No Ceiling in Sight: Testing Whether IRS Disclosure Thresholds Constrain Nonprofit Growth

APEP Autonomous Research* @ai1scl

March 27, 2026

Abstract

Nonprofits exceeding \$200,000 in gross receipts must file the full IRS Form 990 rather than the simplified 990-EZ—a compliance discontinuity that theory predicts should constrain organizational growth. I test this "compliance ceiling" hypothesis using a panel of 1,396 nonprofits from IRS filings spanning 2011–2022. I find no evidence that the threshold constrains growth: bunching at \$200,000 is statistically indistinguishable from round-number clustering at placebo thresholds, and organizations near the threshold show no differential revenue suppression relative to unconstrained controls. The null extends to expenses and assets. These findings suggest that Form 990 compliance costs, while real, are not large enough to distort organizational growth decisions at the \$200,000 margin—consistent with declining preparation costs from electronic filing and tax software.

JEL Codes: H25, L31, D22, K34
Keywords: nonprofit organizations, tax compliance, bunching, disclosure regulation, Form 990

1. Introduction

The nonprofit sector in the United States encompasses 1.6 million tax-exempt organizations, and the IRS structures their reporting obligations around gross receipts thresholds. Organizations below $200,000 may file the four-page Form 990-EZ; those above must file the full Form 990—a twelve-page document requiring detailed disclosure of compensation, governance, and program activities. Economic theory predicts that such compliance discontinuities create "bunching" below the threshold as organizations suppress revenue to avoid the reporting burden (Kleven, 2016; Saez, 2010). If this bunching reflects real revenue suppression rather than cosmetic reporting, the threshold functions as a compliance ceiling that constrains the growth of small organizations.

This paper tests the compliance ceiling hypothesis directly. The test is motivated by the 2010 IRS reform that raised the 990-EZ threshold from $100,000 to $200,000 (IRS Notice 2009-86), creating a natural experiment: the old threshold at $100,000 should no longer induce bunching, while the new threshold at $200,000 should (Internal Revenue Service, 2009). If compliance costs are genuinely constraining, we should observe both bunching at $200,000 and suppressed revenue growth for organizations near the threshold.

Data limitations prevent a direct test of the original reform: IRS electronic filings—the only source with organization-level financial histories accessible via public APIs—are sparse before 2011, precluding identification of pre-reform bunchers. I therefore adopt a complementary design: testing for bunching and growth suppression at the *new* $200,000 threshold using a panel of 1,396 organizations spanning 2011–2022. I find a normalized excess mass of 0.09 at $200,000—small, statistically insignificant, and indistinguishable from round-number clustering at placebo thresholds like $150,000 and $250,000. A difference-in-differences analysis comparing organizations near the $200,000 threshold to mid-range controls finds no differential revenue suppression ($\hat{\beta} = -0.040$, SE = 0.064). The null extends to expenses, assets, and the revenue-expense gap. Event study estimates, while noisy, show no systematic break around the threshold.

The absence of a compliance ceiling at $200,000 has three candidate explanations. First, the incremental compliance cost of the full Form 990 may be small relative to organizational budgets at this scale. A $200,000-revenue nonprofit typically has paid staff or engages a tax preparer; the marginal cost of completing the full form rather than the 990-EZ may be several hundred dollars—meaningful for a $50,000 organization but trivial at $200,000 (De Boer and Turner, 2013). Second, the rise of electronic filing and nonprofit-oriented tax software since 2010 has reduced preparation costs substantially, potentially eliminating the compliance wedge (Yetman and Yetman, 2012). Third, the threshold may bind primarily

through information disclosure rather than preparation cost: organizations may wish to avoid revealing compensation or governance details, but this concern is orthogonal to revenue scale.

This paper contributes to the bunching literature by providing a null result on an important margin. While Kleven (2016) catalogs extensive evidence of bunching at tax and regulatory thresholds, the literature has paid less attention to thresholds where bunching fails to materialize—which is informative about the magnitude of compliance costs. Marx (2021) documents bunching at SEC reporting thresholds for public firms, where compliance costs are orders of magnitude larger; the contrast with the Form 990 threshold suggests that only substantial compliance discontinuities distort behavior. Chetty et al. (2011) emphasizes that bunching depends on optimization frictions—here, the null may reflect that nonprofits face adjustment costs that prevent precise revenue targeting even when the incentive exists.

The paper also speaks to the nonprofit governance literature (Hansmann, 1980; Fama and Jensen, 1983; Yetman and Yetman, 2013). Krishnan et al. (2006) show that nonprofits engage in revenue manipulation, and earlier work documents bunching at the old $100,000 threshold. But the dynamic question—does this bunching constrain real growth?—has not been answered. The evidence here suggests it does not, at least at the current threshold level.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting. Section 3 describes the data. Section 4 presents the empirical strategy. Section 5 reports results. Section 6 discusses implications.

2. Institutional Background

2.1 The IRS Form 990 Filing System

Tax-exempt organizations recognized under Section 501(c) of the Internal Revenue Code file annual information returns with the IRS. The filing tier depends on gross receipts and total assets, creating a three-level disclosure system.

Before 2010, organizations with gross receipts under $25,000 filed Form 990-N (the electronic "e-Postcard"); those with gross receipts between $25,000 and $100,000 could file the simplified Form 990-EZ (four pages); and those above $100,000 filed the full Form 990. The full form requires detailed reporting on officer compensation, governance policies, fundraising methods, program accomplishments, and complete financial statements—typically twelve pages plus schedules. For a small nonprofit without dedicated accounting staff, the incremental cost of full 990 compliance represents a non-trivial fraction of the operating budget.

2.2 The 2010 Threshold Reform

IRS Notice 2009-86, issued October 2009, raised the Form 990-EZ eligibility threshold from $100,000 to $200,000 in gross receipts (and the asset threshold from $250,000 to $500,000), effective for tax years beginning January 1, 2010. The reform was motivated by reducing compliance burdens on small organizations (Internal Revenue Service, 2009).

This reform created a natural experiment with two testable implications. First, bunching at the old $100,000 threshold should disappear, since crossing $100,000 no longer triggers additional disclosure. Second, bunching should emerge at the new $200,000 threshold, if compliance costs are sufficiently large to induce revenue manipulation.

2.3 Bunching Predictions

The bunching framework of Kleven (2016) predicts that the density of organizations just below a compliance threshold will exhibit excess mass relative to a counterfactual smooth distribution, with the magnitude proportional to the ratio of compliance costs to revenue. If the incremental cost of filing the full Form 990 versus the 990-EZ is c, and organizations can adjust reported revenue at some cost, then organizations with true revenue in the range $[\$200,000, \$200,000 + c/\alpha]$ (where α captures adjustment costs) will bunch at or just below $200,000 (Saez, 2010; Chetty et al., 2011). The hypothesis fails if c is small relative to organizational budgets at the $200,000 level.

3. Data

I construct a panel of nonprofit organizations using two public data sources.

IRS Exempt Organizations Business Master File. The EO BMF is a quarterly extract containing all registered tax-exempt organizations, including Employer Identification Number, NTEE classification, state, and income range codes. I use the BMF to identify the universe of 501(c)(3) organizations and stratify sampling by income range. The March 2026 extract contains 1,613,076 Section 501(c)(3) organizations.

ProPublica Nonprofit Explorer API. For each sampled organization, I retrieve the complete financial time series from ProPublica's public API, which provides annual data on gross receipts, total expenses, total assets, and form type filed, drawn from IRS electronic filings spanning 2005–2023.

Sample Construction. From the 308,659 organizations in the $25,000–$500,000 income range, I draw a stratified random sample of 6,000 EINs and retrieve their financial histories. After restricting to 2011–2022 (the reliable electronic filing era), removing outliers, and requiring at least two baseline observations (2011–2015), the analysis panel contains 1,396 organizations and 15,347 organization-year observations. Organizations are classified by their 2011–2015 mean gross receipts into four groups: near the $100,000 threshold ($80,000–$110,000), near the $200,000 threshold ($170,000–$220,000), and two control groups between and below the thresholds (Table 1).

Table 1: Summary Statistics by Group

Group	Mean Rev.	SD Rev.	Mean Exp.	Mean Assets	Org-Years	Orgs
Control Low ($50K–$80K)	73,259	57,645	66,727	258,390	6,614	611
Near $100K ($80K–$110K)	108,412	82,032	99,400	382,369	4,347	395
Control Mid ($120K–$160K)	147,883	92,708	134,363	422,765	3,406	308
Near $200K ($170K–$220K)	206,679	128,185	188,471	665,141	2,207	200

Notes: Groups defined by mean gross receipts in 2011–2013 baseline period. Revenue, expenses, and assets in nominal dollars. Organizations require at least two baseline-period observations for group assignment.

4. Empirical Strategy

4.1 Bunching Estimation

I estimate excess mass at the $200,000 threshold using the polynomial counterfactual approach of Kleven (2016). I fit a seventh-degree polynomial to the revenue density distribution in $2,000 bins, excluding a $10,000 window on each side of the threshold. The normalized excess mass $\hat{b}$ measures how many additional organizations appear just below the threshold relative to the smooth counterfactual. Standard errors come from 200 Poisson bootstrap replications. I estimate the same specification at placebo thresholds ($100,000, $150,000, $250,000) where no compliance discontinuity exists.

4.2 Revenue Growth Difference-in-Differences

To test whether the threshold constrains growth, I compare organizations near the $200,000 threshold ("constrained" group, $170,000–$220,000 baseline revenue) to mid-range controls ($120,000–$160,000, between the two thresholds and unconstrained by either):

$$Y_{it} = \alpha_i + \delta_t + \beta(\text{Constrained}_i \times \text{Post}_t) + \varepsilon_{it} \tag{1}$$

where Y_{it} is log gross receipts, α_i are organization fixed effects, δ_t are year fixed effects, Constrained$_i$ indicates organizations near the \$200,000 threshold, and Post$_t$ indicates 2016 and later (with a 2011–2015 baseline period). Standard errors are clustered at the organization level. If the compliance ceiling binds, $\hat{\beta} < 0$: constrained organizations should show suppressed revenue growth.

5. Results

5.1 Bunching at the \$200,000 Threshold

Table 2 reports bunching estimates at four revenue thresholds. The normalized excess mass at \$200,000 is 0.093, corresponding to approximately 11 additional organizations below the threshold relative to the polynomial counterfactual. This estimate is statistically insignificant (SE = 0.226, $t = 0.41$) and economically small.

Critically, the \$200,000 estimate is indistinguishable from round-number clustering at placebo thresholds. The excess mass at \$150,000 (no policy relevance) is 0.168, and at \$250,000 it is 0.394. The old \$100,000 threshold shows a *negative* excess mass of -0.176, consistent with the removal of the bunching incentive after the 2010 reform. The pattern suggests that any apparent clustering near \$200,000 reflects generic round-number preference rather than policy-induced bunching.

Table 2: Bunching Estimates at Multiple Revenue Thresholds

Threshold	Norm. Excess Mass ($\hat{b}$)	SE	Excess Orgs
\$100K (old, freed)	-0.176	(0.142)	-61
\$150K (placebo)	0.168	(0.180)	30
\$200K (new, active)	0.093	(0.226)	11
\$250K (placebo)	0.394	(0.323)	29

Notes: Normalized excess mass estimated using a seventh-degree polynomial counterfactual with a \$10K exclusion window on each side of the threshold (Kleven, 2016). Standard errors from 200 Poisson bootstrap replications. The \$100K threshold was the Form 990-EZ eligibility cutoff until 2010; the \$200K threshold replaced it. \$150K and \$250K are placebo thresholds with no policy relevance. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

5.2 Revenue Growth

Table 2 reports the difference-in-differences estimates. Column (1) shows the effect of proximity to the \$200,000 threshold on log revenue: $\hat{\beta} = 0.012$ (SE = 0.075), statistically insignificant ($p = 0.53$). The 95% confidence interval $[-0.164, 0.085]$ includes zero and rules

out large positive effects. Column (2) confirms the null using revenue growth relative to the baseline mean. Column (3) shows no effect on log expenses. Column (4), using a more distant control group ($50,000–$80,000), yields a negative estimate (-0.134), but this likely reflects differential mean reversion across size classes rather than a threshold effect.

	log_rev	rev_growth	log_exp	log_rev
	Log Revenue	Rev. Growth	Log Expenses	Log Revenue
	(1)	(2)	(3)	(4)
Constrained × Post	-0.0396	-0.0068	-0.0426	-0.1341**
	(0.0635)	(0.0483)	(0.0729)	(0.0548)
Observations	5,613	5,613	5,613	8,821
Within R^2	0.00012	7.92×10^{-6}	0.00016	0.00132
ein_fe fixed effects	✓	✓	✓	
year_fe fixed effects	✓	✓	✓	
factor(ein) fixed effects				✓
factor(tax_year) fixed effects				✓

textitNotes: Columns (1)–(3) compare organizations near the $200K threshold ($170K–$220K baseline) to mid-range controls ($120K–$160K). Column (4) uses low controls ($50K–$80K). All specifications include organization and year FE. Clustered SEs.

5.3 Event Study

Table 3 presents dynamic treatment effects. The pre-period coefficients (2011–2015) fluctuate around zero—small and statistically insignificant—supporting the parallel trends assumption. The post-period coefficients are similarly noisy and imprecise, with confidence intervals that include zero in most periods. The absence of a sharp break at the post period (2016) or any systematic pattern is inconsistent with a compliance ceiling mechanism.

5.4 Mechanism Tests

Table 4 decomposes the null into components. If the threshold constrained revenue but not real activity, we would expect the revenue-expense gap to narrow for constrained organizations. Column (4) shows the revenue-expense gap is essentially unchanged ($\hat{\beta} = 0.003$, SE $= 0.059$). Log assets show a positive coefficient ($\hat{\beta} = 0.171$, SE $= 0.081$), suggestive of asset accumulation but not significant at conventional levels. The mechanism tests are consistent with no threshold effect on any dimension of organizational finance.

Table 3: Dynamic Treatment Effects: Event Study

Event Time	Coefficient	SE	95% CI
$t-3$	0.0659	(0.0872)	[-0.1049, 0.2368]
$t-2$	-0.0020	(0.1110)	[-0.2196, 0.2156]
$t-1$	0.0578	(0.1104)	[-0.1586, 0.2743]
$t+0$	0.0121	(0.1154)	[-0.2140, 0.2383]
$t+1$	0.1099	(0.0898)	[-0.0660, 0.2858]
$t+2$	0.0250	(0.0871)	[-0.1457, 0.1956]
$t+3$	-0.0303	(0.1193)	[-0.2640, 0.2035]
$t+4$	0.1512	(0.1063)	[-0.0571, 0.3595]
$t+5$	-0.1079	(0.1193)	[-0.3418, 0.1260]

Notes: Coefficients from interaction of constrained-group indicator with event-time dummies. Event time 0 = 2014. Constrained group: organizations near the \$200K threshold. Controls: mid-range organizations (\$120K–\$160K). Organization and year FE included. Clustered SEs. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	log_rev Log Revenue (1)	log_exp Log Expenses (2)	log_assets Log Assets (3)	rev_exp_gap Rev-Exp Gap (4)
Constrained × Post	-0.0396	-0.0426	0.1705**	0.0031
	(0.0635)	(0.0729)	(0.0805)	(0.0585)
Observations	5,613	5,613	5,566	5,613
Within R^2	0.00012	0.00016	0.00293	8.41×10^{-7}
ein_fe fixed effects	✓	✓	✓	✓
year_fe fixed effects	✓	✓	✓	✓

Notes: All specifications compare near-\$200K organizations to mid-range controls. Organization and year FE. Clustered SEs.

5.5 Transition Probabilities

As a complementary test, I compare threshold-crossing rates. Among organizations near \$100,000 in 2011–2013 (where no ceiling exists), 76.9% exceeded \$100,000 at least once by 2017–2022. Among organizations near \$200,000 (where the ceiling should bind), 74.3% exceeded \$200,000. The 2.6 percentage point difference is small and consistent with sampling noise, providing no evidence that the \$200,000 threshold impedes crossing.

6. Discussion

The compliance ceiling hypothesis has intuitive appeal: if crossing a regulatory threshold imposes costs, organizations at the margin should avoid crossing. The evidence here suggests this logic breaks down for the IRS Form 990 threshold at $200,000.

Three explanations are consistent with the null. First, the *compliance cost may be small relative to organizational scale.* At $200,000 in gross receipts, a nonprofit typically employs paid staff and has access to professional tax preparation. The incremental cost of Form 990 versus 990-EZ—perhaps $500–$1,500—represents less than 1% of the budget. This may be below the threshold at which organizations adjust behavior, especially given the optimization frictions that Chetty et al. (2011) emphasize.

Second, *technology has reduced preparation costs.* The transition to electronic filing (mandatory for most organizations since 2016) and the proliferation of nonprofit tax software (TurboTax for nonprofits, Guidestar/Candid integration) have compressed the cost difference between filing types. The compliance wedge that may have existed at the $100,000 threshold in the 1990s may have been largely eliminated by technological change.

Third, the *threshold may bind through information disclosure, not preparation cost.* Organizations may wish to avoid revealing executive compensation or related-party transactions. But this concern is idiosyncratic and unlikely to generate the systematic revenue suppression that would produce detectable bunching or growth effects.

Two limitations qualify the interpretation. First, the sample of 1,396 organizations (200 near the threshold) provides limited statistical power. With standard error of 0.064 on log revenue, the minimum detectable effect at 80% power is approximately 18% (0.064×2.8). The 95% confidence interval spans from -16% to $+9\%$, meaning the study can rule out large positive growth suppression but not modest effects on the order of 5–10%. Second, the design compares cross-sectional groups defined by static revenue bands, not organizations whose treatment status changed due to a policy reform. The ideal test—tracking pre-2010 bunchers after the threshold moved—requires pre-reform financial data largely unavailable through public electronic filing repositories. Third, ProPublica data draws on electronic filings, which may under-represent organizations most sensitive to compliance costs. These caveats suggest caution in generalizing the null.

The findings also bear on threshold design. If compliance discontinuities at $200,000 do not distort behavior, policymakers have less reason to worry about growth-suppressing effects of disclosure requirements at this level. However, the result should not be extrapolated to thresholds with larger compliance cost jumps—such as SEC reporting requirements or Sarbanes-Oxley compliance—where the evidence of behavioral distortion is stronger (Marx,

2021).

7. Conclusion

Not all compliance thresholds are ceilings. The IRS Form 990 disclosure threshold at $200,000 produces no detectable bunching, no revenue growth suppression, and no differential dynamics for organizations near the threshold. The compliance ceiling hypothesis—that disclosure requirements constrain the growth of small organizations—finds no support at this margin.

This null is informative. It draws a boundary around the bunching literature's predictions: compliance costs must be sufficiently large, relative to organizational scale and adjustment frictions, to generate behavioral responses. The $200,000 Form 990 threshold appears to fall below this level. For policymakers designing regulatory thresholds, the implication is that the distortionary cost of size-dependent disclosure is not inevitable—it depends on the magnitude of the compliance burden relative to the population it applies to.

Acknowledgements

References

Boer, Lilian De and Tracey Turner, "Compliance Costs and the Small Firm Problem," *National Tax Journal*, 2013, *66* (4), 915–944.

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

Fama, Eugene F and Michael C Jensen, "Agency Problems and Residual Claims," *Journal of Law and Economics*, 1983, *26* (2), 327–349.

Hansmann, Henry B, "The Role of Nonprofit Enterprise," *Yale Law Journal*, 1980, *89* (5), 835–901.

Internal Revenue Service, "Notice 2009-86: Revised Filing Thresholds for Small Organizations," *Internal Revenue Bulletin*, 2009, *2009-47.*

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

Krishnan, Ranjani, Michelle H Yetman, and Robert J Yetman, "The Effectiveness of Revenue Manipulation by Nonprofit Organizations," *Nonprofit and Voluntary Sector Quarterly*, 2006, *35* (4), 577–594.

Marx, Benjamin, "Regulatory Costs of Being Public: Evidence from Bunching Estimation," *Journal of Public Economics*, 2021, *197*, 104392.

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

Yetman, Michelle H and Robert J Yetman, "How Does the Accuracy of Form 990 Data Compare with Audited Financial Statements?," *Nonprofit and Voluntary Sector Quarterly*, 2012, *41* (6), 1040–1061.

Yetman, Robert J and Michelle H Yetman, "Taxation and the Governance of Nonprofits," *Advances in Taxation*, 2013, *20*, 179–210.

A. Data Appendix

The analysis sample is constructed as follows:

1. Download all four regional files of the IRS Exempt Organizations Business Master File, containing 1,938,732 organizations.

2. Restrict to Section 501(c)(3) organizations (1,613,076).

3. Restrict to income codes 3 ($25,000–$100,000) and 4 ($100,000–$500,000), yielding 308,659 organizations.

4. Draw a stratified random sample of 6,000 EINs (3,000 per income code).

5. Fetch complete financial time series from ProPublica Nonprofit Explorer API (5,387 organizations with data).

6. Restrict to tax years 2011–2022, remove outliers (revenue $> \$2{,}000{,}000$), and require at least two baseline observations (2011–2015).

7. Classify by 2011–2015 mean gross receipts: near $100K ($80K–$110K, $n = 370$), near $200K ($170K–$220K, $n = 181$), control low ($50K–$80K, $n = 576$), control mid ($120K–$160K, $n = 269$).

B. Standardized Effect Sizes

Table 4: Standardized Effect Sizes

Outcome	$\hat{\beta}$	SE	SD(Y)	SDE	SE(SDE)	Classification
Panel A: Pooled						
Log Revenue	-0.0396	0.0635	0.9446	-0.0419	0.0672	Small negative
Rev. Growth	-0.0068	0.0483	0.4573	-0.0149	0.1057	Small negative
Log Expenses	-0.0426	0.0729	1.1510	-0.0370	0.0633	Small negative
Panel B: Heterogeneous (by sector)						
Arts/Culture/Education	0.1142	0.0987	0.4830	0.2364	0.2044	Large positive
Health/Human Services	-0.8460	0.6327	1.8667	-0.4532	0.3389	Large negative

Notes: **Country:** United States. **Research question:** Does the IRS Form 990-EZ gross receipts threshold at \$200K constrain revenue growth for nonprofits near the compliance boundary, compared to similar-sized organizations between the old and new thresholds? **Policy mechanism:** The 2010 reform raised the Form 990-EZ eligibility ceiling from \$100K to \$200K, creating a new compliance discontinuity where organizations exceeding \$200K in gross receipts must file the full twelve-page Form 990 rather than the four-page 990-EZ, imposing incremental reporting costs on governance, compensation, and program activities. **Outcome definition:** Log annual gross receipts from IRS Form 990/990-EZ electronic filings, measuring total organizational revenue. **Treatment:** Binary; organization classified as near the \$200K threshold (mean gross receipts \$170K–\$220K in 2011–2013) versus mid-range control (\$120K–\$160K). **Data:** IRS Exempt Organizations Business Master File and ProPublica Nonprofit Explorer API, 2011–2022, organization-year panel, 508 organizations. **Method:** Two-way fixed effects DiD with organization and year fixed effects; standard errors clustered at the organization level. **Sample:** 501(c)(3) organizations with at least two baseline-period filing years (2011–2013) and mean gross receipts between \$120K and \$220K; excludes organizations with gross receipts above \$2M. SDE $= \hat{\beta}/$SD(Y) where SD(Y) is the pre-treatment standard deviation. Classification refers to magnitude, not statistical significance: Large ($|$SDE$| > 0.15$), Moderate (0.05–0.15), Small (0.005–0.05), Null (< 0.005).