switching cost entirely would raise the negative opinion rate to nearly five times its current level. This is the largest response to any single parameter, confirming that the switching cost is an important barrier to auditor strictness.⁴⁶ We then turn to policy: what combinations of auditor costs would have preserved the pre-reform rate of negative audits, even as the scale discount shrank to its post-reform level? [Figure 4](#) plots the resulting frontier of contagion-free reforms. The actual post-reform cost parameters lie above this frontier, confirming that the combined shift was too large to sustain pre-reform audit rates. However, policies that reduce the recurring cost of negative audits could offset the contagion effect without giving up on the threshold increase. For example, reversing two-thirds of the increase in the recurring cost would fully restore the pre-reform negative audit rate. This could be achieved via a government-sponsored insurance program compensating auditors who lose clients following a negative opinion. The implied subsidy is small. We show that the subsidy amounts to roughly \$12,000 for each negative opinion; applied to the approximately 2,150 negative opinions issued each year in the estimation sample, restoring the pre-reform negative audit rate would cost about \$25 million per year – less than 0.01% of the \$304 billion in federal awards covered by audits in the estimation sample. A policy that neutralizes the contagion effect is therefore inexpensive relative to the oversight it preserves.⁴⁷

⁴⁶Note that in this exercise we vary only the switching cost, and hold all other parameters – including the auditor’s information about clients’ compliance – fixed.

⁴⁷Dollar figures use the normalization in [Appendix A.6.6](#) (one utility unit equals the average annual Single Audit fee). This exercise is positive rather than normative: it quantifies the cost adjustments that would neutralize the behavioral response to the reform, taking the threshold change itself as given.

6 Conclusions

Each year, the federal government distributes hundreds of billions of dollars in awards to nonprofits, universities, and other entities. Recipients spending above a designated threshold must undergo a Single Audit conducted by a private auditor. We study a 2015 reform that raised this threshold from \$500,000 to \$750,000, exempting roughly 14% of recipients from mandatory auditing. Using a comprehensive award-level dataset spanning 1998–2022, we document an unintended contagion effect: auditors who lost low-spending clients became significantly more lenient with their remaining high-spending clients, weakening the monitoring of federal awards worth hundreds of billions of dollars.

Our reduced-form estimates show that treated auditors – those who had at least one client below \$750,000 before the reform – reduced negative audit opinions by up to 23% relative to the pre-reform mean. Documentation of internal-control deficiencies increases, even as overall negative opinions decline. The effect is strongest among auditors with the greatest exposure to the demand shock, in previously stable auditor-client relationships, and in less competitive markets – the settings most disrupted by the reform. External monitoring attenuates the effect: entities subject to IRS Form 990 disclosure show no decline in negative opinions. The contagion effect is transitory, peaking in the immediate post-reform years and dissipating as the market adjusts. These patterns collectively point to a change in auditor incentives rather than improved client compliance. Although the reform exempted only small recipients, the auditors it affected retained far larger clients – more than seventy times the deregulated volume in our sample. Applying our estimates to the awards these auditors continue to oversee, roughly \$6 billion in federal spending each year receives an audit that would otherwise have been negative.

Our findings carry implications for the design of deregulation. Threshold-based regulation is ubiquitous, and proposals to ease the burden on small entities are common across many domains. We show that such reforms can impose externalities on still-regulated participants through the channel of compliance professionals – in our case, auditors, typically small and

specialized firms for whom a retained engagement is a substantial share of revenue, who respond to lost business by softening their assessments. A structural model of the auditor’s decision confirms this interpretation: the reform raised the recurring cost of issuing a negative opinion, lowered the cost of transitioning to one, and shrank auditor economies of scale. On net, the shift in the cost structure generates a decline in negative audits comparable to the one we estimate in the data. The contagion effect need not be inevitable, however. Counterfactual exercises show that coupling deregulation with targeted cost subsidies – for instance, compensating auditors for the effort and client-loss risk associated with negative opinions – can preserve monitoring quality in the still-regulated portion of the market, at a cost equal to a small fraction of the awards at stake.

More broadly, deregulation can weaken oversight whenever the affected market relies on intermediaries whose incentives are shaped by competitive pressure. Rating agencies, food-safety inspectors, and environmental consultants all face analogous tensions between rigor and client retention. Understanding when and how contagion arises in these settings is a promising direction for future work.

References

- Alvero, A., S. Ando, and K. Xiao. 2023. Watch what they do, not what they say: Estimating regulatory costs from revealed preferences. *The Review of Financial Studies* 36:2224–73.
- Arcidiacono, P., and P. B. Ellickson. 2011. Practical methods for estimation of dynamic discrete choice models. *Annual Review of Economics* 3:363–94.
- Armstrong, D. M., S. Glaeser, and J. L. Hoopes. 2025. Measuring firm exposure to government agencies. *Journal of Accounting and Economics* 79:101703–.
- Becker, B., and T. Milbourn. 2011. How did increased competition affect credit ratings? *Journal of Financial Economics* 101:493–514.
- Bolton, P., X. Freixas, and J. Shapiro. 2012. The credit ratings game. *Journal of Finance* 67:85–112.
- Chen, P.-L., J. Hajda, and J. Kalmenovitz. 2024. Escaping pay-for-performance. *Working Paper*.
- Chung, H., and S. Kallapur. 2003. Client importance, nonaudit services, and abnormal accruals. *The Accounting Review* 78:931–55.
- Cuneo, M., J. Leder-Luis, and S. Vannutelli. 2023. Government audits. Working Paper, National Bureau of Economic Research.

- de Chaisemartin, C., and X. D’Haultfœuille. 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review* 110:2964–96.
- DeAngelo, L. E. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3:183–99.
- DeFond, M. L., K. Raghunandan, and K. R. Subramanyam. 2002. Do non-audit service fees impair auditor independence? Evidence from going concern audit opinions. *Journal of Accounting Research* 40:1247–74.
- Donelson, D., J. A. Garfinkel, and S. Roudini. 2024. Regulatory policy enforcement and corporate performance. *Available at SSRN* .
- Duffo, E., M. Greenstone, R. Pande, and N. Ryan. 2013. Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from india. *The Quarterly Journal of Economics* 128:1499–545.
- Ewens, M., and J. Farre-Mensa. 2020. The deregulation of the private equity markets and the decline in ipos. *Review of Financial Studies* 33:5463–509.
- Ewens, M., K. Xiao, and T. Xu. 2024. Regulatory costs of being public: Evidence from bunching estimation. *Journal of Financial Economics* 153:103775–.
- Gopalan, Y., A. Kalda, and A. Manela. 2021. Hub-and-spoke regulation and bank leverage. *Review of Finance* 25:1499–545.
- Harris, E. E., C. M. Petrovits, and M. H. Yetman. 2015. The effect of nonprofit governance on donations: Evidence from the revised form 990. *The Accounting Review* 90:579–610.
- . 2024. Spreading the news: Donor response to disclosures about nonprofit fraud. *Nonprofit and Voluntary Sector Quarterly* 53.
- He, L., and T. M. Whited. 2025. Underrepresentation of women ceos. *Review of Financial Studies, forthcoming* .
- Hirtle, B., A. Kovner, and M. Plosser. 2020. The impact of supervision on bank performance. *The Journal of Finance* 75:2765–808.
- Iliev, P. 2010. The effect of sox section 404: Costs, earnings quality, and stock prices. *The Journal of Finance* 65:1163–96.
- Jayaratne, J., and P. E. Strahan. 1996. The finance-growth nexus: Evidence from bank branch deregulation. *Quarterly Journal of Economics* 111:639–70.
- Kalmenovitz, J. 2021. Incentivizing financial regulators. *The Review of Financial Studies* 34:4745–84.
- . 2023. Regulatory intensity and firm-specific exposure. *The Review of Financial Studies* 36:3311–47.
- Kalmenovitz, J., M. Lowry, and E. Volkova. 2025. Regulatory fragmentation. *The Journal of Finance* 80:1081–126.
- Kalouptsi, M., P. T. Scott, and E. Souza-Rodrigues. 2021. Identification of counterfactuals in dynamic discrete choice models. *Quantitative Economics* 12:351–403.
- Keating, E. K., M. Fischer, T. P. Gordon, and J. Greenlee. 2005. The single audit act: How compliant are nonprofit organizations? *Journal of Public Budgeting, Accounting & Financial Management* 17:285–309.
- Kempf, E., M. Luo, and M. Tsoutsoura. 2025. The political economy of firm networks: CEO ideology and global trade. Working paper, NBER No. 33712.
- Kincaid, J. P., R. P. Fishburne Jr, R. L. Rogers, and B. S. Chissom. 1975. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease

- formula) for navy enlisted personnel. Working Paper.
- Kisgen, D. J., and P. E. Strahan. 2010. Do regulations based on credit ratings affect a firm's cost of capital? *The Review of Financial Studies* 23:4324–47.
- Magnac, T., and D. Thesmar. 2002. Identifying dynamic discrete decision processes. *Econometrica* 70:801–16.
- Matveyev, E., and S. Zeume. 2025. Financing nonprofits. *Available at SSRN 6145866* .
- Mc Laughlin, G. H. 1969. Smog grading-a new readability formula. *Journal of reading* 12:639–46.
- Opp, C. C., M. M. Opp, and M. Harris. 2013. Rating agencies in the face of regulation. *Journal of Financial Economics* 108:46–61.
- Oster, E. 2019. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics* 37:187–204. doi:10.1080/07350015.2016.1227711.
- Reynolds, J. K., and J. R. Francis. 2000. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30:375–400.
- Rust, J. 1987. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica* 55:999–1033.
- . 1994. Estimation of dynamic structural models, problems and prospects: discrete decision processes. In *Advances in Econometrics: Sixth World Congress*, 119–70. Cambridge University Press.
- Swait, J., and J. Louviere. 1993. The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research* 30:305–14.
- Tassin, K. L., T. R. Waymire, and C. S. Hines. 2019. A historical evaluation of the single audit: Thirty years from initial legislation to uniform guidance. *Journal of Governmental and Nonprofit Accounting* 8:21–35.
- Waymire, T. R., T. Z. Webb, and T. D. West. 2018. A comprehensive analysis of findings from single audits: The implications of auditee type and auditor expertise. *Journal of Governmental & Nonprofit Accounting* 7:55–77. ISSN 2155-3815. doi:10.2308/ogna-52163.
- Wooldridge, J. M. 2005. Simple solutions to the initial conditions problem in dynamic, non-linear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics* 20:39–54.

Figure 1: **Event Study: Deregulation and Audit Outcomes**

Results from estimating [Equation \(2\)](#) with pair and period fixed effects. k indexes periods relative to the reform cutoff date of December 26, 2014; each period runs from December 26 to December 25. The sample includes six periods and the omitted period, $k = -1$, is the last pre-reform period (December 26, 2013 – December 25, 2014). Each point represents the interaction coefficient $\hat{\beta}^k$, and whiskers show 95% confidence intervals based on robust standard errors. The red vertical line marks the reform cutoff. See [Section 3.2](#).

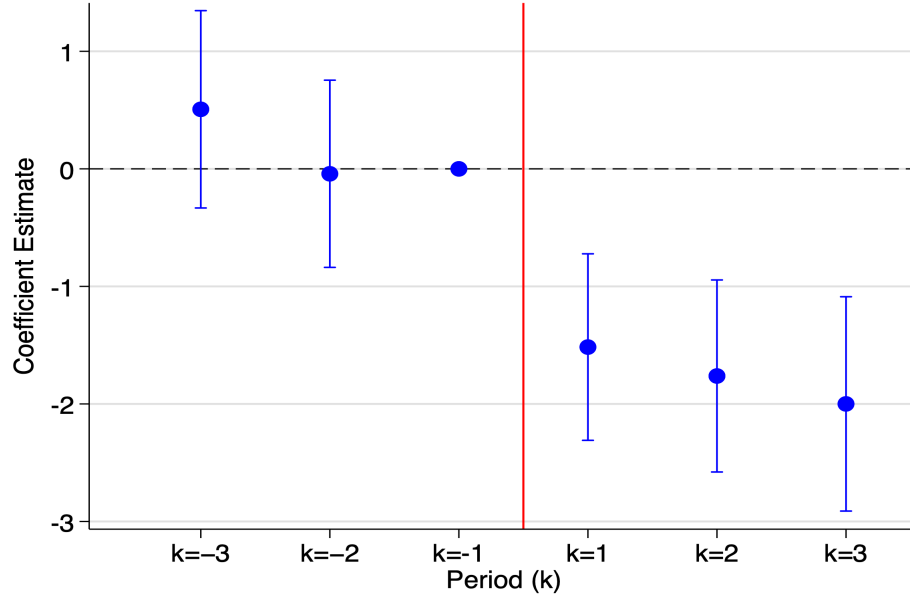
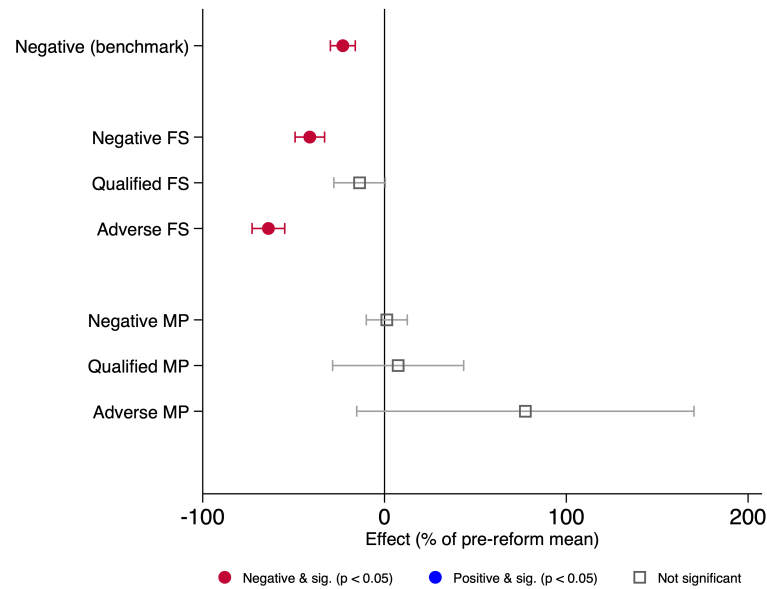


Figure 2: **Decomposition of Audit Outcomes**

Each panel plots the coefficient on $Treated \times Post$ from separate estimations of Equation (1), expressed as a percentage of the outcome's pre-reform mean. Panel A decomposes the aggregate *Negative Audit* indicator into financial-statement (FS) and major-program (MP) opinions, further split into qualified and adverse components. Panel B reports findings-level outcomes: internal-control deficiency, questioned costs, material noncompliance, and award-level findings (total count and share of awards). Filled circles denote significance at the 5% level; hollow squares are not significant. Horizontal bars are 95% confidence intervals. All specifications include pair and period fixed effects and heteroskedasticity-robust standard errors. The full regression results are in Table A.2. See Section 3.4.

Panel A: Opinions



Panel B: Findings

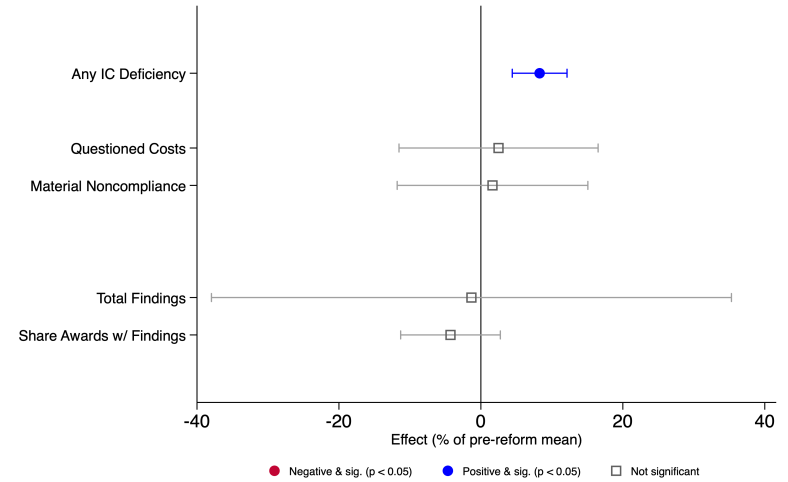


Figure 3: **Textual Analysis of Compliance Cost**

This figure compares the textual properties of the Compliance Supplement in its pre-reform (March 2014) and post-reform (June 2016) editions. All metrics are normalized to the pre-reform value (= 100). The first group reports document length (word count, sentence count). The second group reports five readability indices, where lower values indicate easier readability: the Flesch-Kincaid Grade Level, Gunning Fog Index, SMOG Index, Coleman-Liau Index, and Automated Readability Index. The third group reports structural complexity measures: cross-references per 1,000 words, absolute cross-reference count, type-token ratio, defined-terms count, and passive voice percentage. See [Section 4.4](#) and [Appendix A.2.4](#).

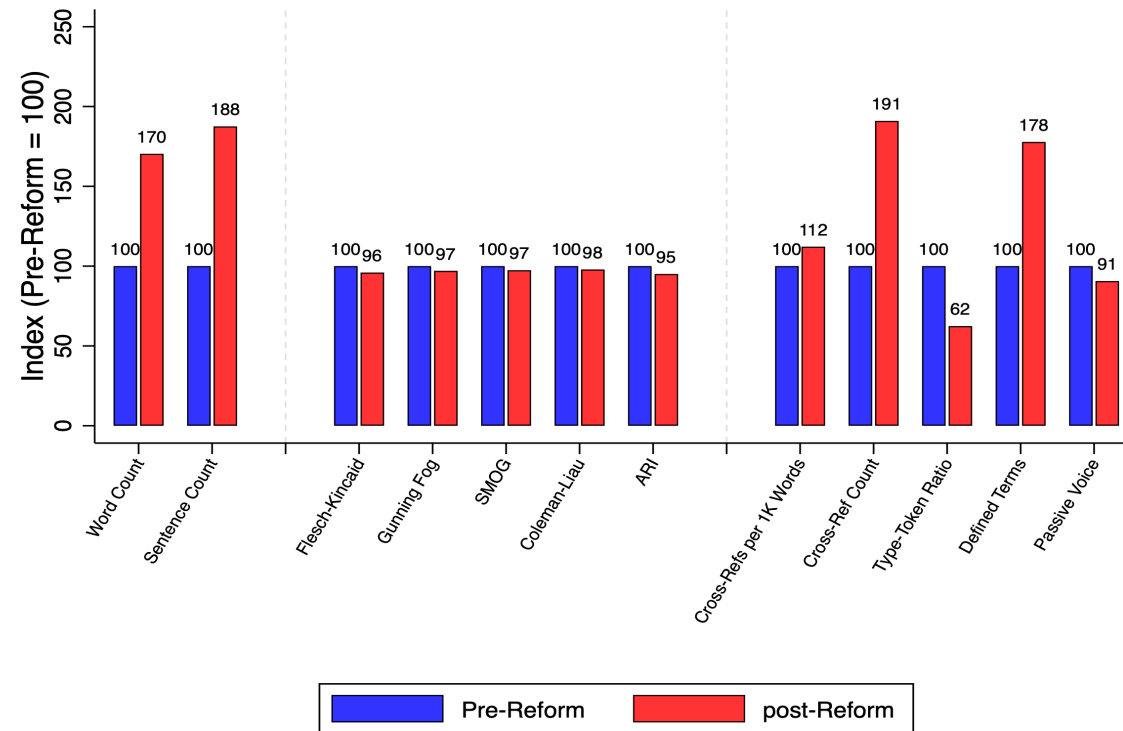


Figure 4: **Alternative Contagion-Free Deregulation Policy**

The figure presents combinations of the recurring cost θ and the switching cost η that preserve the model-implied pre-reform negative audit rate of 7.5%, holding the scale discount δ fixed at its post-reform level (-0.040). The red square indicates the pre-reform parameter estimates from Table A.7, Panel A: $(\theta, \eta) = (0.328, 6.408)$. The blue asterisk indicates the post-reform estimates: $(\theta, \eta) = (0.643, 4.678)$. The top horizontal scale expresses the recurring cost in dollars terms per negative opinion, under the normalization described in Appendix A.6.5. We also mark the distance from the post-reform estimates to the frontier, holding the post-reform switching cost fixed: reducing the recurring cost by the equivalent of roughly \$12,000 per negative opinion (about \$25 million per year) would restore the pre-reform negative audit rate. See Section 5 and Appendix A.6.5.

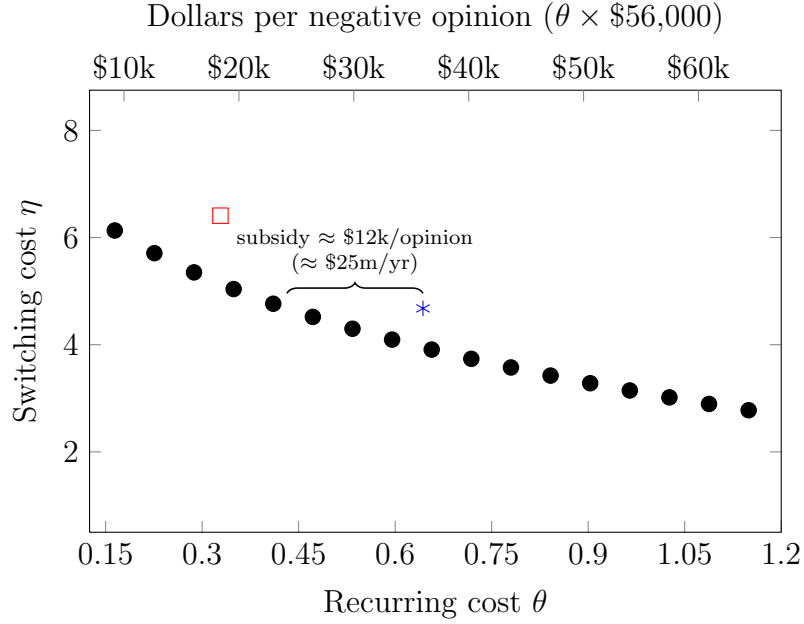


Table 1: **Summary Statistics**

Panel A. Clients. Summary statistics at the client level, 1998–2022. *Amount (\$1,000)* (*Amount (2023 USD; \$1,000)*) is the amount of federal awards spent by the client in thousand USD (adjusted for inflation). Relationship variables include the pair survival rate, client switch rate, and auditor and client attrition rates. Opinion variables include the unqualified opinion rate, the composite negative audit rate, and its decomposition into financial-statement (FS) and major-program (MP) components, as well as qualified and adverse opinion rates. Additional opinion-related variables include the non-negative disclaimer rate. Findings variables include any internal-control deficiency (material weakness or significant deficiency), questioned costs, material noncompliance, any award-level finding, and the share of awards with findings. All indicator variables are scaled 0–100. See [Section 2.2](#).

	Obs.	Mean	SD	Median	Min	Max
Amount (\$1,000)	993,414	7,236	19,762	1,973	310	218,017
Amount (2023 USD; \$1,000)	993,414	9,831	26,069	2,724	518	245,128
Relations:						
Pair Survival Rate	946,253	71.9	44.9	100.0	0.0	100.0
Client Switch Rate	898,031	11.0	31.3	0.0	0.0	100.0
Auditor Attrition Rate	981,329	10.1	30.2	0.0	0.0	100.0
Client Attrition Rate	898,031	8.3	27.6	0.0	0.0	100.0
Opinions:						
Unqualified Opinion	993,414	90.5	29.4	100.0	0.0	100.0
Negative Audit Rate	993,414	9.4	29.1	0.0	0.0	100.0
Non-Negative Disclaimer	993,414	0.2	4.3	0.0	0.0	100.0
Qualified Opinion	993,414	8.0	27.1	0.0	0.0	100.0
Adverse Opinion	993,414	1.7	12.9	0.0	0.0	100.0
Negative FS	993,414	6.3	24.3	0.0	0.0	100.0
Negative MP	993,414	3.6	18.8	0.0	0.0	100.0
Opinion Persistence:						
Pr(Negative Prior Negative)	59,943	68.3	46.5	100.0	0.0	100.0
Pr(Negative Prior Positive)	620,532	2.5	15.7	0.0	0.0	100.0
Findings:						
Any IC Deficiency	993,414	26.9	44.4	0.0	0.0	100.0
Questioned Costs	991,234	4.9	21.6	0.0	0.0	100.0
Material Noncompliance	993,094	4.5	20.7	0.0	0.0	100.0
Any Finding	991,072	7.5	26.3	0.0	0.0	100.0
Share of Awards with Findings	991,072	3.1	14.2	0.0	0.0	100.0

Panel B. Auditors. Summary statistics at the auditor×year level, 1998–2022. *Total Amount (\$1,000)* (*Total Amount (2023 USD; \$1,000)*) is the sum of expenditures (in thousands) across all of an auditor’s clients in a given year (adjusted for inflation). *Number of Clients* is the number of clients an auditor has in a given year. *Survival Rate* is the likelihood an auditor has at least one client next year. Audit outcome variables include whether the auditor issued any negative audit, the count of negative audits, and the negative audit rate across clients. See [Section 2.2](#).

	Obs.	Mean	SD	Median	Min	Max
Total Amount (\$1,000)	216,758	33,162	297,071	3,035	0	25,870,324
Total Amount (2023 USD; \$1,000)	216,758	45,056	391,179	4,408	0	28,379,516
Number of Clients	216,758	5	16	1	0	1,264
Survival Rate	189,494	57.2	49.5	100.0	0.0	100.0
Audit Outcomes:						
Any Negative Audit	216,758	16.0	36.7	0.0	0.0	100.0
Number of Negative Audits	216,758	0	3	0	0	212
Negative Audit Rate	189,494	8.7	24.0	0.0	0.0	100.0

Table 2: **Main Result: Contagion Effect of Deregulation**

We estimate Equation (1) using the baseline sample and estimation window described in Section 3.1. The dependent variable, *Negative Audit*, equals 100 if the auditor provided a negative opinion (qualified or adverse) of the client, and 0 otherwise. $Treated = 1$ if the auditor has at least one client with expenditures below \$750,000 in the last pre-reform period; $Treated = 0$ if the auditor had no such below-threshold clients but had at least one client with expenditures between \$1 million and \$1.5 million in the same period. $Post = 1$ for post-reform periods. Standard errors are heteroskedasticity-robust and reported in parentheses. Pre-Reform Mean reports the mean of the dependent variable in the period immediately preceding the reform, and Effect (%) reports the coefficient on $Treated \times Post$ scaled by that mean. The last row reports Oster (2019) δ^* , computed using column (1) as the baseline and $R^2_{max} = 1.3 \times R^2_{long}$. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. See Section 3.2.

Outcome:	Negative Audit		
	(1)	(2)	(3)
Treated $\times$ Post	-1.359*** (0.272)	-1.749*** (0.278)	-1.852*** (0.282)
Period FE	Yes	Yes	Yes
Auditee FE	Yes	Yes	No
Auditor FE	No	Yes	No
Pair FE	No	No	Yes
Pre-Reform Mean	8.14	8.14	8.08
Effect (%)	-16.69	-21.50	-22.93
Observations	201,068	200,986	194,404
R ²	0.566	0.588	0.616
Oster δ^*	—	-34.99	-13.78

Table 3: **Impact on Auditor Portfolio**

This table estimates Equation (3) at the auditor level, collapsing the client $\times$ period panel to auditor $\times$ period panel. The sample conditions on auditors who remain active after the reform. Panel A captures audit outcomes: an indicator for issuing at least one negative audit (column 1), the count of negative audits (column 2), and the average negative audit rate across clients (column 3). Panel B captures client outflow using backward-looking measures: *Lost Any Client* equals 100 if at least one auditor-client pair from the previous period is absent in the current period (column 4); *Fewer Clients* equals 100 if the auditor's total client count declined (column 5); *Num Clients* is the total number of clients per auditor per event-time period, unconditional (column 6) and conditional on auditors who experienced a net decline (column 7). Standard errors are heteroskedasticity-robust and reported in parentheses. Pre-Reform Mean reports the dependent variable's mean in the period immediately preceding the reform, and Effect (%) scales the coefficient by that mean. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. See Section 3.3.

Panel:	Panel A: Audit Outcomes			Panel B: Client Outflow			
Outcome:	Any Neg. Audit	# Neg. Audits	Neg. Rate	Lost Any Client	Fewer Clients	Num Clients	Num Clients if Fewer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated $\times$ Post	-3.094*** (1.062)	-0.629*** (0.081)	-1.061** (0.424)	13.706*** (1.667)	19.800*** (1.787)	-1.009*** (0.188)	-2.182*** (0.339)
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Auditor FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Reform Mean	29.08	1.22	7.12	58.83	28.46	15.33	17.28
Effect (%)	-10.64	-51.57	-14.90	23.30	69.57	-6.58	-12.63
Observations	14,607	14,607	14,607	14,042	14,042	14,042	4,477
R ²	0.609	0.721	0.574	0.417	0.258	0.975	0.986

Table 4: **Heterogeneous Treatment Effect**

We estimate Equation (1) separately on high- and low-exposure subsamples. All split variables are measured in the period immediately preceding the reform. In each block, treated auditors are split at the median and the full control group is retained in both subsamples. Columns 1–4 split by auditor exposure to the demand shock: the count of at-risk clients (columns 1–2) and the at-risk ratio (columns 3–4). Columns 5–8 split by pre-reform relationship stability: market competition (columns 5–6) and whether the client ever switched auditors during 2012–2015 (columns 7–8). The p -value row tests equality of the $Treated \times Post$ coefficient across subsamples. Standard errors are heteroskedasticity-robust and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. See Section 4.1 and Section 4.2.

Outcome:		Negative Audit						
Group:		Demand Shock Exposure				Pair Stability		
Split:	At-Risk Count		At-Risk Ratio		Market Competition		Never Switched	
	High	Low	High	Low	Stable	Less Stable	Stable	Less Stable
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated $\times$ Post	-2.283*** (0.286)	0.487 (0.355)	-4.346*** (0.482)	-1.627*** (0.283)	-2.374*** (0.333)	0.223 (0.512)	-2.758*** (0.373)	-1.026** (0.434)
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Reform Mean	8.22	6.74	7.88	7.89	8.30	7.14	8.10	8.02
Effect (%)	-27.8	7.2	-55.2	-20.6	-28.6	3.1	-34.1	-12.8
p -value	<0.001		<0.001		<0.001		0.046	
Observations	165,886	48,389	37,197	177,078	132,792	35,108	67,599	100,424
R ²	0.610	0.645	0.654	0.610	0.585	0.605	0.586	0.590

Table 5: **Contagion Effect and External Monitoring**

We estimate Equation (4), augmenting the baseline specification with a triple interaction between *Treated*, *Post*, and *Form990_{c,t}* (an indicator equal to one if auditee *c* has a matching IRS Form 990 filing in period *t*). The coefficient β_1 captures the contagion effect for non-filers; β_2 measures how the effect differs for filers. The row $\beta_1 + \beta_2$ reports the implied net effect for filers, with *p*-values from a test of $H_0: \beta_1 + \beta_2 = 0$ in brackets. Standard errors are heteroskedasticity-robust and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. See Section 4.3.

Outcome:	Negative Audit		
	(1)	(2)	(3)
Treated $\times$ Post (β_1)	-2.568*** (0.425)	-3.365*** (0.437)	-3.568*** (0.445)
Treated $\times$ Post $\times$ Form 990 Filer (β_2)	2.488*** (0.552)	3.728*** (0.566)	3.952*** (0.574)
Treated $\times$ Form 990 Filer	1.294*** (0.497)	-0.595 (0.704)	-0.769 (0.718)
Post $\times$ Form 990 Filer	1.787*** (0.519)	0.664 (0.530)	0.813 (0.537)
Form 990 Filer	-2.333*** (0.509)	-1.078* (0.647)	-0.648 (0.647)
Period FE	Yes	Yes	Yes
Auditee FE	Yes	Yes	No
Auditor FE	No	Yes	No
Pair FE	No	No	Yes
$\beta_1 + \beta_2$	-0.08	0.36	0.38
<i>p</i> -value	[0.82]	[0.31]	[0.29]
Observations	201,068	200,986	194,404
R ²	0.567	0.590	0.618

Table 6: **Alternative Explanations**

Panel A: Sample Composition. We re-estimate Equation (1) under cumulative sample restrictions. Column (1) reproduces the baseline. Column (2) restricts to auditors observed in at least one pre- and one post-reform period; column (3) further requires all six baseline periods. Column (4) additionally restricts to clients observed both before and after the reform. Column (5) retains only clients whose federal expenditures exceed \$750,000 in every period. Each column is a subset of the previous one. Standard errors are heteroskedasticity-robust and reported in parentheses. p -values in brackets test equality with column (1). *, **, and *** denote significance at the 10%, 5%, and 1% levels. See Section 4.6.

Outcome:		Negative Audit			
Sample:	Baseline	+ Balanced Auditors (Pre & post)	+ Balanced Auditors (All Periods)	+ Balanced Clients	+ Always Above \$750K
	(1)	(2)	(3)	(4)	(5)
Treated $\times$ Post	-1.852*** (0.282)	-1.852*** (0.281)	-1.864*** (0.300)	-1.864*** (0.298)	-1.250*** (0.357)
Pair FE	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes
t-1 Mean	8.08	8.11	8.38	8.51	7.44
Effect (%)	-22.9	-22.8	-22.3	-21.9	-16.8
p -value (= col. 1)		[1.000]	[0.977]	[0.976]	[0.186]
Observations	194,404	192,407	181,096	172,565	111,006
R ²	0.616	0.615	0.612	0.605	0.585

Panel B: Size-Based Noncompliance. We estimate Equation (5) in the period 2004–2012, before the baseline estimation window. *Above \$750K* equals one if the client’s federal expenditures exceed \$750,000; this threshold had no legal significance during 2004–2012. Standard errors are heteroskedasticity-robust and reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels. See Section 4.6.

Outcome:	Negative Audit		
	(1)	(2)	(3)
Above \$750K	0.407** (0.165)	0.142 (0.167)	0.140 (0.176)
Auditee FE	Yes	Yes	No
Auditor FE	No	Yes	No
Pair FE	No	No	Yes
Period FE	Yes	Yes	Yes
Observations	355,422	351,134	323,668
R ²	0.553	0.593	0.663

Internet Appendix

A.1 Variable Dictionary

In this section we define all variables used in the empirical analysis, organized by category. Outcome variables and subsample indicators are scaled 0–100; treatment and design indicators are 0/1. Structural parameters are defined in [Appendix A.6.1](#).

Variable	Sample	Description
Treatment and Design		
Treated	Client	Equals 1 if the auditor had at least one client with expenditures below \$750K in the last pre-reform period; equals 0 if the auditor had no such clients but had at least one client between \$1M and \$1.5M.
Post	Client	Equals 1 for all post-reform periods.
Above \$750K	Client	Equals 1 if the client’s federal expenditures exceed \$750K in that year.
JustAbove750	Client	Equals 1 if the client’s federal expenditures fall in a \$50K bin immediately above \$750K, and 0 if they fall between the top of that bin and \$1.1M. Each column of Table A.6 uses a different bin.
Exit	Client	Equals 100 if the entity undergoes a Single Audit in event-time period t but not in $t + 1$, including each entity’s final observed period.
Financial and Size		
Amount (\$1,000)	Client	Federal expenditures of the client in thousands of nominal USD.
Amount (2023 USD; \$1,000)	Client	Federal expenditures of the client in thousands of constant 2023 USD.
Total Amount (\$1,000)	Auditor	Sum of federal expenditures across all of an auditor’s clients in a given year, in thousands of nominal USD.
Total Amount (2023 USD; \$1,000)	Auditor	Sum of federal expenditures across all clients, in thousands of constant 2023 USD.
Number of Clients	Auditor	Number of distinct clients audited in a given year.
Opinions		
Negative Audit (%)	Client	Equals 100 if the auditor issued a qualified or adverse opinion on either the financial statements or any major program; 0 otherwise. Primary outcome variable.
Unqualified Opinion (%)	Client	Equals 100 if the client received unqualified (clean) opinions on both financial statements and all major programs.
Negative FS (%)	Client	Equals 100 if the auditor issued a qualified or adverse opinion on the financial statements.
Qualified FS (%)	Client	Equals 100 if the financial-statement opinion is qualified.

Continued on next page

Variable	Sample	Description
Adverse FS (%)	Client	Equals 100 if the financial-statement opinion is adverse.
Negative MP (%)	Client	Equals 100 if the auditor issued a qualified or adverse opinion on any major program.
Qualified MP (%)	Client	Equals 100 if any major-program opinion is qualified.
Adverse MP (%)	Client	Equals 100 if any major-program opinion is adverse.
Non-Negative Disclaimer (%)	Client	Equals 100 if the auditor issued a disclaimer opinion that is not classified as negative.
Qualified Opinion (%)	Client	Equals 100 if the auditor issued a qualified opinion on either the financial statements or any major program.
Adverse Opinion (%)	Client	Equals 100 if the auditor issued an adverse opinion on either the financial statements or any major program.
Opinion Persistence		
Pr(Negative Prior Negative)	Client	Equals 100 if the pair's next audit is negative, among pair-years whose current audit is negative. Defined for pairs observed in consecutive years.
Pr(Negative Prior Positive)	Client	Equals 100 if the pair's next audit is negative, among pair-years whose current audit is not negative. Defined for pairs observed in consecutive years.
Findings and Internal Controls		
Any IC Deficiency (%)	Client	Equals 100 if the audit identified a material weakness or significant deficiency in internal controls (from either financial statements or major programs).
Questioned Costs (%)	Client	Equals 100 if the audit reported questioned costs for any program.
Material Noncompliance (%)	Client	Equals 100 if the audit identified material noncompliance with federal program requirements.
Any Finding (%)	Client	Equals 100 if the entity has at least one finding across all its federal awards.
Number of Findings	Client	Count of findings across all federal awards in a given year.
Share of Awards with Findings (%)	Client	Percentage of the entity's federal awards that have at least one finding.
Relations		
Pair Survival Rate (%)	Client	Equals 100 if the auditor-client pair observed at time t is also observed at $t + 1$.
Client Switch Rate (%)	Client	Equals 100 if the client changes auditor at $t + 1$, conditional on both client and auditor existing at $t + 1$.
Auditor Attrition Rate (%)	Client	Equals 100 if the auditor ceases to appear in the sample at $t + 1$.
Client Attrition Rate (%)	Client	Equals 100 if the client ceases to appear in the sample at $t + 1$.
Survival Rate (%)	Auditor	Equals 100 if the auditor has at least one client at $t + 1$.
Survived (%)	Client	Equals 100 if the auditor-client relationship survived to $t + 1$.

Continued on next page

Variable	Sample	Description
Exited (%)	Client	Equals 100 if the client exited the Single Audit sample at $t + 1$.
Switched (%)	Client	Equals 100 if the client switched to a different auditor at $t + 1$, conditional on remaining in the sample.
Auditor-Level Outcomes		
Any Negative Audit (%)	Auditor	Equals 100 if the auditor issued at least one negative audit in the event-time period.
Number of Negative Audits	Auditor	Count of negative audits issued by the auditor in the event-time period.
Negative Audit Rate (%)	Auditor	Average negative audit rate across all of the auditor’s clients in the event-time period.
Lost Any Client (%)	Auditor	Equals 100 if at least one auditor–client pair from the previous period is absent in the current period. Backward-looking.
Fewer Clients (%)	Auditor	Equals 100 if the auditor’s total client count declined relative to the previous period.
Num Clients if Fewer	Auditor	Number of clients for auditors who experienced a net decline.
Heterogeneity		
High At-Risk Client Count (%)	Client	Equals 100 if the auditor’s count of clients with expenditures between \$500K and \$750K in the last pre-reform period is above the median among treated auditors. Set to 0 for controls.
High At-Risk Ratio (%)	Client	Equals 100 if the share of at-risk clients relative to total clients in the last pre-reform period is above the median among treated auditors. Set to 0 for controls.
Market Competition (%)	Client	Equals 100 if the entity’s state is in the bottom five deciles of the ratio of unique auditors to unique auditees in the last pre-reform period, indicating a less competitive market. Assigned at the state level and held fixed thereafter.
Never Switched (%)	Client	Equals 100 if the client did not change auditors at any point during 2012–2015. Clients with no switching history over that window are coded 0.
Form 990 Filer (%)	Client	Equals 100 if the auditee has a matching IRS Form 990 filing for the same fiscal year.
Pre-Reform Auditor Characteristics (Robustness)		
# Clients	Auditor	Number of unique clients served by the auditor in the last pre-reform period.
Avg Amount (\$M)	Auditor	Average federal award amount per client in the last pre-reform period, in millions of USD.
Portfolio (\$M)	Auditor	Total dollar value of all audited awards in the last pre-reform period, in millions of USD.
Tenure (Years)	Auditor	Average auditor–client relationship length at the reform date, in years.

A.2 The Single Audit Reform

A.2.1 The Process

The Single Audit has been the principal mechanism for overseeing recipients of federal awards since the Single Audit Act of 1984. Over the following three decades, the audit framework evolved through periodic threshold adjustments and procedural updates, most notably under OMB Circular A-133 (first issued in 1990, revised in 1997 and 2003). By the early 2010s, eight separate OMB circulars governed different aspects of federal award administration.¹ In October 2011, the President issued a memorandum directing OMB to streamline the administrative requirements for federal awards.² OMB established the Council on Financial Assistance Reform (COFAR) to coordinate the effort, and published an Advance Notice of Proposed Guidance in February 2012 (77 FR 11778). The Advance Notice solicited comments on a range of potential changes, including raising the Single Audit threshold to \$1,000,000 – double the then-current level of \$500,000. OMB published a Proposed Guidance in February 2013 (78 FR 7282), which formally proposed a threshold of \$750,000 – a compromise between the status quo and the \$1,000,000 initially floated. After a second round of public comments, the final guidance was published on December 26, 2013 (78 FR 78590). The new audit requirements, including the \$750,000 Single Audit threshold, apply to fiscal years beginning on or after December 26, 2014 (2 CFR §200.110). Most entities transitioned to the new regime for fiscal years starting in calendar year 2015, with the first wave of post-reform audits completed in 2016.

A.2.2 Public Comments

The reform was preceded by a public comment period, in two rounds: first through the 2012 Advance Notice of Proposed Guidance, and then through the 2013 Proposed Guidance, each on its own docket.³ Together the two dockets drew 664 public comments. The Federal Register preamble summarizing this record describes a tension among the stakeholders: many smaller recipients and their representatives favored a higher threshold to reduce compliance costs, while inspectors general and oversight bodies warned that raising it too far would strip audit coverage from a large number of entities.⁴

The comments offer a contemporaneous, independent read on how the market perceived the reform. We therefore collected the full record ourselves, downloading every public submission from both dockets on [regulations.gov](https://www.regulations.gov) and extracting the text of each letter. We classify commenters by organization type and report the composition in [Figure A.4](#). Universities and research institutions filed the most, but the audit profession engaged directly and on

¹The eight circulars were: A-21 (cost principles for educational institutions), A-87 (cost principles for state and local governments), A-110 (administrative requirements for institutions of higher education and nonprofits), A-122 (cost principles for nonprofits), A-89 (federal domestic assistance program information), A-102 (grants to state and local governments), A-133 (audits of states, local governments, and nonprofits), and A-50 (audit follow-up).

²Presidential Memorandum on Administrative Flexibility, Lower Costs, and Better Results for Federal Awards, M-12-01 (Oct. 2011).

³Dockets OMB-2012-0002 (advance notice) and OMB-2013-0001 (proposed rule).

⁴78 FR 78590, at 78593–78594.

the record: the American Institute of Certified Public Accountants (AICPA); state CPA societies in Pennsylvania, Oregon, Illinois, and Virginia; national and regional Single Audit firms including CliftonLarsonAllen, McGladrey, Crowe Horwath, Moss Adams, and Wipfli; and state auditors and federal inspectors general. That the profession engaged the rule at all confirms that the reform was material to the audit market.

We then read the audit community letters in full to record how they viewed the reform. Auditors expected the threshold increase to reduce the number of Single Audits performed – the same demand contraction our design exploits. The Oregon Society of CPAs wrote that raising the threshold “would reduce the number of Single Audits performed” (OMB-2012-0002-0322); the Pennsylvania Institute of CPAs observed that the change would “eliminate the single audit requirement” for affected entities and “could potentially eliminate the need for numerous audits” (OMB-2012-0002-0143); and the AICPA acknowledged that the proposal “would result in fewer grantees being subject to a single audit” (OMB-2012-0002-0250). Crowe Horwath noted that the change would allow grantees “that currently expend over \$500,000 in federal expenditures... to no longer be subject to single audit requirements” (OMB-2013-0001-0248).

The letters also make clear that the engagements at stake are a meaningful source of revenue for the firms that perform them. McGladrey reported that it “currently performs in excess of 600 audits in accordance with OMB Circular A-133 annually” (OMB-2013-0001-0112); Crowe Horwath “currently conducts approximately 200 single audits per year, with such entities . . . receiving . . . in excess of \$5.6 billion in federal funds” (OMB-2013-0001-0248); and Moss Adams performs “approximately 170 audits annually” (OMB-2013-0001-0288). The AICPA placed the whole market at “over 40,000 single audits performed each year by thousands of CPA firms” (OMB-2012-0002-0250). Government audit offices, which face no commercial demand and perform Single Audits under a public mandate, expected no effect: the National Association of State Auditors, Comptrollers and Treasurers wrote that “[b]ecause we conduct our Single Audit at the statewide level, we do not anticipate that these lower thresholds would affect our audit resources” (OMB-2012-0002-0291). We draw on this distinction when we exclude government audit offices from the exposure analysis ([Section 4.1](#)).

Finally, the government’s own paperwork estimates corroborate the scale of the demand contraction. Under the Paperwork Reduction Act, agencies must file public estimates of the burden imposed by federal information collections, including Single Audit reporting. At the proposal stage, OMB projected that the reform would relieve roughly 5,000 entities of the Single Audit requirement while leaving more than 99% of federal award dollars under audit coverage.⁵ In its post-reform filing, OMB recorded a reduction of 3,400 Single Audit submissions per year attributed to the threshold increase.⁶

A.2.3 Key Changes

The reform made changes along several dimensions. We organize them by the expected direction of their effect on compliance burden.

⁵See [Journal of Accountancy](#) (Dec. 2013).

⁶OMB control number 0348-0057, 2016 information-collection review, which attributes the adjustment to the threshold rising “from \$500,000 to \$750,000 starting fiscal year 2015.”

Provisions expected to reduce burden. The most salient change was the increase in the Single Audit threshold from \$500,000 to \$750,000 in annual federal expenditures. OMB estimated that this would exempt 14% of the pre-reform audit population. For entities that remained subject to audit, the reform lowered the minimum percentage of federal expenditures that must be tested as major programs: from 50% to 25% for low-risk auditees and from 50% to 40% for others (2 CFR §200.518). The reform also introduced a de minimis indirect cost rate for any entity that had not previously negotiated a rate with a cognizant agency (2 CFR §200.414(f)). Finally, the consolidation of eight circulars into a single codified framework (2 CFR Part 200) was intended to unify the regulatory structure.

Provisions expected to increase burden. The reform mandated that entities maintain effective internal controls consistent with guidance from the Committee of Sponsoring Organizations (COSO), formalizing what had previously been a general expectation (2 CFR §200.303). Pass-through entities – recipients that distribute federal funds to subrecipients – were subject to expanded monitoring obligations, including pre-award risk assessments, during-award monitoring of subrecipient activities, and timely management decisions on subrecipient audit findings (2 CFR §200.332). The reform also introduced new requirements for conflict-of-interest disclosures (§200.112) and safeguards for personally identifiable information (§200.303(e)), neither of which had direct analogues in the pre-reform circulars.

Provisions with ambiguous effects. The restructuring of the Type A program classification falls under this category. Under A-133, a program was classified as Type A if it exceeded \$300,000 or 3% of total expenditures, whichever was larger. Under the new regulatory framework, the threshold was set on a tiered schedule: \$750,000 for entities with total expenditures up to \$25 million, declining as a percentage of total expenditures for larger entities (2 CFR §200.518). For small and mid-sized entities, the new thresholds were higher, meaning fewer programs were classified as Type A and subject to mandatory testing. For very large entities, the percentage-based thresholds could classify more programs as Type A. Similarly, as we show below, the separation of the Compliance Supplement from the core circular simplified the regulatory text but added a cross-referencing layer: auditors now needed to consult two documents rather than one to determine the applicable audit procedures for a given program.

A.2.4 Quantifying the Costs of Compliance

The Uniform Guidance was explicitly motivated by the desire to reduce administrative burden and streamline compliance.⁷ As mentioned earlier, a careful reading of pre- and post-reform regulatory requirements reveals ambiguous impact. To assess this quantitatively, we compare the textual properties of the Compliance Supplement in its pre-reform (March 2014) and post-reform (June 2016) editions. This is the practitioner-facing document that specifies compliance requirements and audit procedures for federal awards. We focus on Part 3 of the Supplement (“Compliance Requirements”), which contains the generic audit framework

⁷78 FR 78590 (“The reforms...aim to reduce administrative burden...while also guarding against waste, fraud, and abuse.”).

applicable to all federal programs.⁸ Figure 3 summarizes the comparison, with all metrics normalized to the pre-reform value of 100.

The post-reform Part 3 text is substantially longer: 69,672 words across 3,520 sentences, compared to 40,900 words across 1,877 sentences in the 2014 edition – index values of 170 and 188, respectively. The number of sentences grew faster than the number of words, indicating that average sentence length fell slightly (from 21.8 to 19.8 words per sentence). The document expanded primarily by adding more provisions, and those provisions tend to use somewhat shorter sentences.

We compute five standard indices that estimate the education level required to comprehend a text. The *Flesch-Kincaid Grade Level* combines average sentence length and average syllables per word, mapping the result to a U.S. grade level; higher scores indicate more difficult text (Kincaid et al., 1975). The *Gunning Fog Index* weights the proportion of “complex words” (three or more syllables) alongside sentence length; it was designed for business and technical writing. The *SMOG Index* estimates the years of education needed to understand a text based on the count of polysyllabic words per sample of sentences (Mc Laughlin, 1969). The *Coleman-Liau Index* uses character counts rather than syllable counts, while the *Automated Readability Index (ARI)* substitutes characters-per-word for syllables-per-word. All five indices show a modest decline but remain firmly in the “difficult” to “very difficult” range, consistent with specialized legal and regulatory text.

Finally, we examine structural complexity measures. The absolute count of cross-references – instances of terms such as “§,” “section,” “Part,” “subpart,” “paragraph,” “see,” and “refer to” – nearly doubles, from 367 to 701. Normalized by document length, the density of cross-references rises from 9.0 to 10.1 per 1,000 words. This means that, rather than presenting audit requirements as a self-contained document, the post-reform edition directs the reader to external provisions such as program-specific sections and GAGAS. In contrast, the *type-token ratio* – the number of distinct word types divided by total word tokens – falls sharply. This indicates that the expanded post-reform text reuses existing vocabulary rather than introducing proportionally more distinct terms. It suggests that the post-reform text adds provisions that share a common vocabulary. We also find that the number of *defined terms* increases from 9 to 16, reflecting the broader scope of the post-reform framework. The passive voice usage declines modestly (from 37.7% to 34.2%), indicating that the drafting style shifted slightly toward more active prose.

In sum, the textual analysis suggests that the reform did not simplify the regulatory environment for auditors. While sentence-level readability improved marginally, the post-reform Compliance Supplement is 70% longer, nearly twice as cross-referenced, and relies on shorter but more numerous provisions. These properties are consistent with the view that compliance burden – at least for entities that remained subject to audit – was not meaningfully reduced by the reform.

⁸In an untabulated test, we repeat the analysis using the full framework text (Parts 1–3) and obtain qualitatively identical results. Part 4 (program-specific requirements) is excluded because its content varies across vintages for reasons unrelated to the reform as programs are added, removed, and updated annually.

A.3 Data Appendix

A.3.1 Single Audit: Dataset Construction

This section documents the construction of our primary dataset from raw FAC files. The FAC publishes audit data in several interlinked files. Our pipeline proceeds in three steps: (i) import and harmonize the two FAC vintages, (ii) construct audit-opinion indicators, and (iii) build unique identifiers for auditors and auditees.

Step 1: Import and harmonize FAC vintages. The FAC data are published in two distinct vintages. The *historic* vintage (GAO-1) covers audit years 1998–2015 and is organized as a set of annual CSV files (one folder per year), each containing tables such as `ELECAUDITHEADER` (general audit information) and `ELECAUDITS` (federal award details). The *redesigned* vintage (GAO-2) covers 2016 onward and is published as a single CSV per table (`general.csv`, `federal_awards.csv`, `findings.csv`). The two vintages differ in variable names, date formats, and the encoding of indicator variables. We import each vintage separately, rename all variables to a common schema, and standardize formats: dates are converted to Stata date values, indicators are recoded from string (“Y”/“N”, “True”/“False”) to numeric (1/0), and string variables are trimmed and lowered. We then append the two vintages into a single *General* file covering 1998–2022.⁹

Step 2: Construct audit-opinion indicators. The FAC reports audit opinions as string codes. In GAO-1, the *General* file contains opinions on both financial statements (`typereport_fs`) and major programs (`typereport_mp`), encoded as single-letter codes: “U” (unqualified), “Q” (qualified), “A” (adverse), “D” (disclaimer), and “S” (not GAAP, financial statements only).¹⁰ In GAO-2, the *General* file contains financial-statement opinions (`gaap_results`) using full-text labels (e.g., “unmodified_opinion”, “qualified_opinion”), but major-program opinions are *not* reported in the *General* file. Instead, they are stored at the award level in the *Federal Awards* file (`audit_report_type`), available only for awards classified as major programs. We aggregate these to the audit level following the FAC’s aggregation rule: an audit receives an unmodified (“U”) major-program opinion if and only if all underlying awards received unmodified opinions; if any award is qualified, adverse, or disclaimer, the corresponding letters appear in the aggregate code but “U” does not.¹¹ We parse both vintages into a common set of indicator variables: *adverse*, *qualified*, *disclaimer*, and *unqualified*, separately for financial statements (FS) and major programs (MP). Our primary outcome variable, *negative*, equals one if the audit received a qualified or adverse opinion on either FS or MP. In GAO-2, finding-level indicators (material weakness, significant deficiency, questioned costs) are likewise stored in the *Federal Awards* and *Findings* files

⁹A small number of observations with string fields exceeding 100 characters are dropped as data-extraction errors. This affects fewer than 50 records across the full sample.

¹⁰The code “S” is not defined in the official FAC documentation. We confirm its meaning by cross-referencing the GAO-1 codes with the explicit labels used in GAO-2.

¹¹This aggregation rule also applies to GAO-1, where major-program opinions are available both at the audit level (in `ELECAUDITHEADER`) and at the award level (in `ELECAUDITS`, from 2004 onward). We verify that the two sources are consistent.

rather than the *General* file. We aggregate these to the audit level and merge them back into the main panel.

Step 3: Build unique identifiers. Auditees are identified by their Employer Identification Number (EIN), which is consistently populated across years. Some large organizations have multiple divisions that share the same EIN but undergo separate audits; we aggregate these to the EIN level. We use EIN rather than legal name or address because it is stable over time, which is essential for constructing the panel. Auditors are also identified by EIN. Coverage is nearly complete from 2013 onward; in the earlier portion of the sample, some observations have missing auditor EINs. To fill these, we standardize firm names – removing punctuation, legal suffixes, and professional descriptors – and match standardized names to known EINs. Where exact name matches are unavailable, we apply fuzzy string matching (similarity threshold ≥ 0.90) and manually verify borderline cases. Residual unmatched auditors receive unique identifiers based on their standardized names.

Dataset Migration. An earlier draft of this paper relied on the Census Bureau version of the FAC data, which was the only version available at the time. In mid-2023, the GAO assumed responsibility for the FAC and replaced the Census files with two new vintages: GAO-1 (the *historic* archive, covering 1998–2015) and GAO-2 (the *redesigned* format, covering 2016 onward).¹² The current draft uses the GAO vintages exclusively. In addition to the dataset switch, the current draft incorporates substantial feedback from conferences and seminars, refines the empirical design, and includes a number of new tests. Reassuringly, we verified that the underlying data are effectively interchangeable across the two sources. We restrict both datasets to audit years 1998–2022 and exclude audits accepted after August 1, 2023, since those could only appear in the GAO version. After transforming each dataset into a recipient $\times$ year panel (identified by EIN and audit year), the Census panel contains 966,637 observations and the GAO panel contains 966,573 – a difference of less than 0.01%. Merging the two panels yields 966,132 matched observations, with only 946 unmatched (505 Census-only, 441 GAO-only). In our main sample window of 2013–2018, the match rate exceeds 99.96%. For the matched observations, we evaluate the consistency of seven key variables: auditor identifiers (`cpa_ein`, `cpa_state`, `cpa_firm_name`), recipient identifiers (`entity_ein`, `entity_firm_name`, `entity_state`), and total federal expenditures. The overall consistency rate – defined as all seven variables being identical – is 98.3%, rising to 99.9% in the 2013–2018 window. Nearly all discrepancies trace to the auditor’s EIN field in pre-2013 data, where the two sources occasionally record different values. The key results underpinning the earlier draft – the reduction in negative audits among treated auditors, driven primarily by adverse opinions – are exactly reproduced in the GAO data.

¹²We downloaded the Census dataset from <https://facdissem.census.gov/main.aspx> in July 2023 (no longer available). We downloaded the GAO vintages from <https://www.fac.gov/data/download/> (valid as of October 27, 2025). GAO also provides a crosswalk linking the `report_id` identifiers in GAO-2 to the `dbkey` identifiers used in the Census and GAO-1 formats.

A.3.2 Additional Data Sources

We supplement the FAC data with two external sources. IRS Form 990 filings provide financial and organizational characteristics of nonprofit auditees, which we use to construct controls and explore heterogeneity. Audit Analytics provides auditor-level attributes and audit outcomes for publicly traded firms, which we use to benchmark patterns and validate our findings against a well-studied population.

Tax returns We use data from IRS Form 990 filings to obtain financial and organizational characteristics of nonprofit entities, including revenues, expenses, assets, and program composition. These data are constructed from publicly available digitized filings (e.g., IRS Statistics of Income and related processed datasets). We merge the Form 990 data to the Single Audit sample using Employer Identification Numbers (EINs), which serve as the primary unique identifier across both datasets. The merge is implemented at the organization-year level. When multiple filings exist within a fiscal year, we retain the latest submission. In the paper, the tax data are used to construct controls for organizational size and financial structure, and to explore heterogeneity in responses across nonprofits. Because EIN coverage is incomplete and some entities cannot be reliably matched, these variables are used in specifications where the merged sample is available, and results are interpreted accordingly.

Audit Analytics We use Audit Analytics to obtain basic information on auditor characteristics and audit outcomes. In the paper, these data are used primarily to complement our baseline measures with auditor-level attributes (e.g., auditor size and specialization) and to validate patterns in audit findings. Because coverage is limited to larger and more visible entities, these tests are used as supporting evidence rather than in the main specifications. The opinions data are restricted to 10-K filings and excludes supplementary opinions. To address duplicate observations, we retain the most recent filing for each company-fiscal year combination, prioritizing observations with the latest filing date and highest audit fees. The sample covers fiscal years 2001 onwards. Auditor retention is calculated as an indicator variable equal to one if a firm retains the same auditor as in the prior fiscal year, based on matching auditor foreign keys across consecutive years. Restatements are matched to firm-years using the end date of the restated period; adverse restatements are identified using the `RES_ADVERSE` flag, which indicates restatements that worsen previously reported financial figures rather than correct clerical errors or improve reported results. For the matched auditor sample, we merge the Audit Analytics data with the Single Audit data using exact name matches.

A.4 Related Literature

This paper connects several strands of work in finance and economics.

Threshold-based regulation. Our paper is directly relevant to the literature on threshold-based regulation, a ubiquitous instrument across financial regulation. For instance, [Ewens, Xiao, and Xu \(2024\)](#) show that firms manipulate their float to stay below SEC reporting

thresholds; [Iliev \(2010\)](#) exploits the SOX Section 404 compliance threshold to estimate the costs and benefits of internal-control certification; and [Alvero, Ando, and Xiao \(2023\)](#) apply bunching methods to the Dodd-Frank asset thresholds of \$10 billion and \$50 billion, revealing strategic size management by banks and estimating regulatory costs. Papers in this literature focus on how regulated entities respond to thresholds – for instance, by bunching below it or by benefiting from lower compliance costs. Our paper identifies a distinct consequence: threshold increases alter the competitive environment of the intermediaries who help firms comply with the regulation and who certify compliance to third parties, propagating effects to entities that remain above the threshold. This intermediary channel is potentially relevant wherever threshold-based regulation relies on shared compliance professionals – a condition satisfied by the vast majority of thresholds in the Code of Federal Regulations.¹³

Costs of deregulation. A growing literature measures the quantity of regulation and its economic costs ([Kalmenovitz, 2023](#); [Kalmenovitz, Lowry, and Volkova, 2025](#); [Donelson, Garfinkel, and Roudini, 2024](#); [Armstrong, Glaeser, and Hoopes, 2025](#)). These measurements are sometimes used to motivate the case for deregulation, and a complementary literature documents economic gains from reducing regulatory burdens ([Jayaratne and Strahan, 1996](#); [Ewens and Farre-Mensa, 2020](#)). Separate strands study how organizational features of the government shape regulatory outcomes: [Kalmenovitz \(2021\)](#) studies compensation incentives for financial regulators, [Chen, Hajda, and Kalmenovitz \(2024\)](#) examines how regulators escape pay-for-performance constraints, [Gopalan, Kalda, and Manela \(2021\)](#) documents how the geographic distribution of regulatory field offices affects bank leverage, and [Hirtle, Kovner, and Plosser \(2020\)](#) estimates the causal impact of supervisory intensity on bank performance. Our paper identifies a cost of deregulation that is absent from this literature: when oversight is delegated to private intermediaries, reducing the scope of regulation contracts demand for their services and weakens scrutiny of entities that remain regulated.

External monitoring and compliance professionals. A large literature in finance studies how market structure shapes the quality of information produced by private gatekeepers. [Becker and Milbourn \(2011\)](#) show that entry of a third major credit rating agency reduced the quality of ratings from incumbents. [Bolton, Freixas, and Shapiro \(2012\)](#) model how competition among rating agencies erodes information quality as issuers shop for favorable ratings, and [Opp, Opp, and Harris \(2013\)](#) show that regulation based on ratings can distort the incentive to rate accurately. [Kisgen and Strahan \(2010\)](#) provide empirical evidence that ratings-based regulations directly affect firms’ cost of capital. Beyond ratings, a parallel literature studies gatekeepers who inspect or certify the entities that pay them. Closer to our setting, a literature on auditor independence asks whether economic dependence on a client sways reporting: [DeAngelo \(1981\)](#) models the quasi-rents that make incumbents reluctant to lose clients, and [Reynolds and Francis \(2000\)](#), [DeFond, Raghunandan, and Subramanyam \(2002\)](#), and [Chung and Kallapur \(2003\)](#) test whether client importance and fee dependence affect audit opinions. Our paper extends this logic to a new channel: we show that a demand contraction caused by deregulation erodes monitoring quality by raising the value of retaining

¹³A registry of more than 50 such thresholds, together with a catalog of the academic literature studying specific regulatory thresholds, is available at jkalmenovitz.com.

remaining clients. Prior work compares monitors who differ in how much a single client matters to them; we instead use a single deregulatory shock that raises the value of every retained client at once.

Our empirical setting. The market for Single Audit services bridges the literatures on nonprofit finance and governance. [Keating et al. \(2005\)](#) documents compliance rates under the original framework, [Tassin, Waymire, and Hines \(2019\)](#) provides a historical account of Single Audit reforms, and [Cuneo, Leder-Luis, and Vannutelli \(2023\)](#) studies compliance and governance within the system. We use IRS Form 990 filings to construct measures of external monitoring, motivated by [Harris, Petrovits, and Yetman \(2015\)](#) and [Harris, Petrovits, and Yetman \(2024\)](#), who show that donors respond to governance quality and fraud disclosures revealed through Form 990. These filings have also been used to study nonprofit financing more broadly ([Matveyev and Zeume, 2025](#)).

A.5 Strategic Sorting

In this section, we directly examine whether the reform induced strategic sorting around the \$750,000 threshold. Organizations may respond to the reform by reducing their federal expenditures to fall below the threshold and avoid Single Audit requirements. If such behavior differs systematically between treatment and control groups, it could mechanically improve audit outcomes by selectively removing certain types of clients. To test for this possibility, we estimate the following triple-difference specification:

$$\begin{aligned} Exit_{c,a,t+1} = & \alpha + \beta_1 JustAbove750_{c,t} + \beta_2 (JustAbove750_{c,t} \times Treated_a) \\ & + \beta_3 (Treated_a \times Post_t) + \beta_4 (JustAbove750_{c,t} \times Post_t) \\ & + \beta_5 (JustAbove750_{c,t} \times Post_t \times Treated_a) \\ & + \gamma Post_t + \lambda + \varepsilon_{c,a,t} \end{aligned} \tag{A.1}$$

The dependent variable $Exit_{c,a,t+1}$ equals 100 if client c is observed in the Single Audit data at time t but not at $t + 1$. $Post_t$ equals one for fiscal years ending on or after December 25, 2015. The sample is restricted to entities with expenditures between \$750,000 and \$1.1 million, and we partition this range into six \$50,000 bins. $JustAbove750_{c,t}$ identifies entities in each bin just above the threshold, with entities above that bin serving as the reference group. Auditor fixed effects absorb time-invariant differences across auditors, and period fixed effects absorb aggregate shocks common to all observations.¹⁴ The coefficient of interest, β_5 , captures whether entities just above the threshold exit differentially for treated auditors after the reform. If treated auditors' clients are more likely to engage in strategic exit, we would expect $\beta_5 > 0$.

[Table A.6](#), Panel A, reports the results. In the narrowest bin (\$750K–\$800K), β_5 is positive and marginally significant (5.5 percentage points, $p < 0.10$), suggesting some differential exit among treated auditors' clients closest to the threshold. However, this effect is localized:

¹⁴We do not include pair or auditee fixed effects in these specifications, since the exit outcome is terminal: each entity contributes at most one $Exit = 100$ observation before leaving the sample.

across the remaining four bins, β_5 is statistically insignificant with no systematic pattern. Moreover, the main result is robust to excluding all near-threshold clients entirely (column 5 of [Table 6](#), which restricts to entities with expenditures above \$750,000 in every period).

We also examine whether the reform led to a general increase in exits, irrespective of treatment status, by estimating a pooled difference-in-differences specification:

$$\begin{aligned} Exit_{c,a,t+1} = & \alpha + \beta_1 JustAbove750_{c,t} + \beta_4 (JustAbove750_{c,t} \times Post_t) \\ & + \gamma Post_t + \lambda + \varepsilon_{c,a,t} \end{aligned} \quad (A.2)$$

[Table A.6](#), Panel B, reports the results. Exit rates increase sharply among entities closest to the threshold: in the \$750K–\$800K bin, β_4 is 8.1 percentage points (34% of the pre-reform mean); in the \$800K–\$850K bin, the effect is 3.8 percentage points (18%). The pattern attenuates with distance from the threshold but remains statistically significant through the \$900K–\$950K bin (2.2 percentage points), becoming insignificant only in the \$950K–\$1.0M bin.

Taken together, these results suggest that the reform induced a general increase in exits among entities near the threshold, concentrated in the bins closest to \$750,000. The differential exit between treated and control auditors is limited to the narrowest bin and is marginally significant; it disappears entirely in wider bins. Strategic sorting therefore cannot account for the contagion effect documented in the main analysis, which is robust to excluding all near-threshold entities ([Table 6](#), column 5).

A.6 Structural Model

This appendix presents the structural model summarized in [Section 5](#). The baseline model holds client behavior fixed via a private information shock, and attributes changes in audit outcomes entirely to shifts in the auditor’s cost structure. We recover that structure from observed audit decisions, separately for treated and control auditors and for the pre- and post-reform periods; decompose the decline in negative opinions into the contribution of the cost shift and of the changing environment; and evaluate counterfactual policies that would neutralize the contagion effect. We also present three extensions of the baseline model.

A.6.1 Model setup

Time is discrete and the time horizon is infinite. A risk-neutral auditor, who discounts cash flows using discount rate $\rho = 0.95$,¹⁵ is engaged by the client to audit its expenditures of federal awards. Each period, the auditor collects fees proportional to those expenditures, p_t , expressed in inflation-adjusted 2023 USD:

$$\pi(p_t) = \phi \cdot p_t.$$

The auditor decides whether to issue a positive (unqualified) or negative (qualified or

¹⁵As is standard in dynamic discrete choice, the discount factor is calibrated rather than estimated; it is not identified from choice data alone ([Magnac and Thesmar, 2002](#)).

adverse) opinion about the client’s use of federal awards. The decision is denoted by d_t , which equals one if the audit is negative and zero if positive.

Issuing a negative opinion involves a *recurring cost*, captured by $\theta > 0$. It is a fixed cost, paid each time a negative opinion is issued, irrespective of past audit outcomes or the client’s expenditures.¹⁶ It reflects the required effort and the risk of revenue loss: upon receiving a negative audit, the client may seek a different auditor or lose federal funding, either way terminating the relationship. While we cannot observe effort, we document the substantial revenue risk in Table A.5.

Beyond the recurring cost, the auditor pays a one-time cost when switching from a positive opinion to a negative opinion. This *switching cost* is captured by $\eta > 0$. It reflects the additional effort required to install rigorous auditing practices and the additional risk of losing the client. Intuitively, if the auditor already issued a negative opinion in the previous period, then the rigorous practices have already been adopted, and by revealed preference we learn that the risk of client termination is smaller. The switching cost is consistent with the patterns we observe in the data. If the auditor issued a positive opinion at time t , the probability of a negative opinion at time $t + 1$ is 2.5%. Conversely, if the audit outcome at time t was negative, there is a 68.3% chance of a negative audit at time $t + 1$. This persistence suggests that transitioning from positive to negative triggers a switching cost on top of the recurring cost.

The third ingredient is a *scale discount*: the cost of initiating a negative audit decreases with auditor size. The scale discount is represented as $\delta \cdot a_t$, where δ is the structural parameter and a_t is auditor size, measured as the total federal expenditures across all of the auditor’s clients.¹⁷ If $\delta < 0$, larger auditors face lower switching costs as they enjoy economies of scale: they can spread the costs of transitioning to different audit practices across a broader client base.¹⁸

In sum, we model the auditor’s cost function as:

$$c(a_t, d_t, d_{t-1}) = \underbrace{\theta \cdot d_t}_{\text{recurring cost}} + \underbrace{\eta \cdot (1 - d_{t-1}) \cdot d_t}_{\text{switching cost}} + \underbrace{\delta \cdot a_t \cdot (1 - d_{t-1}) \cdot d_t}_{\text{scale discount}}.$$

The state variables are auditor size (a_t) and the current audit outcome which was determined in the previous period (d_{t-1}). Three cases arise. If the auditor issues a positive opinion ($d_t = 0$), there is no cost. If the auditor continues a negative opinion ($d_{t-1} = 1$, $d_t = 1$), only the recurring cost applies. If the auditor switches from positive to negative ($d_{t-1} = 0$, $d_t = 1$), the full switching cost plus the recurring cost apply, partially offset by the scale discount.

¹⁶Appendix A.6.6 estimates a version in which the recurring cost depends on the client’s federal expenditures; the baseline conclusions are unchanged.

¹⁷Measuring auditor size by the logarithm of the number of audited clients instead of dollar volume yields similar estimates.

¹⁸In untabulated regressions we find that large auditors are more likely to initiate a negative opinion after a positive one, consistent with $\delta < 0$. Among continuing relationships whose prior opinion was positive, auditors with an above-median number of clients issue a negative in 2.6% of cases, versus 1.9% for smaller auditors, a statistically significant gap that holds whether we measure size by client count or by total federal awards audited.

A.6.2 The auditor's problem

The auditor chooses audit outcomes to maximize intertemporal utility. The one-period utility function is:

$$u_t(x_t, d_t) = \pi(p_t) - c(a_t, d_t, d_{t-1}),$$

where $x_t \equiv \{p_t, a_t, d_{t-1}\}$ is the vector of state variables observable to both the auditor and the econometrician: client's federal expenditures, auditor portfolio size, and the previous audit outcome.

The auditor's intertemporal utility also depends on a state variable $\varepsilon_t(d_t)$, observed by the auditor but not by the econometrician. Its value depends on the auditor's opinion. This is a standard feature of dynamic discrete choice models (Rust, 1994). Without the random component, the auditor's decision would be fully determined by the observable state variables, which is unrealistic. One can think of $\varepsilon(d_t)$ as a private information shock: for instance, if the client was non-compliant, the auditor receives extra utility from issuing a negative opinion ($\varepsilon_t(1) > 0$) and disutility from issuing a positive one ($\varepsilon_t(0) < 0$), reflecting reputational risk or the chance of being caught by a government watchdog. Client behavior is thus present in the model, but enters through the auditor's information set rather than as an observable state. In Appendix A.6.6, we go one step further and make one component of client behavior explicit, letting the survival of the relationship itself depend on the auditor's opinion.¹⁹

The auditor's total utility is:

$$u_t(x_t, d_t) + \varepsilon_t(d_t) = \begin{cases} \phi p_t - \theta - (\eta + \delta a_t)(1 - d_{t-1}) + \varepsilon_t(1) & \text{if } d_t = 1, \\ \phi p_t + \varepsilon_t(0) & \text{if } d_t = 0. \end{cases}$$

A.6.3 Model solution

Every period, the auditor chooses d_t to maximize the expected present value of future utility. The Bellman equation is:

$$V_t(x_t, \varepsilon_t) = \max_{d_t} \left\{ u_t(x_t, d_t) + \varepsilon_t(d_t) + \rho \mathbb{E} \left[V_{t+1}(x_{t+1}, \varepsilon_{t+1} \mid x_t, \varepsilon_t, d_t) \right] \right\},$$

where $V_t(x_t, \varepsilon_t)$ is the expected discounted utility in state (x_t, ε_t) . Since ε_t is unobservable, we integrate it out following Arcidiacono and Ellickson (2011) to obtain the ex-ante value function:

$$\bar{V}_t(x_t) \equiv \int V_t(x_t, \varepsilon_t) g(\varepsilon_t) d\varepsilon_t.$$

¹⁹The identifying assumption is that the distribution of this signal is unchanged across the pre- and post-reform periods, so that shifts in observed audit behavior are attributed to the auditor's cost structure. Several pieces of evidence support this assumption. First, we find no meaningful differential exit of noncompliant clients around the reform (Section 4.6). Second, when estimating the model, we find that the shift in parameter across periods is absent among control auditors; a market-wide change in compliance would move both groups. Third, we estimate a similar parameter shift in two versions of the model: when the cost of a negative opinion depends on award size, and when there are two types of clients (Appendix A.6.6). This indicates that the effect is not driven by changes in client expenditures or in the composition of the pool of clients.

The conditional value function – the present value of choosing d_t and behaving optimally from period $t + 1$ onward – is:

$$v_t(x_t, d_t) \equiv u_t(x_t, d_t) + \rho \int \bar{V}_{t+1}(x_{t+1}) f(x_{t+1} | x_t, d_t) dx_{t+1}.$$

The decision rule $d_t^*(x_t, \varepsilon_t)$ solves:

$$d_t^*(x_t, \varepsilon_t) = \arg \max_{d_t} [v_t(x_t, d_t) + \varepsilon_t(d_t)]. \quad (\text{A.3})$$

The probability that the auditor chooses d_t given state x_t is found by integrating the decision rule over the regions of $\varepsilon_t(d_t)$ for which $d_t^*(x_t, \varepsilon_t) = d_t$:

$$\mathbb{P}_t(d_t | x_t) = \int \mathbb{I} \left\{ \arg \max_{d_t} [v_t(x_t, d_t) + \varepsilon_t(d_t)] = d_t \right\} g(\varepsilon_t) d\varepsilon_t.$$

We assume that ε is independent and identically distributed extreme value type I with scale parameter σ . The conditional choice probability then admits the closed-form logit solution:

$$\mathbb{P}_t(d_t | x_t) = \frac{\exp[v_t(x_t, d_t)]}{\exp[v_t(x_t, 0)] + \exp[v_t(x_t, 1)]}. \quad (\text{A.4})$$

A.6.4 Estimation and Identification

Estimation. We proceed in two steps. First, we estimate the transition matrix of auditor size a_t from the data, using a kernel density estimator with 25 evenly spaced points of support. Auditor size is measured as the logarithm of the sum of real federal expenditures across the auditor’s clients, winsorized at the 1st and 99th percentiles. Second, we estimate the preference parameters $\{\sigma, \phi, \theta, \eta, \delta\}$: the scale parameter of the unobserved utility (σ), auditor fees (ϕ), recurring cost (θ), switching cost (η), and scale discount (δ). Since the audit decision is based on comparing the utility of a positive opinion to that of a negative opinion, the comparison is unaffected by rescaling utilities or adding a constant. We therefore normalize profits to zero ($\phi = 0$) and scale the remaining structural parameters by the scale parameter of the private information shock σ .²⁰ With $\phi = 0$, the client’s expenditures do not enter the auditor’s effective decision problem, and the state reduces to $x_t \equiv \{a_t, d_{t-1}\}$. We estimate $\{\theta, \eta, \delta\}$ using the nested fixed point (NFXP) method of Rust (1987, 1994): we solve the decision problem in Equation (A.3) and use its solution to obtain the conditional choice probabilities in Equation (A.4), which form the likelihood function for maximum partial likelihood estimation.

Identification. The model has three structural parameters: the recurring cost θ , the switching cost η , and the scale discount δ . As with any estimation procedure, the effectiveness of NFXP hinges on whether the data-generating process contains enough information to

²⁰As explained in He and Whited (2025), scaling parameters by σ is equivalent to scaling utility. The parameter estimates thus capture their importance relative to the dispersion of the unobserved shock, rather than their objective value.

identify the model’s parameters. In our setting, identification is obtained by requiring that these parameters influence the auditor’s decision to issue a negative opinion, and that they do so in a distinctive way.

First, the recurring cost θ is identified from the overall rate of negative opinions. Mechanically, the probability of a negative opinion declines in θ , since θ is incurred whenever $d_t = 1$ regardless of the previous outcome or auditor size. Because θ enters the cost function symmetrically across states, it shifts $P(d_t = 1 \mid d_{t-1} = 0)$ and $P(d_t = 1 \mid d_{t-1} = 1)$ in the same direction, leaving the gap between them largely unchanged. Next, both η and δ govern the cost of *initiating* a negative opinion (switching from $d_{t-1} = 0$ to $d_t = 1$) but not the cost of *continuing* to issue one. They are therefore jointly identified from the difference in choice probabilities across the two lagged states: Figure A.1 shows that there is a 65-percentage-point persistence gap between $P(d_t = 1 \mid d_{t-1} = 0)$ and $P(d_t = 1 \mid d_{t-1} = 1)$. However, the two parameters are separately identified because they interact differently with auditor size a_t . The switching cost η is a fixed cost that does not depend on auditor size a_t ; increasing η therefore widens the persistence gap uniformly. The scale discount δ , by contrast, is multiplied by a_t ; when $\delta < 0$, the upward-sloping relation between $P(d_t = 1 \mid d_{t-1} = 0)$ and auditor size steepens, while $P(d_t = 1 \mid d_{t-1} = 1)$ remains flat (Figure A.1). This differential size gradient is uniquely attributable to δ .

In sum, θ governs the level of negative audits, η the persistence gap between states, and δ the size gradient in switching behavior.

A.6.5 Results

Parameter Estimates. Table A.7, Panel A, presents the parameter estimates and their standard errors.²¹ The first row reports the full-sample estimates. The recurring cost θ is positive and statistically significant, confirming that issuing a negative opinion is costly in terms of effort and client-retention risk. The switching cost η is also positive and significant, and substantially larger in magnitude, indicating that the cost of initiating a negative opinion, rather than the cost of maintaining one, is the dominant force shaping audit decisions.²² The scale discount δ is negative and significant, meaning that larger auditors face lower switching costs. These full-sample estimates form the baseline for the counterfactual exercises below.

The second block of Panel A compares pre-reform and post-reform estimates, pooling all auditors.²³ When comparing across subsamples, we account for potential differences in the scale of the private information shock using the procedure of Swait and Louviere (1993). The Swait-Louviere test rejects the null of equal parameters across subsamples ($p < 0.001$). The recurring cost θ nearly doubled (from 0.328 to 0.643), meaning the baseline penalty for issuing a negative opinion rose sharply – consistent with the increased value of retaining remaining clients. The switching cost η fell by roughly one-quarter (from 6.408 to 4.678), which on its own would make negative opinions easier to initiate, but this is more than offset by the other shifts. The scale discount δ shrank by more than half (from -0.108 to -0.040),

²¹Standard errors treat each auditor–client pair as an independent unit and are computed from the outer product of pair-level scores, the likelihood analog of clustering by pair.

²²In principle, η could partly reflect persistent heterogeneity in client compliance rather than a switching cost; we address this possibility in Appendix A.6.6.

²³We use the same structure described in Section 3.1.

meaning large auditors lost much of their cost advantage. These shifts have offsetting effects on audit stringency – the higher recurring cost directly discourages negative opinions, while the lower switching cost makes it cheaper to transition to a negative opinion – but the net effect is a decline in the model-implied rate of negative audits.

Next, we estimate the model separately for treated and control auditors, before and after the reform. The results are in the bottom two blocks of [Table A.7](#), Panel A. For treated auditors, the recurring cost rises from 0.311 to 0.686 and the Swait-Louviere test overwhelmingly rejects parameter equality across periods ($LR = 350.7$, $p < 0.001$); for control auditors, the estimates move little and equality cannot be rejected ($LR = 3.42$, $p = 0.18$). The difference in the recurring cost increase across the two groups is itself significant ($t \approx 4.7$). The structural shift is thus a treatment effect more than a market-wide trend. This exercise confirms that the decline in negative audits is a treatment effect on treated, because any time-series change around the reform – such as different auditing standards – would have shifted the parameters in the control group as well.

Finally, we decompose the model-implied decline in negative audits. Moving from the pre-reform to the post-reform cost parameters, while holding the pre-reform environment fixed, lowers the negative audit rate from 7.5% to 6.5%, while allowing the environment to change as well lowers it to 4.5%. The cost structure shift alone thus accounts for roughly a third of the model-implied decline, with the remaining two-thirds operating through changes in the distribution of state variables. Within the cost shift, the higher recurring cost and the flattened scale discount each subtract about 1.6 percentage points from the rate while the lower switching cost adds back 2.2; the three forces partially offset, which is why the net effect coexists with large individual parameter movements.

Model Fit. [Table A.7](#), Panel B, assesses model fit by comparing data-implied and model-implied probabilities of negative audits across combinations of the two state variables: previous opinion and auditor size.²⁴ The model closely reproduces the key patterns. If the auditor issued a positive opinion today, the probability of a negative opinion tomorrow is 2.6% in the estimation sample and 2.5% in the model; if the auditor issued a negative opinion, the probability of continuing negative is 67.8% in the data and 67.9% in the model. Both probabilities vary with auditor size, and the model captures this heterogeneity as well.²⁵

Inspecting the Mechanism. [Figure A.1](#) plots the model-implied probability of a negative audit as a function of auditor size, separately for each lagged outcome. In Panel A ($d_{t-1} = 0$), the probability of initiating a negative opinion rises with auditor size, reflecting the scale

²⁴Model-implied probabilities are computed exactly: starting each auditor–client pair at its observed initial state, we propagate the model’s distribution over states forward through the estimated choice probabilities and the size transition matrix for the pair’s observed panel length, and aggregate the implied probabilities across pairs.

²⁵The full-sample negative audit rate in the structural estimation is 8.1%, lower than the 9.4% reported in [Table 1](#). The difference reflects sample restrictions required for estimation, most importantly the need to observe a lagged outcome for each auditor–client pair. Because negative opinions predict relationship dissolution ([Table A.5](#)), conditioning on pair survival mechanically selects for pairs with better audit histories. Re-estimating the baseline difference-in-differences on this restricted subsample yields a treatment effect of 24%, confirming that the sample selection does not attenuate the contagion effect.

discount δ : larger auditors face lower switching costs. In Panel B ($d_{t-1} = 1$), the continuation probability declines mildly with auditor size. This is a forward-looking effect: by switching back to a positive opinion today, the auditor forgoes the option value of continuation but gains flexibility, since re-entering the negative state later is cheaper when the net switching cost ($\eta + \delta \cdot a_t$) is smaller. Larger auditors, with a more favorable re-entry cost, are therefore more willing to revert.

Figure A.2 shows how each structural parameter affects the likelihood of negative audits in isolation. Starting from the pre-reform baseline, we vary one parameter at a time while holding the others constant. Two key insights emerge. First, negative audits decrease with both the recurring cost θ and the switching cost η , but the slope of η is substantially steeper: switching costs have a larger quantitative impact on auditor decisions. Second, negative audits increase with the scale discount δ , since a larger discount offsets the costs of issuing a negative opinion. Given these sensitivities, the substantial shifts documented in Table A.7 explain the decline in negative audits after the reform. The recurring cost rose, the switching cost fell, and the scale discount shrank – jointly pushing in the direction of fewer negative audits. The cost shifts that drive the contagion – a higher baseline penalty for negative opinions and the erosion of large auditors’ cost advantage – are consistent with the demand contraction documented in the reduced-form analysis.

Counterfactual Policies. We assess the quantitative importance of each parameter through two sets of counterfactual experiments, using the full-sample estimates from Table A.7. In the first set, we compute elasticities: the percentage change in the negative audit rate from a 1% change in each parameter. A 1% increase in η reduces negative audits by 2.1%, compared to 0.6% for θ and 0.5% for δ . This confirms that switching cost is an important driver of auditor leniency. Multiplying these percentage changes by the average annual volume of awards covered by audits in the estimation sample (\$304 billion), a 1% increase in η corresponds to \$540m in annual award volume reclassified from negative to positive opinions, compared to \$160m for θ and \$120m for δ . Thus, although the stakes per auditor could seem modest, each opinion certifies millions of dollars of federal spending, so small shifts in individual audit costs translate into large changes in the volume of awards under rigorous oversight. In the second set, we shut down each parameter entirely. Setting $\eta = 0$ (no switching cost) raises the negative audit rate to 39.5%; setting $\theta = 0$ (no recurring cost) raises it to 19.0%; and setting $\delta = 0$ (no scale discount) lowers it to 5.7%. These magnitudes underscore that the switching cost is the key barrier to initiating negative audits – without it, auditors would issue negative opinions roughly 40% of the time.

We use the model to characterize contagion-free deregulation reforms. We ask: holding the scale discount δ at its post-reform level (-0.040), what combinations of θ and η would preserve the pre-reform average negative audit rate? Figure 4 plots this frontier. The actual post-reform parameters (blue star) lie above the frontier, confirming that the combined cost shift was too large to sustain pre-reform audit rates. The actual pre-reform parameters (red square) are also above the frontier, but the pre-reform negative audit rate was higher because the substantially larger scale discount ($\delta = -0.108$) offset the cost burden. For example, reversing roughly two-thirds of the increase in the recurring cost – while leaving the post-reform switching cost and scale discount unchanged – would fully restore the pre-reform

negative audit rate.²⁶

Finally, [Figure A.3](#) holds θ fixed and varies (η, δ) ; the results highlight that achieving any given negative audit rate requires trading off switching costs against the scale discount. In principle, a contagion-free reform could pair the threshold increase with policies that reduce θ – for instance, a government-sponsored insurance program compensating auditors who lose clients following a negative opinion. Alternatively, policies that restore the scale discount, such as subsidized consolidation of small audit practices, could offset the contagion effect through the δ channel.

Estimated dollar values. We can express these estimates in dollars. In the extension in [Appendix A.6.6](#), the auditor earns a fee flow of one utility unit per period, so one unit corresponds to the average annual Single Audit fee of approximately \$56,000.²⁷ Under this setting, the reform raised the effective price of issuing a negative opinion from roughly \$18,000 to \$36,000 per opinion per year. The counterfactual above then prices the fix: reversing two-thirds of the increase in the recurring cost amounts to a subsidy of roughly \$12,000 per negative opinion. Auditors in the estimation sample issued about 2,150 negative opinions per year before the reform, and applying the restored pre-reform rate to post-reform audit volumes implies a similar count. Thus, restoring the pre-reform negative audit rate would cost approximately \$25 million per year – less than 0.01% of the \$304 billion in federal awards these audits cover.²⁸

A.6.6 Extensions

Client Loss Risk. In the baseline model, the recurring cost θ is interpreted as a risk of losing the client due to negative audit. We develop a version that makes this risk explicit, letting the relationship terminate stochastically: after the auditor issues opinion d , the pair survives into the next period with probability $s(d)$. The auditor earns a fee flow of one unit per period while the relationship lasts, and termination is an absorbing state with value zero.²⁹ The conditional value functions become $U_d(x, \ell) = u_d(x, \ell) + \rho s(d) \mathbb{E}[V(x', d) | x]$, so choosing a negative opinion endogenously forfeits $\rho[s(0) - s(1)] \mathbb{E}[V']$ of continuation value. We calibrate $s(d)$ from the observed one-year pair survival rates by audit opinion,³⁰ separately for the pre- and post-reform windows. We then re-estimate θ twice, once where pair survival rate depends on audit outcomes and once where that rate is fixed.³¹ By comparing the two

²⁶Re-running this exercise on treated auditors only yields similar results. We prefer the pooled estimates because the policy instruments considered are market-wide and cannot condition on reform exposure.

²⁷This number is derived from the limited sample of Single Audit recipients who also file Form 990.

²⁸The switching cost translates to several annual fees under the same calibration; we treat this as an upper bound, since part of η may reflect persistent heterogeneity in compliance. See [Appendix A.6.6](#).

²⁹Normalizing the continuation value of a terminated relationship to zero follows the convention in exit and default models (e.g., [Magnac and Thesmar, 2002](#); [Kalouptsi, Scott, and Souza-Rodrigues, 2021](#)).

³⁰We report two sets of results, one based on raw survival rates and another after absorbing entity-state and period fixed effects.

³¹The two specifications describe the same choices. The difference is that the extension reallocates the recurring cost into a part explained by the measured client-loss risk and a residual, and the split is determined by the observed survival gap $s(0) - s(1)$.

estimates, we can split the recurring cost into a component explained by the client-loss risk and a residual that the priced risk does not capture.

Table A.8, Panel A, reports the result. In the table, *total* captures the recurring cost in a version where negative audit does not affect termination risk; *priced risk* is the forfeited relationship value induced by a negative opinion, in a version where such opinion increases the termination risk;³² *residual* is the recurring cost re-estimated in that version, capturing the deterrent that remains once the forfeited value is charged; and $\mathbb{E}[V]$ is the model-implied value of a surviving relationship – the expected present value of its future fee flow, computed under the estimated choice probabilities with continuation discounted by $\rho s(d)$, and expressed in annual fees (since the fee flow is one unit per period). We find that the reform increased the recurring cost by 0.45, consistent with the baseline model. Moreover, the increase in priced client-loss risk accounts for roughly 40% of the rising costs. This means that auditor leniency is driven, to first order, by the present value of the client relationship at risk.³³ Finally, the model-implied value of a surviving relationship rises from 3.6 to 4.8 annual fees. This means that the client relationship an auditor risks by issuing a negative opinion became more valuable after the reform: relationships survive longer, so the same negative opinion jeopardizes a larger stream of future fees.³⁴

Client Size. The baseline model excludes client characteristics from the cost function: with fees normalized, the client’s expenditures drop out of the auditor’s problem. In practice, though, opinions may respond to the client’s scale. We therefore estimate an extended model in which the recurring cost depends on the client’s expenditures, $\theta(p_t) = \theta_0 + \theta_P (p_t - \bar{p})$, with p_t the logarithm of the client’s real expenditures and $\bar{p}$ its sample mean. Client expenditures become a second observable state, with a joint transition process combining the auditor-size kernel with the within-pair expenditure transition; the specification nests the baseline at $\theta_P = 0$. Table A.8, Panel B, reports the results. Two findings stand out. First, the key pattern remains intact: following the reform, at the mean expenditure level the recurring cost doubles ($0.332 \rightarrow 0.663$) and the switching cost falls. Second, in the pooled sample θ_P is essentially zero, which explains why the one-dimensional model fits well. However, the pooled zero masks an informative sign reversal. Before the reform, clients with larger expenditures carried a slightly higher cost of a negative opinion ($\theta_P = 0.019$); after the reform the slope turns sharply negative ($\theta_P = -0.082$). In other words, post-reform leniency concentrates on *small* clients. This is consistent with our preferred interpretation: the potential loss of small, near-threshold clients motivated auditors to become more lenient.

³²The price is $\rho[s(0) - s(1)] \mathbb{E}[V']$, reflecting the additional probability that the relationship ends after a negative opinion, multiplied by the discounted value of the relationship the auditor would otherwise keep.

³³The post-reform level reinforces this point: once the risk is priced, the residual recurring cost falls from 0.53 to 0.03, meaning that nearly the entire post-reform deterrent is the relationship value at risk. Before the reform there is little to decompose – the total is only 0.08 – and the priced risk of 0.26 exceeds it, leaving a slightly negative residual.

³⁴Note that the level estimates in this model and the baseline model are not directly comparable, because when the relationship can be terminated, continuation values are discounted by $\rho s(d)$ rather than ρ , which shifts the level of the recurring cost.

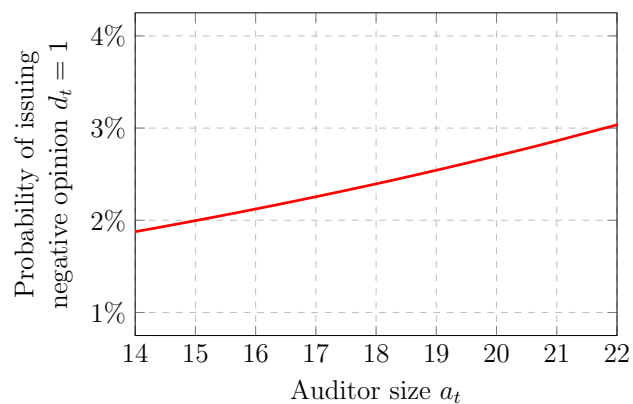
Heterogeneous Clients. We noted above that audit outcomes are persistent. This persistence may be driven by high switching costs, but it can also reflect permanent differences across clients, whereby some are more prone to negative opinions. To account for this possibility, we estimate a finite-mixture version of the model with two types of clients, $k \in \{1, 2\}$. They differ only in the recurring cost, which is θ_1 for the low-cost type and $\theta_1 + \mu$ for the high-cost type; all other model features are common across types. Because the partial likelihood conditions on each pair’s first observed outcome d_{i0} , we follow Wooldridge (2005) and let the type probability depend on that initial condition, $\Pr(k = 2 \mid i) = \Lambda(\alpha_0 + \alpha_1 d_{i0})$. The pair likelihood is $L_i = \sum_k \pi_k(d_{i0}) \prod_t P_k(d_{it} \mid x_{it}, d_{i,t-1})$, estimated by direct maximum likelihood with multiple starting values; standard errors are clustered by pair.

Table A.8, Panel C, reports the results. The switching cost η remains large and precisely estimated, albeit smaller than the baseline model. The recurring cost θ rises after the reform by approximately 0.3 for both types of clients, an order of magnitude matching the baseline model estimates. We also find that the reform’s impact is concentrated in the treatment group and absent from the control group. In particular, a likelihood-ratio test of parameter equality across the pre- and post-reform periods rejects overwhelmingly for treated auditors ($LR = 580.5$, $p < 0.001$) but not for control auditors ($LR = 12.0$, $p = 0.06$), where the six-parameter joint test is estimated on a much smaller sample. Thus, this version of the model leaves the substantive conclusions unchanged: the switching cost remains large and the recurring cost still rises across the reform within each type.

Figure A.1: **Policy Functions**

The figures show the model-implied probability of issuing a negative opinion tomorrow (d_t), as a function of two state variables: the audit outcome today (d_{t-1}) and the auditor size based on federal expenditures across the auditor's clients (a_t). The parameter values are from the full-sample estimates in [Table A.7](#), Panel A. Panel A corresponds to the auditor issuing a positive opinion today; Panel B corresponds to the auditor issuing a negative opinion today. See [Appendix A.6.5](#).

Panel A: Positive opinion issued today $d_{t-1} = 0$



Panel B: Negative opinion issued today $d_{t-1} = 1$

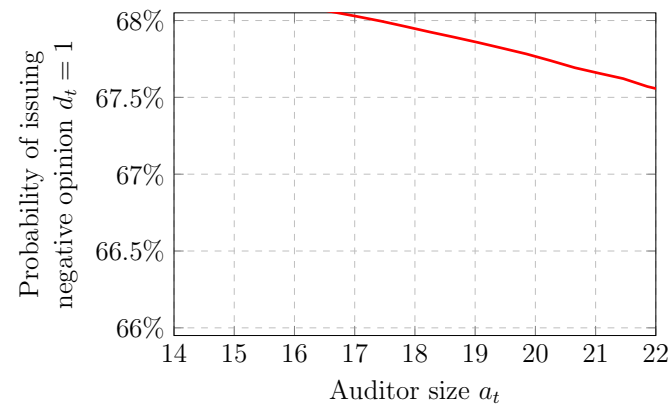


Figure A.2: **Comparative Statics**

The figure shows how the probability of a negative audit responds to each structural parameter in isolation. For each parameter, we compute the negative audit probability when the parameter changes by the amount shown on the x -axis (percentage change relative to the baseline), while keeping all other parameters at their baseline values. The coordinate (0%, 8.3%) corresponds to the baseline scenario. Panel A varies the recurring cost θ , Panel B varies the switching cost η , and Panel C varies the scale discount δ . See [Appendix A.6.5](#).

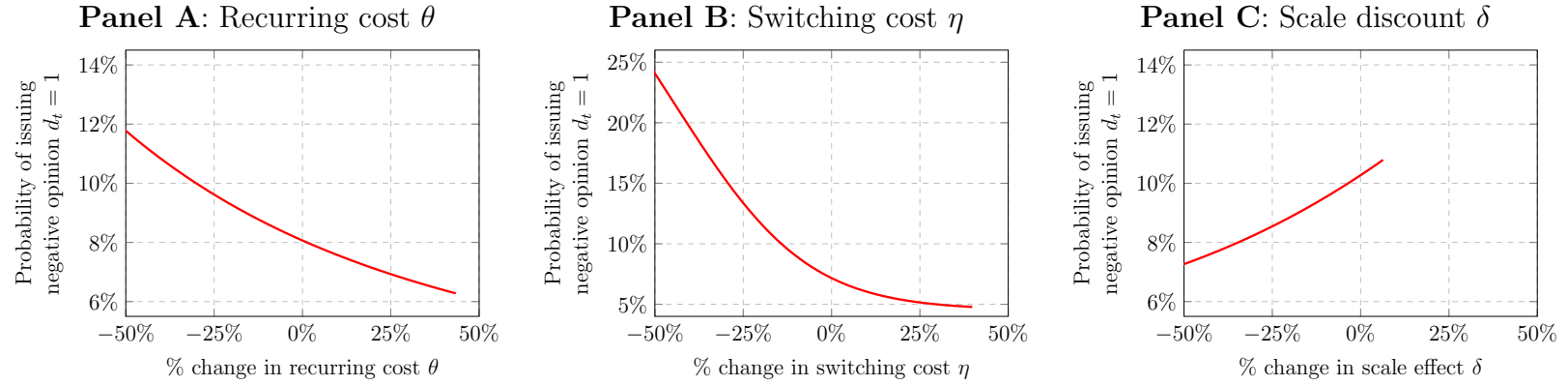


Figure A.3: **Alternative Deregulation Policies: Additional**

The figure presents combinations of the switching cost η and the scale discount δ that achieve the model-implied post-reform negative audit rate of 4.5%, holding the recurring cost θ fixed. The black dots use $\theta = 0.643$ (the post-reform estimate); the red squares use $\theta = 0.328$ (the pre-reform estimate). The blue asterisk indicates the actual post-reform parameter values: $(\eta, \delta) = (4.678, -0.040)$. See [Section 5](#) and [Appendix A.6.5](#).

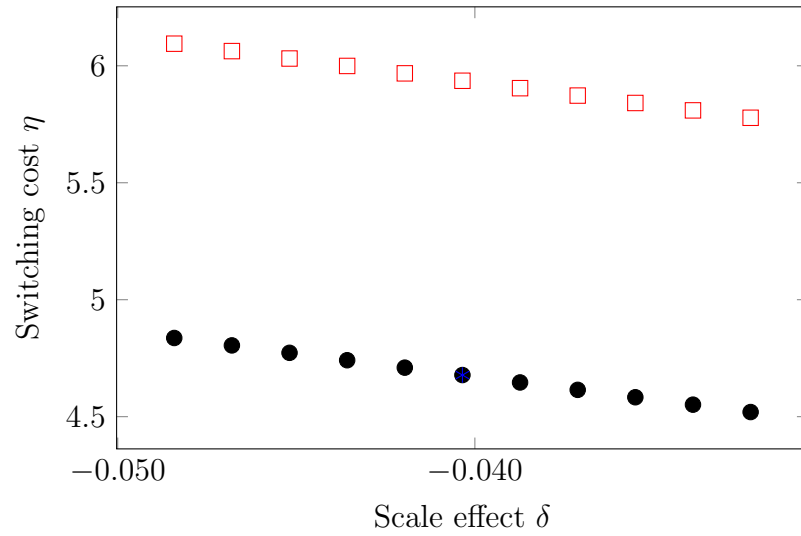


Figure A.4: **Composition of Rulemaking Commenters**

The figure classifies the 664 public comments filed on the Single Audit reform this paper exploits, by the type of commenter. The categories are: *University / research* – universities, colleges, research institutes, laboratories, and medical or research centers; *State / local government* – states, counties, cities, townships, and their departments, agencies, boards, and school districts; *Nonprofit / recipient* – foundations, charities, associations, coalitions, and community organizations; *Individual / self* – submissions filed without a named organization; *Tribal government* – tribes, nations, and pueblos; *Audit / CPA firm* – public accounting firms; *Federal agency* – federal departments and agencies; *For-profit / consultant* – private companies and consultancies; *Govt auditor (state/local)* – state and local audit offices, auditors general, and comptrollers; *CPA society* – the AICPA and state CPA societies; *Federal IG / oversight* – federal inspectors general; and *Other* – submissions not matched by any rule. See [Appendix A.2.2](#).

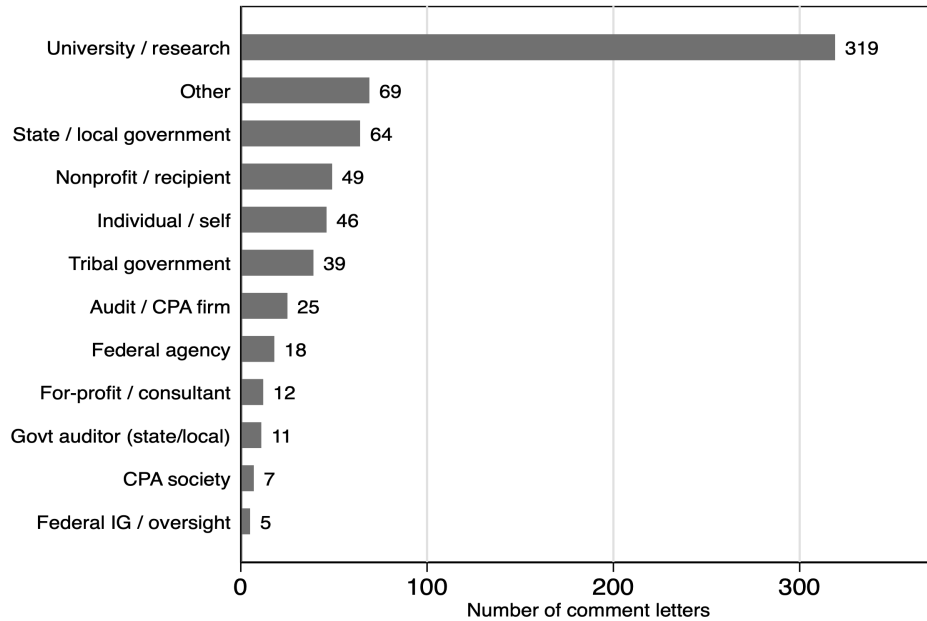


Table A.1: **Robustness: Pre-Reform Observable Differences**

This table augments the main specification from column (3) of [Table 2](#) by sequentially adding interactions between *Post* and pre-reform auditor characteristics: number of clients, average award amount per client (in \$M), total portfolio value (in \$M), and average auditor-client tenure (in years). Standard errors are heteroskedasticity-robust. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. See [Section 3.2](#).

Outcome:	Negative Audit				
	(1)	(2)	(3)	(4)	(5)
Treated $\times$ Post	-1.956*** (0.279)	-1.972*** (0.285)	-1.939*** (0.286)	-1.942*** (0.286)	-1.934*** (0.286)
Post $\times$ # Clients		0.023 (0.094)	-0.307*** (0.099)	-0.167 (0.161)	-0.141 (0.160)
Post $\times$ Avg Amount			1.314*** (0.095)	1.452*** (0.152)	1.463*** (0.151)
Post $\times$ Portfolio				-0.228 (0.181)	-0.233 (0.181)
Post $\times$ Tenure					0.119 (0.098)
Period FE	Yes	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes	Yes
Pre-Reform Mean	8.15	8.15	8.15	8.15	8.15
Effect (%)	-23.99	-24.19	-23.78	-23.82	-23.72
Observations	168,023	168,023	168,023	168,023	168,023
R ²	0.588	0.588	0.589	0.589	0.589

Table A.2: **Contagion Effect: Additional Audit Outcomes**

Panel A. We estimate Equation (1) replacing the dependent variable with each opinion type. Column 1 reproduces the main outcome. Columns 2–4 decompose into financial-statement opinions: any negative FS opinion, qualified FS, and adverse FS. Columns 5–7 report the corresponding major-program opinions. Standard errors are heteroskedasticity-robust. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. See Section 3.4.

Outcome:	Negative Audit	Negative FS	Qualified FS	Adverse FS	Negative MP	Qualified MP	Adverse MP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated $\times$ Post	-1.852*** (0.282)	-1.914*** (0.193)	-0.323* (0.170)	-1.645*** (0.118)	0.052 (0.239)	0.008 (0.020)	0.003* (0.002)
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Reform Mean	8.08	4.66	2.35	2.58	4.16	0.11	0.00
Effect (%)	-22.93	-41.08	-13.72	-63.86	1.26	7.50	77.47
Observations	194,404	194,404	194,404	194,404	194,404	194,404	194,404
R ²	0.616	0.680	0.669	0.683	0.538	0.365	0.388

Panel B. We estimate Equation (1) on internal-control findings and award-level audit outcomes. *IC Deficiency* equals 100 if the entity has any material weakness or significant deficiency. *Questioned Costs* and *Material Noncompliance* are analogous indicators. *Total* is the count of findings across all awards, and *Share* is the percentage of awards with at least one finding. Standard errors are heteroskedasticity-robust. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. See Section 3.4.

	IC Deficiency	Questioned Costs	Material Noncompl.	Award Findings	
				Total	Share
	(1)	(2)	(3)	(4)	(5)
Treated $\times$ Post	2.170*** (0.516)	0.096 (0.276)	0.067 (0.282)	-0.021 (0.291)	-0.333 (0.279)
Period FE	Yes	Yes	Yes	Yes	Yes
Pair FE	Yes	Yes	Yes	Yes	Yes
Pre-Reform Mean	26.21	3.86	4.11	1.56	7.79
Effect (%)	8.28	2.49	1.64	-1.33	-4.28
Observations	193,539	193,539	193,539	193,539	193,539
R ²	0.632	0.456	0.597	0.658	0.594

Table A.3: **Correlation between Audit Outcomes**

Pairwise correlations among six audit outcomes: *Negative Audit*, *Adverse Opinion*, *Qualified Opinion*, *IC Deficiency*, *Questioned Costs*, and *Material Noncompliance*. Computed on the common sample of $N = 193,560$ observations. *** denotes significance at the 1% level. See [Section 3.4](#).

	Negative Audit	Adverse Opinion	Qualified Opinion	IC Deficiency	Questioned Costs	Material Noncompl.
Negative Audit	1.000					
Adverse Opinion	0.486***	1.000				
Qualified Opinion	0.896***	0.115***	1.000			
IC Deficiency	0.314***	0.139***	0.293***	1.000		
Questioned Costs	0.277***	0.066***	0.292***	0.267***	1.000	
Material Noncompl.	0.296***	0.156***	0.269***	0.262***	0.155***	1.000

Table A.4: **Contagion Effect: Excluding Large Auditors**

The table re-estimates the baseline specification (Equation (1)), with pair and period fixed effects, on subsamples that progressively exclude the largest auditors. Column 1 is the full sample. Columns 2–5 restrict to auditors whose total federal awards audited in the last pre-reform period fall below \$1 billion, \$500 million, \$100 million, and \$50 million, respectively. *Pre-Reform Mean* is the negative-audit rate in the last pre-reform period within each column’s sample, and *Effect* expresses the coefficient as a percentage of that mean. Robust standard errors in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels. See Section 4.1.

Outcome:		Negative Audit			
Sample:	Baseline (all auditors)	Fed $\leq$ \$1B	Fed $\leq$ \$500M	Fed $\leq$ \$100M	Fed $\leq$ \$50M
	(1)	(2)	(3)	(4)	(5)
Treated $\times$ Post	-1.852*** (0.282)	-2.382*** (0.295)	-1.623*** (0.302)	-2.173*** (0.369)	-1.448*** (0.412)
Pair FE	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes
Pre-Reform Mean	8.08	8.16	7.63	7.90	7.36
Effect (%)	-22.9	-29.2	-21.3	-27.5	-19.7
Observations	194,404	137,169	121,190	69,760	46,612
R ²	0.616	0.622	0.625	0.625	0.650

Table A.5: **Costs of Issuing a Negative Audit**

The table reports regressions of auditor-client relationship outcomes on an indicator for a negative audit opinion. The dependent variables are indicators (scaled 0–100) for whether the relationship survived to the next period (columns 1 and 4), whether the client exited the Single Audit sample (columns 2 and 5), and whether the client switched auditors conditional on remaining in the sample (columns 3 and 6). *Negative Audit* equals 100 if the auditor issued a qualified or adverse opinion. Within each panel, columns 1–3 report the raw difference in means and columns 4–6 add entity-state and period fixed effects. Panel A restricts the sample to the pre-reform window (2013–2015); Panel B uses the full panel (1998–2021). Standard errors are heteroskedasticity-robust. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels. See [Section 4.4](#).

Specification:	Difference in Means			Entity State + Period FE		
Outcome:	Survived	Exited	Switched	Survived	Exited	Switched
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Pre-reform window (2013–2015)</i>						
Negative Audit	-6.405*** (0.456)	5.849*** (0.398)	1.612*** (0.361)	-5.757*** (0.465)	4.876*** (0.391)	1.715*** (0.383)
Sample Mean	77.95	13.74	10.05	77.95	13.74	10.05
Effect (%)	-8.22	42.58	16.03	-7.38	35.49	17.06
Observations	126,816	126,816	109,395	126,816	126,816	109,395
R ²	0.002	0.002	0.000	0.017	0.027	0.008
<i>Panel B. Full panel (1998–2021)</i>						
Negative Audit	-5.318*** (0.164)	3.861*** (0.129)	3.048*** (0.155)	-3.223*** (0.167)	2.579*** (0.131)	1.645*** (0.156)
Sample Mean	71.91	12.95	17.87	71.91	12.95	17.87
Effect (%)	-7.39	29.81	17.06	-4.48	19.91	9.21
Observations	946,253	946,253	823,693	946,252	946,252	823,692
R ²	0.001	0.001	0.001	0.037	0.020	0.041
Period FE	No	No	No	Yes	Yes	Yes
Entity State FE	No	No	No	Yes	Yes	Yes

Table A.6: Strategic Sorting

Panel A: Triple-Difference. We estimate Equation (A.1) in a sample restricted to organizations with \$750,000–\$1,100,000 in federal expenditures. *Exit* equals 100 if the entity undergoes a Single Audit in period t but not in $t+1$. *JustAbove750* equals one if the entity's expenditures fall within the bin indicated in the column header. Each column corresponds to a separate \$50,000-wide bin. All regressions include auditor and period fixed effects. Standard errors are heteroskedasticity-robust. *, **, and *** denote significance at the 10%, 5%, and 1% levels. See Appendix A.5.

Outcome:	Exit				
	750K-800K	800K-850K	850K-900K	900K-950K	950K-1.0M
<i>JustAbove750</i> = 1:					
<i>JustAbove750</i> = 0:	800K-1.1M	850K-1.1M	900K-1.1M	950K-1.1M	1.0M-1.1M
	(1)	(2)	(3)	(4)	(5)
<i>JustAbove750</i> $\times$ <i>Treated</i> $\times$ <i>post</i> (β_5)	5.453* (2.862)	-1.841 (2.649)	-3.802 (2.573)	0.809 (2.582)	3.095 (2.643)
<i>JustAbove750</i> $\times$ <i>post</i> (β_4)	1.981 (2.746)	5.195** (2.534)	4.965** (2.456)	1.087 (2.454)	-3.090 (2.501)
<i>Treated</i> $\times$ <i>post</i> (β_3)	0.651 (1.075)	0.938 (1.139)	1.628 (1.246)	1.028 (1.379)	0.863 (1.653)
<i>JustAbove750</i> $\times$ <i>Treated</i> (β_2)	-0.827 (1.403)	-3.905*** (1.387)	-0.896 (1.400)	-2.953** (1.399)	-1.152 (1.463)
<i>JustAbove750</i> (β_1)	5.850*** (1.315)	6.387*** (1.301)	2.181* (1.311)	3.453*** (1.297)	1.151 (1.350)
Auditor FE	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes
Pre-Reform Mean	23.27	20.35	18.72	17.60	16.85
Effect (%)	23.44	-9.04	-20.31	4.60	18.37
Observations	90,996	76,388	62,019	48,263	35,280
R ²	0.27	0.28	0.30	0.32	0.34

Panel B: Pooled Difference-in-Differences. We estimate Equation (A.2), which pools all entities regardless of treatment status. The sample, outcome, bin definitions, and fixed effects are the same as in Panel A. See Appendix A.5.

Outcome:	Exit				
<i>JustAbove750</i> = 1:	750K-800K	800K-850K	850K-900K	900K-950K	950K-1.0M
<i>JustAbove750</i> = 0:	800K-1.1M	850K-1.1M	900K-1.1M	950K-1.1M	1.0M-1.1M
	(1)	(2)	(3)	(4)	(5)
<i>JustAbove750</i> $\times$ <i>post</i> (β_4)	8.074*** (0.683)	3.815*** (0.656)	1.957*** (0.647)	2.187*** (0.671)	-0.277 (0.718)
<i>JustAbove750</i> (β_1)	3.620*** (0.336)	2.514*** (0.339)	0.904*** (0.346)	0.584 (0.367)	0.357 (0.400)
Auditor FE	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes
Pre-Reform Mean	23.93	20.78	19.23	17.75	16.91
Effect (%)	33.74	18.36	10.18	12.32	-1.64
Observations	142,120	118,388	95,526	73,719	53,218
R ²	0.27	0.28	0.30	0.31	0.33

Table A.7: **Cost of Auditing: Structural Estimation**

Panel A. This table presents the structural parameter estimates from the dynamic discrete choice model described in [Appendix A.6.1](#). θ is the recurring cost of a negative audit, η is the cost of switching from a positive to a negative audit, and δ is the scale discount. The first row reports full-sample estimates; the remaining blocks compare the pre-reform ($k = -3$ to -1) and post-reform ($k = 1$ to 3) subsamples, first pooling all auditors and then separately for treated and control auditors, classified as in [Section 3.1](#). When comparing subsample estimates, we use the scale-factor adjustment of [Swait and Louviere \(1993\)](#) to account for differences in the dispersion of the private information shock; for each pre/post comparison, the likelihood-ratio statistic and p -value from the first-stage Swait-Louviere test are reported. Standard errors (in parentheses) are clustered by auditor-client pair. See [Appendix A.6.5](#).

Parameter estimates			
	θ	η	δ
Full sample	0.306 (0.003)	5.636 (0.075)	-0.065 (0.004)
<i>All auditors</i>			
Pre-reform	0.328 (0.009)	6.408 (0.217)	-0.108 (0.011)
Post-reform	0.643 (0.013)	4.678 (0.230)	-0.040 (0.012)
SL test	$LR = 311.5, p < 0.001$		
<i>Treated auditors</i>			
Pre-reform	0.311 (0.010)	6.034 (0.271)	-0.082 (0.014)
Post-reform	0.686 (0.015)	4.568 (0.275)	-0.039 (0.014)
SL test	$LR = 350.7, p < 0.001$		
<i>Control auditors</i>			
Pre-reform	0.466 (0.035)	6.849 (0.865)	-0.148 (0.047)
Post-reform	0.560 (0.045)	5.708 (1.007)	-0.087 (0.056)
SL test	$LR = 3.42, p = 0.18$		

Panel B. This table reports the probability of issuing a negative opinion in the model and in the data, using the full-sample estimates. The subsamples are based on the two state variables: last period’s audit outcome (positive or negative) and auditor size, where “Large” and “Small” indicate the above- and below-median split based on total federal expenditures across the auditor’s clients. See [Appendix A.6.5](#).

Probability of negative opinion (full sample)		
	Model	Actual
Full sample	0.083	0.081
Positive opinion	0.025	0.026
... and large auditor	0.028	0.028
... and small auditor	0.022	0.022
Negative opinion	0.679	0.678
... and large auditor	0.677	0.650
... and small auditor	0.680	0.714

Table A.8: **Cost of Auditing: Structural Estimation – Extensions**

Panel A. Pricing the Client-Loss Risk. This table splits the recurring cost of a negative audit into a part explained by measured client-loss risk and a residual. *Total* is θ estimated with the relationship-survival channel switched off, with survival held at its average level so that the opinion has no termination consequence. *Residual* is θ estimated with the channel switched on, where the model charges the auditor the forfeited continuation value whenever it issues a negative opinion; it is therefore the part of the recurring cost that measured client loss does not explain. *Priced risk* is that forfeited value, $\rho[s(0) - s(1)] \mathbb{E}[V']$, and $\mathbb{E}[V]$ is the model-implied value of a surviving relationship, both in annual fees. The upper block calibrates $s(d)$ to the observed one-year pair survival rates by opinion; the lower block calibrates the gap to the entity-state and period fixed-effects estimate within each window, leaving $s(0)$, the total, and $\mathbb{E}[V]$ unchanged by construction. The switching cost and scale discount are nearly invariant across these specifications and thus not reported. Standard errors (in parentheses) are clustered by auditor-client pair and treat $s(d)$ as known. See [Appendix A.6.6](#).

	$s(0)$	$s(1)$	θ total	θ residual	Priced risk	$\mathbb{E}[V]$ (fees)
<i>Raw survival rates</i>						
Pre-reform	0.757	0.683	0.082 (0.012)	-0.249 (0.014)	0.256	3.64
Post-reform	0.828	0.731	0.534 (0.015)	0.029 (0.016)	0.442	4.78
<i>Fixed-effects-adjusted survival rates</i>						
Pre-reform	0.757	0.694	0.082 (0.012)	-0.202 (0.013)	0.219	3.64
Post-reform	0.828	0.738	0.534 (0.015)	0.067 (0.016)	0.409	4.78

Panel B. Observable Client Characteristics This table presents structural parameter estimates from the extended model in which the recurring cost depends on the client’s federal expenditures, $\theta(p_t) = \theta_0 + \theta_P (p_t - \bar{p})$, where p_t is the logarithm of the client’s real federal expenditures and $\bar{p}$ its sample mean. The specification nests the baseline model at $\theta_P = 0$ and the likelihood-ratio test compares the two. Estimates are reported for the full sample, the pre-reform subsample ($k = -3$ to -1), and the post-reform subsample ($k = 1$ to 3). Standard errors (in parentheses) are clustered by auditor-client pair. See [Appendix A.6.6](#).

	Full Sample	Pre-Reform	Post-Reform
θ at mean expenditures	0.306 (0.003)	0.332 (0.009)	0.663 (0.014)
θ_P (expenditure slope)	0.0005 (0.002)	0.019 (0.006)	−0.082 (0.008)
η	5.640 (0.076)	6.565 (0.221)	4.126 (0.236)
δ	−0.066 (0.004)	−0.116 (0.011)	−0.012 (0.012)
LR test vs. baseline ($df = 1$)	0.0	9.3	95.1

Panel C. Permanent Unobserved Heterogeneity This table presents structural parameter estimates from the finite-mixture model with two permanent types of auditor-client pairs. The types differ only in the recurring cost: θ_1 for the low-cost type and $\theta_2 = \theta_1 + \mu$ for the high-cost type; all other model features are kept the same. The type probability is conditioned on the pair’s first observed outcome to address initial conditions (Wooldridge, 2005), and $\mathbb{E}[\pi_2]$ is the population share of the high-cost type. Estimates are reported for the full sample, the pre-reform subsample ($k = -3$ to -1), and the post-reform subsample ($k = 1$ to 3). Standard errors (in parentheses) are clustered by auditor-client pair. See Appendix A.6.6.

	Full Sample	Pre-Reform	Post-Reform
θ_1 (low type)	0.074 (0.005)	−0.324 (0.043)	−0.042 (0.041)
$\theta_2 = \theta_1 + \mu$ (high type)	1.246	2.137	2.473
μ	1.172 (0.019)	2.461 (0.117)	2.516 (0.108)
η	4.887 (0.088)	5.154 (0.289)	3.290 (0.298)
δ	−0.081 (0.005)	−0.153 (0.014)	−0.062 (0.015)
$\mathbb{E}[\pi_2]$ (high-type share)	0.82	0.91	0.93

Table A.9: Select Enforcement Actions Relating to Federal Awards

The table lists select federal enforcement actions involving federal awards and grants between 2010 and 2019. The enforcement actions are obtained from two sources: the U.S. Department of Justice news archive ([Link](#)), and the enforcement database of the Office of Inspector General of the U.S. Department of Health and Human Services ([Link](#)). We retain releases in which the conduct concerns a federal award or grant and exclude routine award announcements and unrelated matters. Multiple releases describing the same scheme (e.g., indictment, plea, conviction, sentencing, or settlement) are consolidated into a single matter. See [Section 4.5](#).

Year	Entity/Person(s)	Type	Allegation(s)	DOJ Action	Source
2010	Educational Broadcasting Corp.	Nonprofit	Charged improper expenditures to federal awards.	Settlement	DOJ Link
2011	Institute of International Education	Nonprofit	Failed to comply with grant requirements and repeatedly made false claims for payment.	Settlement	DOJ Link
2013	Science Applications International Corp.	For Profit	Charged inflated prices under grants to train first responder personnel to prevent and respond to terrorism attacks.	Settlement	DOJ Link
2014	AA Solar Inc.; Joseph Kozicki	For Profit	Submitted false documentation to DCEO and used some of the grant money for other purposes.	Criminal indictment	DOJ Link
2015	PricewaterhouseCoopers LLP	Auditor	Had extensive deficiencies in a prior PwC audit of Big Brothers Big Sisters of America's (BBBSA) compliance with the laws, regulations, and grant requirements for its federal programs. The original report was withdrawn and replaced.	DOJ OIG Quality Control Review	DOJ Link
2017	Milton S. Eisenhower Foundation; Alan Curtis	Nonprofit	Submitted false claims for payments of grant funds awarded by the U.S. Department of Justice.	Settlement	DOJ Link
2018	Beth N. Sallee	Individual	Devised a scheme and artifice to defraud and to obtain money and property by means of material false and fraudulent pretenses, representations, and promises.	Criminal charges	DOJ Link
2019	Duke University	University	Violated the False Claims Act by submitting applications and progress reports that contained falsified research on federal grants to the National Institutes of Health and to the Environmental Protection Agency.	Settlement	DOJ Link

Table A.10: **Random Sample of 10 Treated Auditors**

The table describes a random sample of 10 treated auditors and summarizes the accounting services offered by each auditor. See [Section 4.5](#).

Auditor	City	Source	Specialty
Paul H. Koehler	Lincoln, NE	Link	Government and nonprofit services
Gregory, Sharer & Stuart	St. Petersburg, FL	Link	Not-for-profit, manufacturing and health care, real estate and construction, and insurance
Sciarani & Co.	Yerington, NV	Link	Accounting services, taxation, and governmental audits
James G. Zupka	Garfield Heights, OH	Link	Government and not-for-profit auditing, accounting, and tax preparation
Mense, Churchwell & Mense	Joplin, MO	Link	Audit, tax, bookkeeping
Fortenberry & Ballard	Brandon, MS	Link	Audit, tax, payroll
Donald E. Curtis	Beebe, AR	Link	Audit and other services for cities, utilities, and education
Vine Dahlen	Lynnwood, WA	Link	Tax planning and compliance, auditing and assurance, accounting, and consulting
Clarke Whitney	Bremerton, WA	Link	Business (e.g., nonprofit organizations), tax, and QuickBooks
Roman Toro & Co.	Yauco, PR	Link	Single-audit tax services, accounting and auditing, non-traditional engagements, and write-up services