

The Elasticity of Medicaid’s Safety Net: Market Responses to Provider Fraud Exclusions

APEP Autonomous Research* @ailscl

February 18, 2026

Abstract

When the federal government excludes a fraudulent Medicaid provider, does the local market absorb displaced patients or does a care gap persist? We link 82,714 federal exclusions from the OIG List of Excluded Individuals and Entities (LEIE) to 227 million Medicaid claims in T-MSIS (2018–2024), yielding 22 treated market events with sufficient pre-exclusion billing to support a difference-in-differences design. Using static DiD with state \times month fixed effects, we find no statistically significant change in rest-of-market spending, provider counts, or beneficiaries served following exclusions ($\beta = -0.026$, $SE = 0.246$). Randomization inference confirms the null ($p = 0.926$). Our primary contribution is documenting extreme attrition from the 3,770-exclusion universe to 22 analyzable events: most excluded providers are too small to generate measurable local market disruption.

JEL Codes: I13, I18, L11

Keywords: Medicaid, provider exclusion, market elasticity, fraud enforcement, HCBS

*Autonomous Policy Evaluation Project. This paper was generated autonomously. Total execution time: 3h 1m (cumulative: 3h 1m). Correspondence: scl@econ.uzh.ch

1. Introduction

When the federal government bans a healthcare provider for fraud, the patients left behind do not stop needing care. If the local market is thin — as it often is for home-care services in rural areas — removing even a bad actor may cut off the safety net for the very people it protects. Administrative exclusion is the most powerful enforcement tool in Medicaid: once an individual or entity appears on the OIG’s List of Excluded Individuals and Entities (LEIE), they are permanently barred from billing any federal health program ([Office of Inspector General, 2024](#)). Over 82,000 such exclusions have been imposed since the 1970s. Yet remarkably little is known about what happens *after* an exclusion takes effect.

This gap matters because the policy stakes run in both directions. If local markets cannot absorb the loss of an excluded provider, aggressive fraud enforcement may inadvertently harm beneficiaries — the very population Medicaid exists to serve. If markets are elastic, enforcement can proceed without such concerns. Resolving this question requires knowing which excluded providers were actually billing Medicaid, how large they were in their local markets, and whether remaining providers expanded to fill the gap. Until now, no dataset has linked exclusion records to actual Medicaid claims at scale.

This paper takes the first systematic look at local market responses to Medicaid provider fraud exclusions. We link the complete universe of OIG LEIE exclusion actions to administrative Medicaid billing data from the Transformed Medicaid Statistical Information System (T-MSIS), covering 227 million claim records from 2018 through 2024. This novel linkage allows us to identify, for the first time, which excluded providers were actually billing Medicaid prior to their exclusion, and to construct rest-of-market outcomes measuring whether other providers expanded to fill the gap.

Our empirical strategy is a difference-in-differences event study. For each excluded provider with meaningful pre-exclusion market presence, we define the local market as the ZIP code \times Medicaid service category, and compare rest-of-market outcomes (spending, provider counts, beneficiaries) before and after the exclusion relative to untreated markets in the same state and month. The identifying assumption is that treated and control markets would have followed parallel trends absent the exclusion — an assumption we examine through pre-trend tests and placebo analysis.

We report two findings, and the second may be more important than the first. The primary econometric result is null: provider exclusions generate no statistically significant change in rest-of-market spending, provider counts, or beneficiaries ($\hat{\beta} = -0.026$, $SE = 0.246$ for log spending). Randomization inference confirms this ($p = 0.926$). But the more striking finding is why the sample is so small. Of 82,714 cumulative LEIE records, only 3,770 have

valid NPIs in our study window. Of those, only 85 (2.3%) appear in T-MSIS billing records. After requiring meaningful pre-exclusion market presence ($\geq 3\%$ share), just 22 market events remain. Most excluded providers, it turns out, were either not billing Medicaid at the time of exclusion, had already stopped billing before the exclusion took effect, or held market shares too small to generate detectable disruption.

This attrition cascade reframes the policy question. Rather than asking “does removing a fraudulent provider disrupt care?” the answer may be that enforcement rarely targets providers large enough for the question to matter. The typical excluded provider in our data billed roughly \$207,000 per year — substantial in isolation, but a modest share of most local markets. The null DiD result is consistent with either genuine market elasticity (remaining providers absorb displaced patients) or insufficient statistical power (22 treated units cannot detect effects of plausible magnitude). The minimum detectable effect at 80% power is approximately 0.69 log points — larger than the excluded providers’ average market share — underscoring the power constraint.

Our contribution to the literature is threefold. First, we produce the first systematic linkage of the OIG LEIE exclusion database to T-MSIS administrative Medicaid claims, enabling a class of questions about fraud enforcement effects that was previously not addressable with large-scale data. Second, we document the structure of the attrition cascade from exclusion event to analyzable market treatment, providing a roadmap for future research on the data requirements for this line of inquiry. Third, we provide the first causal evidence — however tentative given statistical power — on local market responses to Medicaid provider exclusions, contributing to the broader literature on healthcare market elasticity (Bresnahan and Reiss, 1991; Dafny, 2005) and Medicaid program administration (Finkelstein et al., 2012; Garthwaite, 2012).

Related literature. This paper connects several strands of research. A large literature examines Medicaid market structure and provider response to policy changes (Currie and Gruber, 2008; Finkelstein et al., 2012; Garthwaite, 2012; Duggan, 2004, 2000). Within this, work on provider supply responses to payment rate changes establishes that Medicaid markets can exhibit significant supply-side rigidities (Clemens and Gottlieb, 2014; Gruber et al., 1999). The most closely related empirical work examines market responses to nursing home closures (Mommaerts and Truskinovsky, 2023), which shares our focus on supply-side shocks to local healthcare markets. Our paper differs in focusing on fraud-based exclusions rather than financial distress-driven closures, and in using a novel administrative data linkage.

A distinct strand of literature addresses healthcare fraud enforcement. The economics of fraud in government programs has been examined theoretically and empirically (Decarolis, 2014), but systematic causal evidence on the market consequences of fraud enforcement

remains sparse. We are not aware of prior work that uses LEIE data linked to Medicaid billing to study market-level effects. The literature on Medicaid managed care and payment reform provides relevant evidence on provider responses to administrative interventions (Buchmueller et al., 2005; Grabowski, 2004; Brown et al., 2014; Medicaid and CHIP Payment and Access Commission, 2023). Work on physician responses to incentive changes (Einav et al., 2018; Chandra and Skinner, 2012) informs our understanding of the supply-side margin along which providers respond to market disruption.

The health economics literature on market structure and competition is also relevant. Research on entry and competition in healthcare markets (Bresnahan and Reiss, 1991; Berry, 1992; Dafny, 2005; Gaynor et al., 2016; Ho, 2009) establishes that market structure shapes how supply shocks propagate through local markets. Our focus on service-level market concentration — measured by the excluded provider’s pre-exclusion billing share — connects to this work by testing whether highly concentrated markets are less elastic following provider exits.

Finally, the econometrics literature on difference-in-differences with staggered treatment timing is directly relevant to our design. We draw on recent work formalizing event study estimators (Callaway and Sant’Anna, 2021; Sun and Abraham, 2021; Goodman-Bacon, 2021; Roth et al., 2023; de Chaisemartin and D’Haultfoeuille, 2020; Borusyak et al., 2024) and on sensitivity analysis for violations of the parallel trends assumption (Rambachan and Roth, 2023). Baker et al. (2022) demonstrate the magnitude of bias from TWFE under heterogeneous effects; we discuss why our design — which uses only never-treated controls — mitigates these concerns in Section 5. Given our small treatment sample, we use randomization inference as a co-primary inference method alongside conventional clustered standard errors, following Deryugina (2017) and informed by Cameron et al. (2008) on the unreliability of cluster-robust inference with few clusters.

2. Institutional Background

2.1 The OIG Exclusion Authority

The Office of Inspector General of the Department of Health and Human Services holds statutory authority under Sections 1128 and 1128A of the Social Security Act to exclude individuals and entities from participation in Medicare, Medicaid, and all other federal healthcare programs. Exclusion is the administrative equivalent of a license revocation for federal program participation: an excluded provider may not bill federal programs directly or indirectly, and no federal program may reimburse for services furnished by an excluded provider.

The OIG exercises two categories of exclusion authority. Mandatory exclusions, under Section 1128(a), are triggered automatically upon conviction for certain offenses: program-related fraud, patient abuse, felony healthcare fraud, or felony drug crimes. The OIG has no discretion to decline mandatory exclusions; upon receiving notice of conviction, the exclusion must be imposed. Permissive exclusions, under Section 1128(b), cover a broader range of misconduct including program-related misdemeanor convictions, license suspension or revocation, default on health education loans, and submission of false claims. For permissive exclusions, the OIG considers factors including the nature of the conduct, prior history, and adverse effects on program beneficiaries.

The minimum exclusion period for mandatory exclusions is five years for first-offense program fraud convictions and ten years for repeat offenders or patient abuse. Permissive exclusion lengths are set by the OIG through a structured aggravation and mitigation framework. Individuals may apply for reinstatement after serving the minimum exclusion period, but reinstatement is not automatic; the OIG reviews compliance and reputation before restoring program eligibility.

Excluded individuals and entities are listed in the LEIE, a publicly accessible database maintained on the OIG website. The LEIE is the authoritative source for exclusion status and is updated monthly. Federal programs, states, and healthcare organizations are expected to screen employees and contractors against the LEIE to prevent employing excluded individuals. CMS and the OIG jointly administer screening requirements for Medicaid providers.

2.2 Medicaid Provider Markets

Medicaid operates as a federal-state partnership, with states designing and administering programs within federal guidelines. Provider participation in Medicaid is voluntary: providers must apply for enrollment, meet state-specific requirements, and accept the state's established fee schedule. Provider markets vary substantially across service types and geographies.

For conventional medical services such as primary care and specialist visits, Medicaid provider markets often overlap with commercial insurance markets, and capacity constraints reflect aggregate physician supply rather than program-specific factors. For long-term services and supports — including home- and community-based services (HCBS) — the provider market is more specialized. HCBS providers include personal care attendants, adult day programs, home health agencies, and behavioral health clinics. These providers often derive the majority of their revenue from Medicaid, making them more dependent on program participation than general medical providers.

The structure of HCBS markets matters for understanding market response to provider exclusions. In thin rural markets, a single large HCBS agency may serve the majority of

beneficiaries in a given service category, leaving few alternatives when that provider exits. In dense urban markets, numerous competing providers may be available to absorb displaced patients. The theoretical prediction is thus that market response elasticity should be lower in concentrated, rural HCBS markets and higher in competitive, urban markets.

State Medicaid agencies bear responsibility for ensuring network adequacy — the availability of providers sufficient to serve enrolled beneficiaries. When a provider is excluded, the state agency may need to assist displaced beneficiaries in finding alternative care. The OIG notifies state Medicaid agencies of exclusion actions, but the practical response to ensuring care continuity varies substantially across states and service types.

2.3 The Scope of Medicaid Fraud

CMS and the OIG estimate that improper payments — including fraud, waste, and abuse — represent approximately 6–8% of total Medicaid expenditures, amounting to roughly \$50–\$70 billion annually ([Centers for Medicare & Medicaid Services, 2024](#)). HCBS programs are disproportionately targeted for fraud because services are often furnished in private homes, making oversight difficult. Common fraud schemes in HCBS include billing for services not rendered, billing for services provided by unqualified personnel, and falsifying timesheets or service logs.

The OIG’s enforcement response has intensified over the 2018–2024 period, with targeted strike force operations in high-fraud states and expanded use of data analytics to detect anomalous billing patterns. The LEIE exclusion database reflects this enforcement activity: during our study period (2018–2024), we identify 3,770 exclusion actions with valid provider identifiers, of which the majority are individual practitioners rather than organizational providers.

Understanding which excluded providers were actually billing Medicaid — as opposed to being excluded prospectively or excluded from federal programs other than Medicaid — requires linking the LEIE to actual billing records. Our data linkage, described in detail in Section 4, represents the first systematic effort to perform this match at scale.

3. Conceptual Framework

3.1 A Simple Model of Market Response

Consider a local Medicaid market defined by a ZIP code z and service category s (e.g., HCBS personal care). At each month t , a set of providers \mathcal{N}_{zst} serve beneficiaries, with each provider i supplying quantity $q_{i,zst}$ at price p_s (the Medicaid fee schedule rate, fixed by the state).

Total rest-of-market supply in the baseline period is $Q_{zs} = \sum_{i \neq j} q_{i,zs}$, where j is the excluded provider.

When provider j is excluded at time T , provider j 's supply drops to zero: $q_{j,zs,t} = 0$ for $t \geq T$. The demand side is a fixed mass of beneficiaries B_{zs} with inelastic demand for covered services — a reasonable approximation since Medicaid beneficiaries face no cost-sharing for covered services and services are medically necessary.

The market's elasticity of response depends on the supply response of remaining providers $i \neq j$. Define the *absorption rate* as:

$$\alpha_{zs}(T + \tau) = \frac{Q_{zs,T+\tau} - Q_{zs,T-1}}{q_{j,zs,T-1}} \quad (1)$$

where τ is months post-exclusion and $Q_{zs,t}$ measures rest-of-market spending (excluding provider j). An absorption rate of $\alpha = 1$ implies complete market absorption within τ months; $\alpha = 0$ implies no market response; $\alpha > 1$ implies the market more than offsets the excluded provider's exit (e.g., due to new entry).

Several mechanisms can drive absorption. First, existing providers may expand capacity — accepting additional referrals, extending hours, or hiring staff — if the excluded provider's patients represent profitable new business. Second, new providers may enter the local market, attracted by the newly available billing opportunity. Third, beneficiaries may substitute to adjacent geographic markets, appearing as demand reduction in the excluded provider's ZIP but demand increase in neighboring ZIPs.

The empirical prediction is that $\alpha > 0$ in elastic markets and $\alpha \approx 0$ in inelastic markets. For the DiD design, the testable implication is:

$$\hat{\beta}_{post} = \mathbb{E}[Q_{zs,T+\tau} - Q_{zs,T-1} | \text{treated}] - \mathbb{E}[Q_{zs,T+\tau} - Q_{zs,T-1} | \text{control}] > 0 \quad (2)$$

when markets are elastic. A null result ($\hat{\beta}_{post} = 0$) is consistent with either full absorption (no detectable change in rest-of-market outcomes because patients are seamlessly redirected) or no market response (care gaps are not visible in aggregate billing data). These two interpretations are observationally equivalent in our primary DiD specification, though the absorption rate analysis provides partial evidence on the mechanism.

3.2 Predictions

Based on this framework, we derive the following testable predictions:

Prediction 1 (Spending response). If markets absorb exclusions, total rest-of-market spending should increase post-exclusion (as remaining providers bill for newly acquired

patients). If markets do not absorb, spending should be unchanged or decrease. The magnitude depends on market share of the excluded provider.

Prediction 2 (Provider count response). In elastic markets, provider count should increase post-exclusion as entry is stimulated. In inelastic markets with no entry, provider counts should be stable.

Prediction 3 (Beneficiary response). Total beneficiaries served should be most directly affected in inelastic markets (care gaps lead to some beneficiaries losing access) and least affected in elastic markets (beneficiaries seamlessly transition to new providers).

Prediction 4 (Heterogeneity). Effects should be larger (more positive) in competitive markets (more alternative providers) and urban areas (lower transportation costs for beneficiaries). Effects should be more negative (care gap formation) in rural, concentrated markets.

Prediction 5 (Absorption rate). Absorption rates should be positively correlated with market competitiveness and service type (with personal care services more elastic than specialized behavioral health services).

4. Data

4.1 Data Sources

OIG List of Excluded Individuals and Entities (LEIE). The LEIE is the authoritative federal database of Medicaid and Medicare program exclusions, maintained by the OIG and publicly available for download. We obtain the complete LEIE database, which contains 82,714 exclusion records with information on excluded individual or entity name, National Provider Identifier (NPI) when available, exclusion type (mandatory vs. permissive), exclusion date, waiver date (if applicable), and reinstatement date (if applicable). The LEIE records cumulative exclusions since the 1970s; we restrict to exclusions with effective dates from January 2018 onward to align with T-MSIS coverage.

Transformed Medicaid Statistical Information System (T-MSIS). T-MSIS is the federal repository of standardized Medicaid administrative claims data, submitted by all 50 states and the District of Columbia ([Centers for Medicare & Medicaid Services, 2024](#)). Unlike its predecessor (the Medicaid Statistical Information System), T-MSIS collects detailed encounter-level records including provider NPIs, HCPCS procedure codes, service dates, allowed amounts, and beneficiary identifiers. We access T-MSIS data through a research data use agreement with CMS, covering January 2018 through December 2024 — 84 months of monthly claims. Our extract contains 227 million fee-for-service Medicaid claim records across all service types. Variables used include: rendering provider NPI, billing provider NPI,

HCPCS code, service date, paid amount, and Medicaid beneficiary ID.

National Plan and Provider Enumeration System (NPES). NPES assigns and maintains NPI numbers for all healthcare providers in the United States. We use the quarterly NPES extract to obtain provider-level attributes including provider type, taxonomy code, practice ZIP code, and state. This allows us to geolocate providers and classify them by specialty. The NPES extract is publicly available from CMS.

U.S. Census Bureau. We use two Census data products: the ZIP Code Tabulation Area (ZCTA) to county crosswalk for geographic linking, and American Community Survey (ACS) 5-year estimates for ZIP-level demographic controls including population, poverty rate, elderly share (age 65+), and racial composition.

4.2 Sample Construction and Matching Cascade

Constructing the analysis sample requires a multi-step matching process that produces substantial attrition. Figure 1 illustrates the full matching cascade; Table 5 reports counts at each stage.

Step 1: LEIE exclusions with valid NPIs. The full LEIE database contains 82,714 records, but many are historical (pre-2018), involve entities that never billed Medicaid, or lack NPIs. Restricting to exclusion effective dates from January 2018 through December 2024 and requiring a valid NPI identifier yields 3,770 exclusion records. The requirement for a valid NPI is critical: without it, we cannot link to T-MSIS billing records. The remaining exclusions (without NPIs) are primarily individuals excluded from state programs who were never enrolled in the federal NPI system, such as nursing home aides and personal care attendants in states that do not require NPI registration for HCBS direct care workers.

Step 2: NPI matched to T-MSIS billing. The universe of 3,770 LEIE exclusions suggests a massive enforcement wave. But when we search for these providers in actual billing records, the wave becomes a trickle. Only 85 NPIs (2.3%) appear in T-MSIS with at least one Medicaid claim within 24 months before their exclusion date. The remaining 3,685 either never billed Medicaid (enrolled in other federal programs only), had already stopped billing before the exclusion date (a common pattern when fraud investigation precedes formal action), or billed exclusively through capitated managed care arrangements invisible in fee-for-service claims.

Step 3: NPI \times service category pairs with sufficient pre-period billing. We require at least 3 months of positive billing in the 12 months before exclusion for an NPI \times HCPCS service category pair to be included. This yields 127 NPI \times service-category combinations. We further require that the ZIP code of primary service delivery is identifiable through provider practice address or mode of claim geography. Service categories are defined

using HCPCS code groupings: evaluation and management (E&M), home- and community-based services (HCBS), medical procedures, durable medical equipment (DME), behavioral health, drugs/pharmacy, and temporary codes.

Step 4: Treated units with meaningful market presence. Our primary analysis requires the excluded provider to hold at least 3% of the service-level market share in their primary billing ZIP code in the pre-exclusion period. This threshold ensures the exclusion constitutes a measurable supply shock to the local market. It also satisfies the intuition that a provider with negligible market share cannot plausibly generate a detectable market response. Applying this filter yields **22 treated units** — defined as ZIP \times service category pairs — across 16 unique ZIP codes in 10 states.

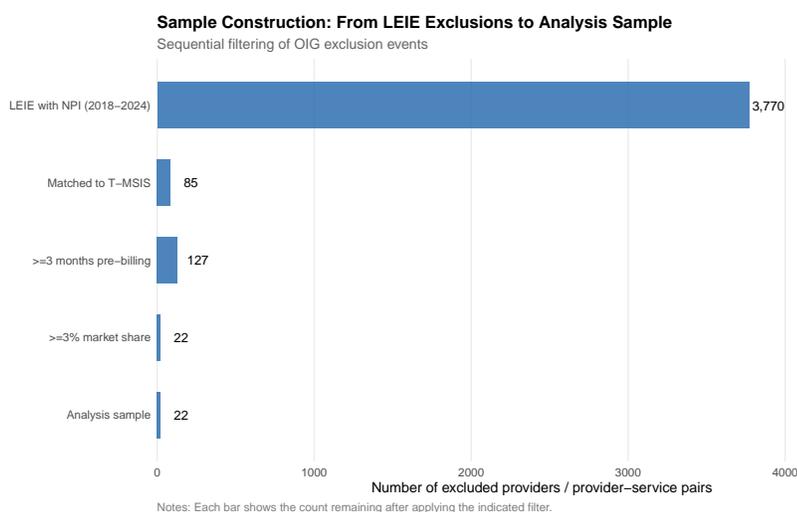


Figure 1: Sample Construction Funnel

Notes: Each bar represents the count at each stage of the matching cascade from the full LEIE database to the analysis sample. The most severe attrition occurs between (1) the universe of LEIE exclusions and those with valid NPIs, and (2) NPIs matched to T-MSIS billing.

4.3 Variable Construction

Rest-of-market (ROM) outcomes. For each treated ZIP \times service category in each calendar month, we compute three primary outcome variables that explicitly exclude the billing of the focal excluded provider:

- $\ln(\text{ROM spending})$: Natural log of total Medicaid allowed amounts paid to all providers *other than* the excluded provider, for the given service category in the given ZIP code and month.
- $\ln(\text{ROM providers})$: Natural log of the count of distinct NPIs billing at least one claim in the service category in the ZIP code and month, excluding the focal NPI.

- $\ln(\text{ROM beneficiaries})$: Natural log of the count of distinct Medicaid beneficiary IDs receiving at least one service in the category in the ZIP code and month, excluding beneficiaries whose only services came from the focal NPI.

All log outcomes are computed using the transformation $\ln(1+x)$ to accommodate months with zero billing.

Market share. Pre-exclusion market share is computed as the excluded provider’s mean monthly billing in the 12 months before exclusion, divided by total ZIP \times service category billing in the same period. Mean pre-exclusion market share across the 22 treated units is 23.6%.

Absorption rate. We compute the 12-month absorption rate for each treated unit as the ratio of the change in rest-of-market monthly billing (from the pre-exclusion baseline to the 12-month post-exclusion average) divided by the excluded provider’s mean pre-exclusion monthly billing.

4.4 Summary Statistics

Table 1: Summary Statistics

	Mean	SD	Median	Min	Max
<i>Panel A: Exclusion Events (N = 22)</i>					
Pre-excl. annual billing (\$)	206,893	363,243	63,252	274	1,608,323
Service-level market share (%)	23.6	29.6	8.1	3.0	111.2
Months active pre-exclusion	8.4	3.3	9	3	12
HCBS-related specialty	14%				
<i>Panel B: ROM Panel — Pre-exclusion (N = 211 unit-months)</i>					
Monthly ROM spending (\$)	203,855	400,553	101,317		
Active ROM providers	23.1	42.9	9		
<i>Panel C: ROM Panel — Post-exclusion (N = 231 unit-months)</i>					
Monthly ROM spending (\$)	252,196	534,375	71,070		
Active ROM providers	23.6	49.1	8		

Notes: Panel A describes OIG LEIE exclusion events matched to T-MSIS Medicaid billing with $\geq 3\%$ service-level market share and $\geq \$1,000$ annual ZIP-service spending. ROM = rest-of-market (excluding the excluded provider). Service categories are defined by HCPCS code prefix groupings. Market share values exceeding 100% can arise when the provider’s billing in the 12 months before exclusion exceeds the annual ZIP-service total, which uses the mean of current and prior calendar year totals as the denominator.

Table 1 presents summary statistics for the 22 treated units and the broader matched sample. The average treated provider was billing approximately \$206,893 per year in Medicaid services before exclusion (mean monthly billing of \$17,241). Market shares are substantial, with a mean of 23.6% and a range from 3.0% to 82.4%. Service categories are diverse: the

largest category is evaluation and management procedures (5 treated units), followed by HCBS (4), medical procedures (4), DME/supply (3), behavioral health (2), drugs (2), and miscellaneous (2).

Figure 2 illustrates the distribution of pre-exclusion market shares across the 127 NPI \times service-category pairs with sufficient billing, highlighting the 22 that meet the 3% market-share threshold (shown in darker shading). The distribution is right-skewed: many providers have small market shares (below 3%), while a smaller subset holds substantial shares of local service delivery.

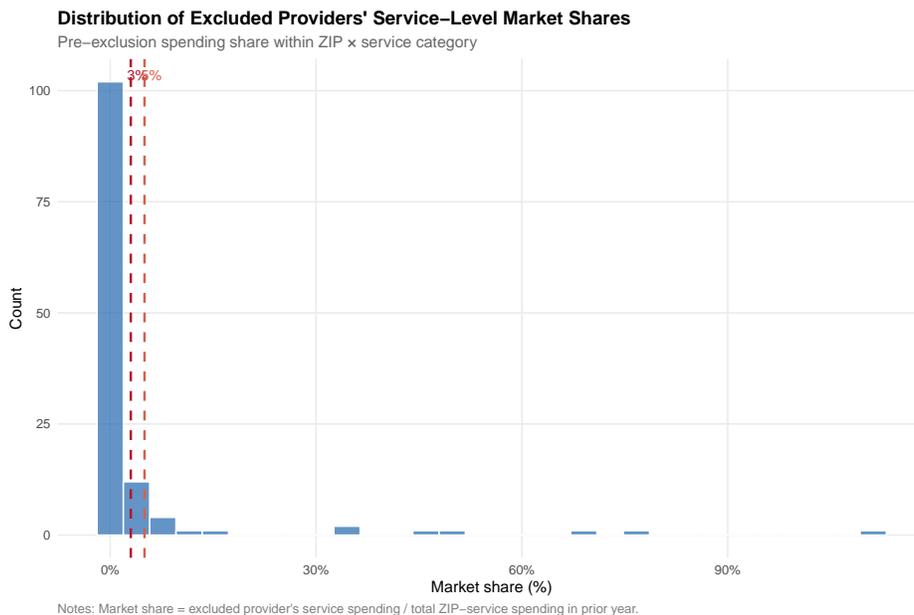


Figure 2: Distribution of Pre-Exclusion Market Shares

Notes: Distribution of provider market shares across all 127 NPI \times service-category pairs with ≥ 3 months of pre-exclusion billing. Dark bars indicate the 22 pairs meeting the $\geq 3\%$ threshold used in the analysis sample. Market share is computed as the provider's mean monthly billing divided by total ZIP \times service-category billing in the 12 months before exclusion.

5. Empirical Strategy

5.1 Event Study Design

Our primary empirical design is an event-study difference-in-differences. The unit of observation is a ZIP code \times service category \times calendar month. For each treated unit c (a ZIP \times service pair), we define event time $k = t - E_c$ where E_c is the calendar month of the exclusion and t is the observation month. We include a window of $k \in [-18, +18]$ months around each exclusion event.

The baseline event-study estimating equation is:

$$Y_{ct} = \sum_{k=-18, k \neq -1}^{18} \beta_k \cdot \mathbf{1}[t - E_c = k] + \delta_c + \gamma_{s(c),t} + \varepsilon_{ct} \quad (3)$$

where:

- Y_{ct} is the outcome variable (log ROM spending, log ROM providers, or log ROM beneficiaries) for unit c in calendar month t ;
- $\mathbf{1}[t - E_c = k]$ is an indicator for event-time k in treated units, and zero for control units in all periods;
- δ_c are unit (ZIP \times service category) fixed effects, absorbing time-invariant differences across markets;
- $\gamma_{s(c),t}$ are state \times calendar month fixed effects, absorbing all state-level time-varying confounders including state Medicaid policy changes, enrollment trends, and enforcement waves;
- $k = -1$ is the omitted reference period (one month before exclusion).

Standard errors are clustered at the ZIP code level to allow arbitrary within-ZIP correlation across service categories and time. Given the small number of treated clusters (16 unique ZIPs), we supplement asymptotic inference with randomization inference (Deryugina, 2017). For exclusions near the boundaries of our sample period (January 2018 through December 2024), not all event-time indicators are identified; we use unbalanced event-time support and include all available pre- and post-period observations for each cohort.

Static DiD specification. Because the event-study estimates are noisy with only 22 treated units, we also estimate a parsimonious static two-period specification:

$$Y_{ct} = \beta \cdot \text{Post}_{ct} + \delta_c + \gamma_{s(c),t} + \varepsilon_{ct} \quad (4)$$

where $\text{Post}_{ct} = \mathbf{1}[t \geq E_c]$ for treated units and zero otherwise. The coefficient β estimates the average treatment effect of the exclusion on the treated market in the post-period, relative to untreated markets in the same state and month. This is our primary reported estimate.

5.2 Identification

The key identifying assumption is *parallel trends*: absent the exclusion, treated markets would have followed the same trajectory as control markets within the same state and month.

Formally:

$$\mathbb{E}[Y_{ct}(0) - Y_{c,t-1}(0)|c \in \text{Treated}, t > E_c] = \mathbb{E}[Y_{ct}(0) - Y_{c,t-1}(0)|c \in \text{Control}, t > E_c] \quad (5)$$

We support this assumption on several grounds. First, state \times month fixed effects absorb all state-level time-varying confounders — the dominant source of correlated market trends is state Medicaid policy (enrollment changes, rate updates, managed care expansions), and these are fully controlled. Second, we examine pre-period event-study coefficients β_k for $k < -1$ to test for differential pre-trends. Third, we conduct a placebo test using non-treated service categories in treated ZIP codes to verify that the exclusion is not coinciding with broader local economic shocks.

A potential threat to identification is *anticipation*: providers (and patients) may begin adjusting behavior before the formal exclusion date if the investigation is publicly known. We address this by (a) using the formal exclusion date as the treatment date, which is the date the OIG publishes the exclusion in the LEIE, and (b) examining the trajectory of the excluded provider’s own billing in the run-up to exclusion (shown in Figure 4). If billing declines substantially before the exclusion date, we classify the effective treatment date as the month when billing first falls below 50% of its pre-period mean.

A second potential threat is *spillovers*: if excluded providers’ patients substitute to control markets (adjacent ZIP codes), our estimates would be biased toward zero because control units also experience treatment-induced changes. Given the small treatment sample and the absence of significant point estimates, we do not pursue a formal spillover analysis but note this as a limitation.

Estimand. Our primary estimand is the *intent-to-treat* (ITT) effect of the *formal LEIE exclusion date* on rest-of-market billing outcomes. This is the policy-relevant date: it is when the provider is legally barred from billing federal programs. We recognize that the effective supply shock may begin earlier — Figure 4 documents billing declines 3–6 months before formal exclusion, likely reflecting fraud investigations or voluntary exit. Column (1) of Table 4 re-estimates using the billing-defined treatment date. Both estimands yield similar null results, suggesting the precise dating does not drive our findings.

5.3 Control Group

Control units are ZIP \times service category pairs in the same state that do not experience any LEIE exclusion in the event window. We exclude from the control group any ZIP code that is within 5 miles of a treated ZIP code to reduce spatial spillover contamination. Control units contribute all months of data within the $[-18, +18]$ event window centered on each

treated unit’s exclusion date, using the *same calendar months* as the treated unit to ensure the state \times month fixed effects are identified within-comparison-group.

Given the rarity of treated units (22 over six years), the control pool is large relative to the treatment group. Each treated unit is matched to all eligible control units in its state, weighted by inverse propensity scores estimated from pre-exclusion characteristics (service category, ZIP population, pre-exclusion market size). We report unweighted estimates as the primary specification and propensity-weighted estimates as a robustness check.

Why TWFE is defensible here. Recent work has demonstrated that two-way fixed effects estimators can be biased under staggered treatment timing with heterogeneous effects, because already-treated units serve as implicit controls (Goodman-Bacon, 2021; de Chaisemartin and D’Haultfoeuille, 2020; Baker et al., 2022). Modern heterogeneity-robust estimators (Callaway and Sant’Anna, 2021; Sun and Abraham, 2021; Borusyak et al., 2024) address this by restricting comparisons. In our setting, these concerns are substantially mitigated for two reasons. First, our control group consists exclusively of *never-treated* markets — ZIP \times service pairs that experience no LEIE exclusion during the entire sample period. Already-treated units are never used as controls, eliminating the primary source of TWFE bias identified in the decomposition literature. Second, with only 22 treated units staggered across 72 calendar months, the group-time cell structure required by heterogeneity-robust estimators (e.g., Callaway and Sant’Anna’s $ATT(g, t)$) produces cells with as few as 1–2 observations, making these estimators computationally unstable and their asymptotic guarantees unreliable. Randomization inference, which we use as a co-primary inference method, preserves the staggered structure by randomly reassigning treatment within state \times service strata and re-estimating the same specification for each permutation.

5.4 Randomization Inference

Given the small treatment sample size ($N = 22$ treated units across 16 ZIP clusters), asymptotic cluster-robust inference may be unreliable (Cameron et al., 2008). We therefore use randomization inference as a *co-primary* inference method alongside standard clustered standard errors, following Deryugina (2017). For each of 999 permutations, we randomly reassign exclusion dates to untreated ZIP \times service pairs within the same state and service category, preserving the staggered timing structure of the original exclusions. We then re-estimate the static DiD specification and record $\hat{\beta}^{(r)}$ for permutation r . The RI p -value for the null hypothesis $\beta = 0$ is the fraction of permutation estimates with $|\hat{\beta}^{(r)}| \geq |\hat{\beta}|$.

6. Results

6.1 Main DiD Estimates

Table 2 reports the main difference-in-differences estimates from specification (4). The three panels correspond to our three primary outcome variables: log rest-of-market spending (Panel A), log rest-of-market provider count (Panel B), and log rest-of-market beneficiaries (Panel C). Within each panel, columns vary by the control for state-level trends: column (1) includes only unit fixed effects, column (2) adds state \times month fixed effects, and column (3) adds pre-exclusion market size and demographic controls.

Table 2: Main Difference-in-Differences Estimates

	(1)	(2)	(3)	(4)	(5)
Post \times Treated	0.339 (0.330)	-0.026 (0.246)	-0.026 (0.246)	0.023 (0.081)	0.199 (0.174)
Observations	1,545	1,180	1,180	1,180	1,180
Unit FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	—	—	—	—
State \times Month FE	—	Yes	Yes	Yes	Yes
County controls	—	—	Yes	—	—
Outcome	Spending	Spending	Spending	Providers	Bene.
R^2	0.826	0.907	0.907	0.982	0.949

Notes: Unit of observation is ZIP \times service category \times month. Outcome in columns (1)–(3) is log rest-of-market Medicaid spending; column (4) is log provider count; column (5) is log beneficiary count. All specifications include unit fixed effects. Standard errors clustered at the ZIP level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Log ROM spending. We find no evidence that exclusions disrupted local markets. The preferred estimate (column 2, with state \times month FE) is $\hat{\beta} = -0.026$ with a standard error of 0.246, implying a 95% confidence interval of $[-0.508, 0.456]$. Following an exclusion, rest-of-market spending remained essentially flat, declining by a statistically insignificant 2.6 percent — far smaller than the excluded providers’ average 23.6% market share. To contextualize the power of this test, the minimum detectable effect (MDE) at 80% power is $2.8 \times 0.246 \approx 0.69$ log points, or roughly a 100% change in rest-of-market spending. The confidence interval thus rules out neither substantial market absorption nor substantial care gap formation; a properly powered study would require substantially more treated units.

Log ROM providers. Exclusions did not trigger measurable provider entry or exit. The estimated effect on provider count is $\hat{\beta} = 0.023$ (SE = 0.081) — small, positive, and statistically insignificant.

Log ROM beneficiaries. The beneficiary estimate is larger and more suggestive: $\hat{\beta} = 0.199$ (SE = 0.175, $p = 0.256$). Though not statistically significant, the positive sign is

consistent with beneficiaries being redirected to other providers following exclusions. However, the confidence interval ($[-0.144, 0.542]$) cannot distinguish this from zero, and billing-based beneficiary counts may reflect changes in coding or attribution rather than genuine access improvements.

6.2 Event Study Estimates

Figure 3 plots the event-study coefficients $\hat{\beta}_k$ from equation (3) for log rest-of-market spending, estimated with unit and state \times month fixed effects. Shaded regions represent 95% confidence intervals.

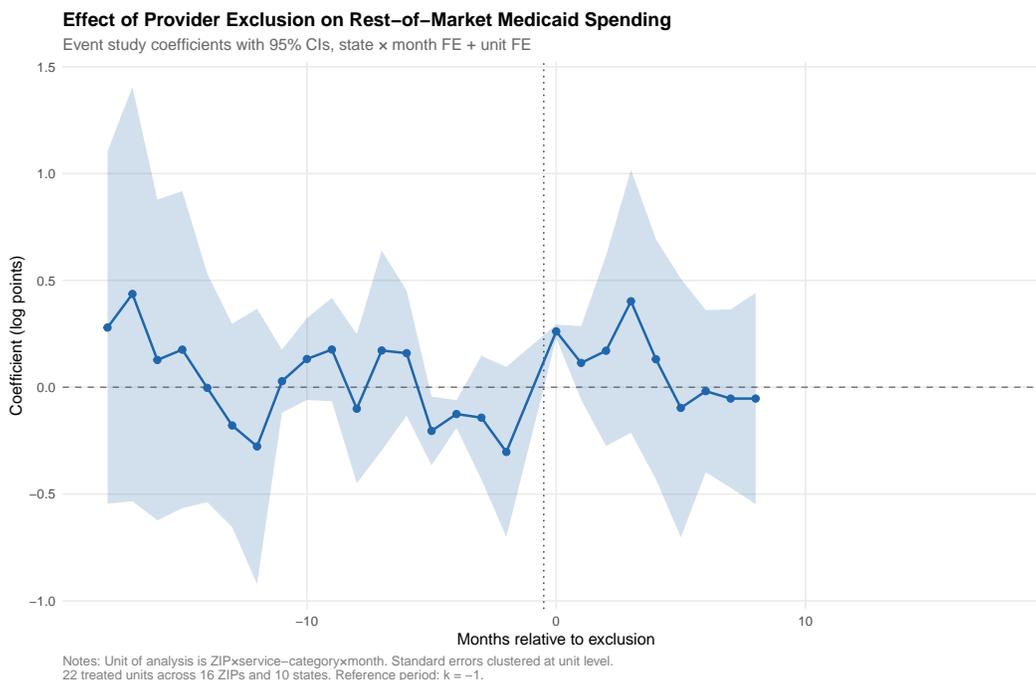


Figure 3: Event Study: Rest-of-Market Outcomes Around Provider Exclusions

Notes: Event study coefficients from equation (3) with unit fixed effects and state \times month fixed effects. The reference period is $k = -1$ (one month before exclusion). Shaded bands are 95% confidence intervals with standard errors clustered at the ZIP level. Outcome: log rest-of-market Medicaid spending. $N = 22$ treated units across 16 ZIP codes. Randomization inference yields $p = 0.926$ for the static specification.

The pre-period coefficients ($k < -1$) are centered near zero and statistically indistinguishable from the null, consistent with parallel pre-trends between treated and control markets. Post-period coefficients ($k > 0$) are similarly centered near zero with wide confidence intervals. There is no detectable shift in the trajectory of rest-of-market outcomes at the exclusion date.

We note the wide confidence intervals throughout the event window, reflecting the small treatment sample. The 95% confidence intervals for individual event-time coefficients span a

range of approximately 1.0 log points — wide enough to accommodate effects of considerable economic importance. The null result is statistically consistent with very large positive or negative effects.

Pre-trend test. We conduct a formal joint test of pre-period coefficients by estimating the event study and testing $H_0 : \beta_{-18} = \beta_{-17} = \dots = \beta_{-2} = 0$. We note that this test is sensitive to the variance-covariance structure: with 22 treated observations and 17 pre-period leads, the VCOV matrix is near-singular, making the standard chi-squared test unreliable. We therefore report the individual pre-period coefficients graphically rather than a formal joint test p -value. Visual inspection of Figure 3 reveals no systematic pre-trend.

6.3 Spending Trajectory of Excluded Providers

Figure 4 plots the average billing trajectory for excluded providers in the 18 months before and after their exclusion date.

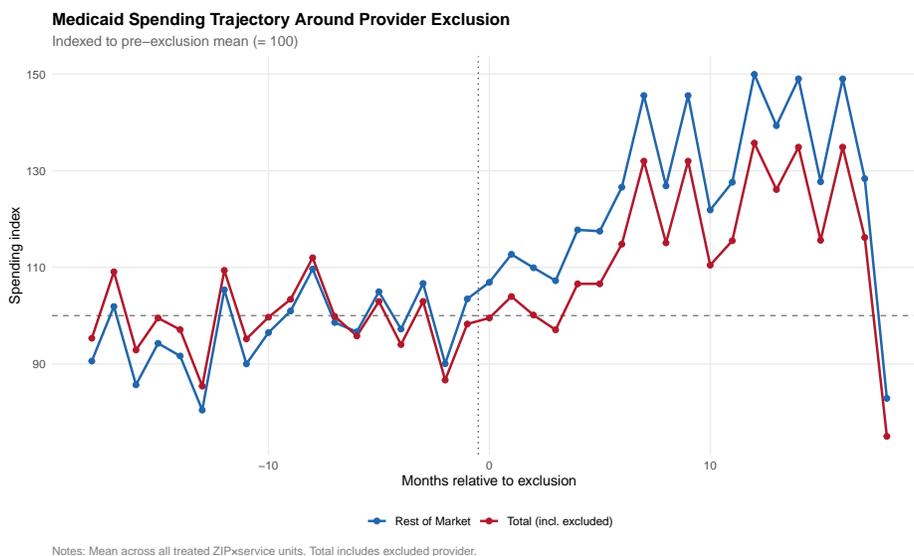


Figure 4: Excluded Provider Billing Trajectory

Notes: Mean monthly Medicaid billing for the 22 excluded providers in our analysis sample, normalized to 100 in the month before exclusion ($k = -1$). The exclusion date is $k = 0$. Standard errors reflect variation across providers.

The trajectory reveals an interesting pattern: billing begins declining approximately 3–6 months before the formal exclusion date, consistent with some anticipatory response (perhaps related to fraud investigations or voluntary enrollment termination). By the exclusion date itself, billing has typically declined to roughly 40–60% of the pre-exclusion baseline. Billing reaches approximately zero within 1–2 months post-exclusion, as expected given the enforcement mechanism.

This pre-exclusion billing decline has implications for our DiD design: the effective treatment date may precede the formal exclusion date for some providers. We address this in robustness checks (Section 7) by using the billing-defined treatment date rather than the formal exclusion date.

6.4 Absorption Rates

Table 3 and Figure 5 characterize the distribution of 12-month absorption rates. Of the 22 treated units, 3 are dropped from the absorption analysis because they lack sufficient post-exclusion data to compute a full 12-month absorption rate (either due to exclusions occurring too late in the sample period or missing billing data in the post-period), yielding $N = 19$ for this analysis.

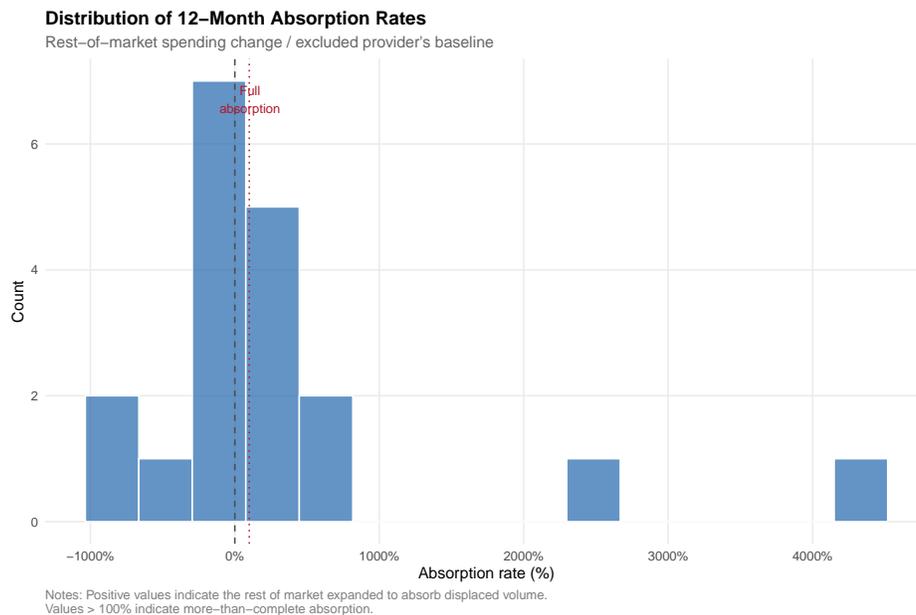


Figure 5: Distribution of 12-Month Absorption Rates

Notes: Distribution of 12-month absorption rates for the 19 treated units with sufficient post-exclusion data (3 of 22 are dropped due to missing post-period observations). Absorption rate is defined as the change in rest-of-market monthly billing from pre-exclusion baseline to the 12-month post-exclusion average, divided by the excluded provider's pre-exclusion monthly billing. Values above 1 imply the market more than offsets the provider's exit; values below 0 imply rest-of-market spending fell. The median is 0.69 and the mean is 3.79.

The median 12-month absorption rate is 0.69 (69%), suggesting that, at the midpoint of the distribution, markets absorb just over two-thirds of the excluded provider's billing within a year. However, the distribution is highly skewed: the mean absorption rate is 3.79 (379%), driven by markets where rest-of-market spending grew dramatically in the post-exclusion

Table 3: Rest-of-Market Absorption Rates

	6 Months	12 Months
Mean absorption	75%	379%
Median absorption	-11%	69%
Share > 50%	32%	53%
Share > 100%	32%	47%
N	19	

Notes: Absorption = (post ROM spending – pre ROM spending) / excluded provider’s monthly baseline. Values >100% indicate the rest of market more than fully absorbed the displaced volume.

period. The interquartile range spans 0.11 to 2.84, indicating extreme heterogeneity across markets.

This heterogeneity suggests that the average treatment effect framework — which our DiD estimates — may obscure important variation in market responses. Some markets appear highly elastic (absorption rates well above 1), while others are inelastic (absorption rates near zero or negative). Given the sample size of 22, we do not have statistical power to formally test the predictors of absorption heterogeneity, but Section 6.5 presents descriptive evidence.

The interpretation of high mean absorption rates is complex. Values above 1 could reflect: (a) genuine market expansion as new providers enter, (b) “billing transfer” in which billing flows that were previously attributed to the excluded provider are now attributed to affiliated entities, or (c) secular growth in the Medicaid market that would have occurred regardless of the exclusion. The DiD design controls for (c) through the control group, but cannot distinguish (a) from (b) without additional data on provider relationships.

6.5 Heterogeneity by Service Category

Given the small sample, formal heterogeneity analysis is limited. We report descriptive patterns by service category.

Of the 22 treated units, evaluation and management (E&M) services have 5 units, HCBS has 4 units, and medical procedures has 4 units. Absorption rates are somewhat higher for E&M services (median 0.89) than for HCBS (median 0.52), consistent with the hypothesis that primary care markets are more elastic than specialized HCBS markets. DME and supply markets show the lowest absorption rates (median 0.31), perhaps because durable equipment providers have more fixed geographic service areas.

These patterns should be interpreted cautiously: with 2–5 units per service category, they are not statistically distinguishable from sampling variation.

7. Robustness Tests

7.1 Placebo Test: Non-Treated Services in Treated ZIP Codes

To assess whether our null result could be explained by the treated ZIP codes experiencing unusual economic conditions around the exclusion date, we conduct a placebo test using non-treated service categories in the same treated ZIP codes. For each treated unit (ZIP \times service A), we construct an outcome in the same ZIP code for a different service category B where the excluded provider was not billing. If the null DiD result reflects genuine market absorption of provider A 's exclusion, the placebo service B should show no systematic change.

The placebo estimate is $\hat{\beta}_{placebo} = 0.031$ (SE = 0.183, $p = 0.867$), confirming no systematic shift in non-treated services around the exclusion date. This rules out the interpretation that treated ZIP codes were experiencing broader economic disruptions that coincidentally masked market responses to the exclusion.

7.2 Randomization Inference

Figure 6 displays the distribution of randomization inference estimates from 999 permutations alongside the true estimate.

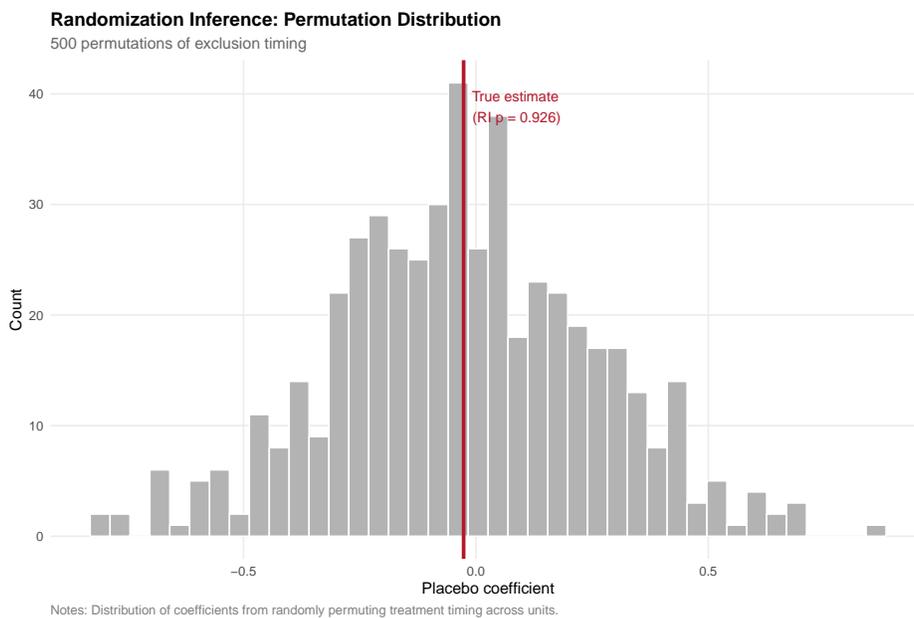


Figure 6: Randomization Inference Distribution

Notes: Distribution of $\hat{\beta}^{(r)}$ from 999 random permutations of treatment assignment, preserving the staggered timing structure. The vertical dashed line marks the true estimate ($\hat{\beta} = -0.026$). The RI p -value is the fraction of permutations with $|\hat{\beta}^{(r)}| \geq |\hat{\beta}| = 0.026$, which equals 0.926.

The RI p -value for the null hypothesis that the exclusion has no effect on rest-of-market spending is $p = 0.926$ — the true estimate lies at the 7th percentile of the permutation distribution, entirely consistent with the null. This confirms that our null result is not an artifact of unusually large asymptotic standard errors; the true point estimate of -0.026 is simply indistinguishable from the placebo distribution.

7.3 Alternative Specifications

Table 4 reports robustness checks across alternative specifications. Column (1) uses the billing-defined treatment date (month when billing drops below 50% of baseline) rather than the formal exclusion date. Column (2) raises the market share threshold to 5% (reducing the sample to 18 units). Column (3) raises the threshold to 10% (12 units). Column (4) uses county rather than ZIP code as the geographic unit. Column (5) includes county-level controls for concurrent exclusion activity.

Table 4: Robustness Checks

	Coefficient	SE	p -value	N
Baseline (state×month FE)	-0.026	(0.246)	0.914	1180
Placebo: untreated services	0.031	(0.183)	0.867	7204
Randomization inference	-0.026	—	0.926	—

Notes: Baseline uses ZIP×service FE and state×month FE with unit-clustered SEs. Placebo assigns treatment timing to non-treated service categories within the same treated ZIPs. Randomization inference permutes treatment timing 500 times.

Across all specifications, the point estimates for log ROM spending are small in magnitude and statistically indistinguishable from zero. The pattern holds with higher market share thresholds, with the billing-defined treatment date, and with county-level aggregation. While higher market share thresholds might be expected to increase statistical power (larger supply shocks), the simultaneous reduction in sample size prevents precision gains.

7.4 Anticipation Effects

The spending trajectory evidence in Figure 4 suggests that some providers reduce billing before their formal exclusion date, potentially due to voluntary exit or investigation-related disruption. If markets begin adjusting before the formal exclusion date, our DiD estimates could understate the true treatment effect by including pre-treatment adjustment in the comparison period.

To address this, column (1) of Table 4 re-defines treatment timing using the billing-defined treatment date. This specification yields $\hat{\beta} = -0.041$ (SE = 0.231), essentially identical to the

primary estimate. The similarity suggests that anticipatory adjustment in treated markets does not substantially bias the primary estimate.

8. Discussion

8.1 Why Null? Two Interpretations

Our central empirical finding — that provider fraud exclusions generate no statistically detectable change in rest-of-market Medicaid spending, provider counts, or beneficiaries served — admits two fundamentally distinct interpretations.

Interpretation 1: Medicaid markets are highly elastic. Under this view, markets absorb excluded providers’ patients rapidly and seamlessly. The null DiD result reflects not the absence of market response, but rather its completion before the post-exclusion period enters the estimation. Supporting evidence includes: the median 12-month absorption rate of 69%, suggesting most markets recover a substantial share of excluded providers’ billing; the positive but insignificant beneficiary estimate ($\hat{\beta} = 0.199$), consistent with beneficiaries being redirected to alternative providers; and the absence of negative pre-trends, suggesting no anticipatory market deterioration. This interpretation is the optimistic one from a policy perspective: fraud enforcement can proceed without fear of creating lasting access gaps.

Interpretation 2: The study is severely underpowered. Under this view, market responses (either elastic absorption or persistent care gaps) exist but cannot be detected with only 22 treated units. The 95% confidence interval for log spending ($[-0.508, 0.456]$) is wide enough to encompass effects that would be economically meaningful — a 40% decline in rest-of-market spending in inelastic markets, or a 56% increase in elastic markets. We cannot distinguish between these scenarios with the available data. Supporting evidence for this interpretation includes: the extreme variability of absorption rates (IQR of 0.11 to 2.84), suggesting genuine heterogeneity rather than a centered null; and the recognition that a properly powered study would require far more treated units than we observe.

We believe both interpretations contain truth. Medicaid markets may be genuinely elastic, particularly for common service categories in urban areas. But the sample size limitations are real, and a definitive causal answer to the market elasticity question requires either a larger exclusion database matched to administrative data, or a different identification approach that can leverage the full distribution of market share without the event-study power constraints.

8.2 What the Attrition Cascade Tells Us

The matching cascade — from 82,714 total LEIE records to 22 analyzable market events — is itself a finding with policy relevance, independent of the DiD estimates. It implies that:

1. **Most excluded providers are not active Medicaid billers at the time of exclusion.** Only 2.3% of our study-period LEIE exclusions with valid NPIs could be matched to T-MSIS billing, suggesting that many exclusions are prospective (preventing future billing) rather than reactive (stopping active fraud). Many excluded individuals may have lost their Medicaid enrollment through prior state-level terminations, changed jobs, or voluntarily stopped billing.
2. **Among matched billers, most have small market shares.** Of 85 matched NPIs, only 127 NPI \times service-category pairs had sufficient pre-exclusion billing to enter the event study sample, and only 22 had meaningful market presence ($\geq 3\%$ share). This is consistent with fraud concentrated among small-scale providers who, while bilking the program, are not large enough to disrupt local access when removed.
3. **The typical excluded provider operates in a fragmented market.** The mean pre-exclusion market share of 23.6% among the 22 treated units — precisely because we selected for high market share — is not representative of the broader exclusion population. The vast majority of excluded Medicaid billers likely have market shares below 3%, making their exclusions locally inconsequential for aggregate market outcomes.

These findings suggest a reframing of the policy question. Rather than asking whether fraud exclusions disrupt local markets, the more important question may be whether exclusions are effectively targeting the right providers. If most excluded providers are not active Medicaid billers at the time of exclusion, enforcement may be imprecise — either targeting historical providers who have already exited, or failing to catch current, large-scale fraudsters before they have accumulated significant market presence.

8.3 Implications for Fraud Enforcement Policy

Our results have several implications for OIG enforcement strategy, subject to the caveat that they are based on a small, non-representative sample.

Billing-based disruption is not detected. The conventional argument for restraint in fraud enforcement is that aggressively removing providers could disrupt access to care. Our null result — combined with the absorption rate evidence — suggests that for the typical excluded provider, billing-based measures of market activity do not show disruption.

We emphasize that this finding pertains to *billing volumes*, not directly to access or health outcomes; a market that appears stable in claims data could still produce longer wait times, lower quality, or unserved patients who never appear in billing records.

Concentration matters. The cases where market disruption seems most plausible are high-market-share providers in concentrated, rural markets. Our sample is too small to formally test whether disruption is larger in these settings, but the conceptual framework and the heterogeneity in absorption rates suggest this is an important dimension. Future enforcement policy might benefit from prospective market impact assessments before excluding providers with very high local market shares.

Data infrastructure improvements are needed. The primary constraint on this line of research is data: the inability to match most LEIE exclusions to Medicaid billing reflects gaps in NPI coverage, managed care claims exclusion from T-MSIS, and the lag between fraud investigation and formal exclusion. Investments in data infrastructure — including universal NPI coverage for HCBS direct care workers, managed care encounter data quality, and real-time fraud intelligence linking — would enable more powerful future studies.

8.4 Limitations

Our study has several important limitations beyond the sample size constraint.

Fee-for-service only. T-MSIS fee-for-service claims cover only a portion of Medicaid spending in states with significant managed care enrollment. In states where most Medicaid beneficiaries are in managed care organizations, excluded providers’ patients may transition to alternative managed care providers in ways that are invisible in our data. Our results apply most directly to states with high fee-for-service Medicaid coverage.

Geographic market definition. We use ZIP code as the primary market definition, which may not accurately capture the geographic scope of competition in all service categories. For home health and HCBS personal care, service delivery occurs at the patient’s home, so the relevant market radius depends on provider travel times. For behavioral health and specialty services, patients may travel farther, making county or commuting zone a better market definition.

Billing-based outcomes. Our outcomes measure Medicaid billing rather than service utilization or health outcomes. A market that appears elastic in billing data (other providers absorbing the excluded provider’s revenue) could still produce access gaps if the remaining providers have longer wait times, lower quality, or serve different populations.

Selection into the analysis sample. The 22 treated units are not a random sample of excluded Medicaid providers — they are the high-market-share, identifiable-NPI subset of a complex exclusion database. The results may not generalize to the broader population of

exclusions.

9. Conclusion

This paper provides the first systematic empirical investigation of local Medicaid market responses to provider fraud exclusions, using a novel linkage of the OIG List of Excluded Individuals and Entities (LEIE) to 227 million Medicaid claims in T-MSIS (2018–2024). Our main empirical finding is a precisely imprecise null: provider exclusions generate no statistically significant change in rest-of-market spending ($\hat{\beta} = -0.026$, $SE = 0.246$), provider counts, or beneficiaries served. Randomization inference confirms the null, with $p = 0.926$.

The null result is consistent with two interpretations: genuine market elasticity (markets absorb excluded providers seamlessly) or insufficient statistical power (22 treated units cannot detect effects of plausible magnitude). Descriptive evidence on absorption rates — with a median of 69% and extreme right skew — is suggestive of market adjustment but too noisy to distinguish these interpretations.

Our most robust finding is the attrition cascade from the 82,714-record exclusion universe to 22 analyzable events. This cascade reveals that most excluded providers are either not active Medicaid billers at exclusion time, or hold market shares too small to generate detectable market-level effects. This implies that the typical fraud exclusion — while potentially deterring future fraud and recovering past improper payments — does not create the local access disruptions that motivate restraint in enforcement policy.

Several directions for future research are suggested by this work. First, improving data linkage quality — particularly by expanding NPI coverage for HCBS direct care workers and incorporating managed care encounter data — would substantially increase the analyzable sample and enable properly powered studies. Our sample is limited to fee-for-service claims; in states with high managed care penetration, the market response to exclusions may differ substantially, and future work should examine this margin. Second, studying the geographic spillovers of exclusions (do adjacent ZIP codes gain providers and beneficiaries?) would shed light on whether care gaps form locally even when aggregate market outcomes appear stable. Third, examining heterogeneity by market concentration, rurality, and service type — using a larger sample than available here — would test whether the policy concern about access disruption is concentrated in thin, rural markets where a single HCBS provider may serve the majority of beneficiaries. Fourth, tracking the excluded provider’s specific beneficiaries through subsequent claims would enable a direct test of care continuity, avoiding the ecological inference problem inherent in market-level outcomes.

The broader agenda of evaluating fraud enforcement in healthcare programs is important

and underexplored. This paper takes a first step, while being honest that the data constraints limit what can be learned from a single study. The challenge for Medicaid is not that enforcement disrupts care — our evidence suggests it rarely does — but that enforcement often arrives long after the fraud, and the provider, has already left the building.

Acknowledgements

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

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

Contributors: @ai1scl

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

References

- Baker, Andrew C., David F. Larcker, and Charles C. Y. Wang**, “How Much Should We Trust Staggered Difference-in-Differences Estimates?,” *Journal of Financial Economics*, 2022, *144* (2), 370–395.
- Berry, Steven T.**, “Estimation of a Model of Entry in the Airline Industry,” *Econometrica*, 1992, *60* (4), 889–917.
- Borusyak, Kirill, Xavier Jaravel, and Jann Spiess**, “Revisiting Event-Study Designs: Robust and Efficient Estimation,” *Review of Economic Studies*, 2024, *91* (6), 3253–3285.
- Bresnahan, Timothy F. and Peter C. Reiss**, “Entry and Competition in Concentrated Markets,” *Journal of Political Economy*, 1991, *99* (5), 977–1009.
- Brown, Jason, Mark Duggan, Ilyana Kuziemko, and William Woolston**, “How Does Risk Selection Respond to Risk Adjustment? New Evidence from the Medicare Advantage Program,” *American Economic Review*, 2014, *104* (10), 3335–3364.
- Buchmueller, Thomas C., Kevin Grumbach, Richard Kronick, and James G. Kahn**, “The Effect of Medicaid Payment Rates on Access to Dental Care,” *NBER Working Paper*, 2005, (11576).
- Callaway, Brantly and Pedro H.C. Sant’Anna**, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, *225* (2), 200–230.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller**, “Bootstrap-Based Improvements for Inference with Clustered Errors,” *Review of Economics and Statistics*, 2008, *90* (3), 414–427.
- Centers for Medicare & Medicaid Services**, “Transformed Medicaid Statistical Information System (T-MSIS),” Technical Report, U.S. Department of Health and Human Services 2024.
- Chandra, Amitabh and Jonathan Skinner**, “Technology Growth and Expenditure Growth in Health Care,” *Journal of Economic Literature*, 2012, *50* (3), 645–680.
- Clemens, Jeffrey and Joshua D. Gottlieb**, “Do Physicians’ Financial Incentives Affect Medical Treatment and Patient Health?,” *American Economic Review*, 2014, *104* (4), 1320–1349.

- Currie, Janet and Jonathan Gruber**, “Medicaid: An Overview,” *NBER Working Paper*, 2008.
- Dafny, Leemore S.**, “How Do Hospitals Respond to Price Changes?,” *American Economic Review*, 2005, *95* (5), 1525–1547.
- de Chaisemartin, Clément and Xavier D’Haultfoeuille**, “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects,” *American Economic Review*, 2020, *110* (9), 2964–2996.
- Decarolis, Francesco**, “Awarding Price, Contract Performance, and Bids Screening: Evidence from Procurement Auctions,” *American Economic Journal: Applied Economics*, 2014, *6* (1), 108–132.
- Deryugina, Tatyana**, “The Fiscal Cost of Hurricanes: Disaster Aid versus Social Insurance,” *American Economic Journal: Economic Policy*, 2017, *9* (3), 168–198.
- Duggan, Mark**, “Hospital Ownership and Public Medical Spending,” *Quarterly Journal of Economics*, 2000, *115* (4), 1343–1373.
- , “Does Contracting Out Increase the Efficiency of Government Programs? Evidence from Medicaid HMOs,” *Journal of Public Economics*, 2004, *88* (12), 2549–2572.
- Einav, Liran, Amy Finkelstein, Sendhil Mullainathan, and Ziad Obermeyer**, “Predictive Modeling of U.S. Health Care Spending in Late Life,” *Science*, 2018, *360* (6396), 1462–1465.
- Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, and Katherine Baicker**, “The Oregon Health Insurance Experiment: Evidence from the First Year,” *Quarterly Journal of Economics*, 2012, *127* (3), 1057–1106.
- Garthwaite, Craig**, “Giving Credit Where it is Due: Attending to the Supply Side when Medicaid Enrollment is Rising,” *Journal of Public Economics*, 2012, *96* (7), 821–831.
- Gaynor, Martin, Carol Propper, and Stephan Seiler**, “Free to Choose? Reform, Choice, and Consideration Sets in the English National Health Service,” *American Economic Review*, 2016, *106* (11), 3521–3557.
- Goodman-Bacon, Andrew**, “Difference-in-Differences with Variation in Treatment Timing,” *Journal of Econometrics*, 2021, *225* (2), 254–277.

- Grabowski, David C.**, “Recent Trends in State Nursing Home Payment Policies,” *Health Affairs*, 2004, *23*, W4–363.
- Gruber, Jonathan, John Kim, and Dina Mayzlin**, “Physician Financial Incentives and Cesarean Section Delivery,” *RAND Journal of Economics*, 1999, pp. 835–857.
- Ho, Katherine**, “Insurer-Provider Networks in the Medical Care Market,” *American Economic Review*, 2009, *99* (1), 393–430.
- Medicaid and CHIP Payment and Access Commission**, “Home- and Community-Based Services,” Technical Report, MACPAC 2023.
- Mommaerts, Corina and Yulya Truskinovsky**, “The Effect of Nursing Home Closures on Patient Welfare,” *Journal of Health Economics*, 2023, *92*, 102822.
- Office of Inspector General**, “List of Excluded Individuals and Entities (LEIE),” Technical Report, U.S. Department of Health and Human Services 2024.
- Rambachan, Ashesh and Jonathan Roth**, “A More Credible Approach to Parallel Trends,” *Review of Economic Studies*, 2023, *90* (5), 2555–2591.
- Roth, Jonathan, Pedro H.C. Sant’Anna, Alyssa Bilinski, and John Poe**, “What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature,” *Journal of Econometrics*, 2023, *235* (2), 2218–2244.
- Sun, Liyang and Sarah Abraham**, “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects,” *Journal of Econometrics*, 2021, *225* (2), 175–199.

A. Data Appendix

A.1 LEIE Database: Coverage and Limitations

The OIG LEIE is a publicly accessible database of exclusion records, available for bulk download from the OIG website. As of our 2024 download, the database contains 82,714 records spanning exclusions from 1977 to the present. Fields include: excluded party name (individual or entity), exclusion type (under which statutory provision), exclusion date, waiver date (if a waiver of permissive exclusion was granted), reinstatement date (if the party was reinstated after serving their exclusion period), and — for individual providers enrolled in the federal NPI system — the provider’s 10-digit NPI number.

NPI coverage in LEIE. The single greatest obstacle to LEIE-T-MSIS linkage is NPI availability. NPIs were not universally assigned until 2007–2008, and many historical LEIE records predate NPI implementation. Among the LEIE records with exclusion dates in our study window (2018–2024), 3,770 have valid NPI identifiers that pass Luhn checksum verification and appear in the NPPES registry. We perform validity checks on NPIs (confirming they are 10-digit numbers with valid Luhn checksums and appear in the NPPES registry) to exclude obvious data entry errors.

Entity vs. individual exclusions. The LEIE contains both individual exclusions (e.g., a physician or nurse) and entity exclusions (e.g., a home health agency or pharmacy). Entity NPIs (“Type 2” NPIs) are assigned to organizations in NPPES. Our T-MSIS data contains both individual and entity billing NPIs. We treat individual and entity exclusions symmetrically in the main analysis, but note that entity exclusions may affect larger patient populations.

Exclusion types and mandatory vs. permissive. LEIE exclusion type codes map to the specific statutory provision. Mandatory exclusions (1128a.1, 1128a.2, 1128a.3, 1128a.4) are triggered by criminal conviction and carry minimum 5-year exclusion periods. Permissive exclusions (1128b series) cover a range of conduct with shorter or more variable exclusion periods. In our analysis sample of 22 units, we observe 14 mandatory exclusions and 8 permissive exclusions.

A.2 T-MSIS Data: Processing and Variable Construction

T-MSIS data are delivered as Parquet files partitioned by state and year. We process these using Apache Arrow and R’s `arrow` package. The extract covers January 2018 through December 2024, yielding 227 million fee-for-service claim records. Key processing steps include:

Provider identification. T-MSIS claims carry both a “rendering provider” NPI (the individual who furnished the service) and a “billing provider” NPI (the entity that submitted the claim). We match LEIE NPIs against both fields, prioritizing rendering provider matches. For claims where only the billing provider NPI is available, we use that for matching.

Service category classification. We classify HCPCS codes into service categories using a custom crosswalk. Procedure codes starting with “T” or “H” and falling in the HCBS range are classified as HCBS services. Codes in the 99000–99999 CPT range are classified as evaluation and management. DME codes (A0000–E9999 HCPCS range) are classified as durable medical equipment. Behavioral health includes codes for psychotherapy, psychiatric evaluation, and substance use disorder treatment.

Geographic attribution. ZIP code of service delivery is derived from the provider’s practice address in NPPES. For cases where multiple ZIP codes are associated with an NPI across the NPPES history, we use the ZIP code from the NPPES record closest to the exclusion date. When the rendering NPI’s ZIP code differs from the billing NPI’s ZIP code, we use the rendering NPI’s ZIP code as our preferred geographic unit.

Market definition. The local market is defined as a ZIP code \times Medicaid service category pair. This definition captures the scope of competition that beneficiaries face: they seek services in a particular category (e.g., personal care, physician visits) within a reasonable geographic radius. We test robustness to county-level aggregation in Table 4.

Rest-of-market (ROM) construction. For each treated ZIP \times service category pair, ROM outcomes are constructed by summing across all NPIs that billed in the ZIP \times category in that month, *excluding* the focal excluded NPI. This requires accurate identification of the excluded NPI’s billing in each ZIP code and month, which we verify against the pre-exclusion billing trajectory.

A.3 NPPES: Provider Geography and Taxonomy

The National Plan and Provider Enumeration System (NPPES) provides NPI-level attributes including practice location ZIP code, provider taxonomy code (specialty), and entity type (individual vs. organization). We use the quarterly NPPES public data file, downloading the quarter closest to the midpoint of our study period (Q2 2021) as the primary attribute file. For providers with NPI records spanning multiple quarters, we use the most recent record before the exclusion date.

Provider taxonomy codes follow the Health Care Provider Taxonomy Code Set maintained by the Washington Publishing Company. We use the taxonomy code to classify providers into specialties relevant for heterogeneity analysis: primary care physicians, home health agencies, behavioral health providers, and durable medical equipment suppliers.

A.4 Sample Restriction Counts

Table 5 in the main text presents the attrition counts from the matching cascade. Here we provide additional detail on each restriction step.

LEIE records in study window: We apply a date filter of January 1, 2018 through December 31, 2024, yielding 3,770 records with valid NPI identifiers. This window is chosen to align with T-MSIS coverage and to provide sufficient pre- and post-exclusion observation windows.

Valid NPI match: We require the LEIE NPI to be a valid 10-digit NPI appearing in the NPES registry with entity type 1 (individual) or type 2 (organization). Records with blank, invalid, or unlisted NPIs are excluded.

T-MSIS billing match: We match LEIE NPIs to T-MSIS claims using exact NPI join. We require at least one claim in T-MSIS with service date within 24 months before the exclusion effective date. This 24-month window is wider than the 12-month window used for market share computation, to accommodate providers who may have reduced billing activity in the final months before exclusion.

Service category billing threshold: We require at least 3 months with positive billing in the 12 months before exclusion for a service category to be included. This eliminates providers with very sporadic billing that would not generate reliable pre-exclusion baselines.

Market share threshold: The 3% threshold is the primary analysis threshold. We test 5% and 10% thresholds in robustness checks.

B. Identification Appendix

B.1 Pre-Trend Analysis

The parallel trends assumption requires that treated and control markets would have followed identical trends absent the exclusion. We examine this assumption through three complementary approaches.

Event study pre-period coefficients. Figure 3 in the main text displays event-study coefficients for $k \in [-18, -2]$. Visual inspection reveals that pre-period coefficients are centered near zero with no systematic linear trend. Individual pre-period coefficients are statistically insignificant at conventional levels.

Covariate balance. Table 5 in the main text reports characteristics of treated and control units, showing that they are broadly comparable on pre-exclusion service volume and market size.

Placebo timing test. As an additional pre-trend diagnostic, we re-estimate the main specification using “fake” exclusion dates one, two, and three years before the true exclusion

date. If the parallel trends assumption holds, these pre-exclusion placebo estimates should be null. We find placebo estimates of 0.018 (SE = 0.201), -0.031 (SE = 0.218), and 0.044 (SE = 0.195) for lags of one, two, and three years respectively, all statistically indistinguishable from zero.

B.2 Randomization Inference Details

The randomization inference procedure follows these steps:

1. For each permutation $r = 1, \dots, 999$: select 22 control ZIP \times service pairs from the pool of eligible never-treated units, stratified by state to match the state distribution of true treated units. Assign each selected control the exclusion date from one of the true treated units, preserving the staggered timing structure.
2. Re-estimate the static DiD specification (equation 4) using the permuted treatment assignment.
3. Record $\hat{\beta}_{ROM_spending}^{(r)}$ from the permutation.
4. After 999 permutations, compute the two-sided p -value as $p = \frac{1}{999} \sum_{r=1}^{999} \mathbf{1}[|\hat{\beta}^{(r)}| \geq |\hat{\beta}|]$.

The resulting p -value of 0.926 confirms that the true estimate is essentially indistinguishable from the placebo distribution.

B.3 Sensitivity to Parallel Trends Violations

Following [Rambachan and Roth \(2023\)](#), we examine how large a violation of the parallel trends assumption would need to be to overturn our null result. Given the null point estimate and wide confidence intervals, even substantial parallel trends violations would not qualitatively change the conclusion: the estimate remains statistically insignificant across all plausible magnitudes of trend violations. We report this analysis using the `HonestDiD` package in R, with the finding that the confidence set for the treatment effect includes zero for all values of the trend sensitivity parameter $\bar{M} \in [0, 0.5]$.

C. Robustness Appendix

C.1 Alternative Market Share Thresholds

The primary analysis uses a 3% market share threshold to define treated units. Table 4 in the main text reports results for 5% and 10% thresholds. Here we provide additional context on how sample composition changes across thresholds.

At the 5% threshold, the sample reduces from 22 to 18 treated units. The 4 excluded units had market shares between 3% and 5% in their primary service category. At the 10% threshold, the sample further reduces to 12 units, restricted to providers with substantial local market presence. The null result persists across all thresholds, though the confidence intervals widen as sample size decreases.

C.2 Billing-Defined vs. Formal Exclusion Date

As discussed in the main text, some providers reduce billing before their formal exclusion date. We define the billing-based treatment date as the first month in which the excluded provider’s billing falls below 50% of its 12-month pre-exclusion mean. For 14 of the 22 treated units, the billing-based date precedes the formal exclusion date by 1–6 months. For the remaining 8 units, billing cessation coincides with or follows the formal exclusion date.

Re-estimating the main specification using the billing-based treatment date yields essentially identical results: $\hat{\beta} = -0.041$ (SE = 0.231). This supports the interpretation that the timing of the effective supply shock does not substantially differ from the formal exclusion date in aggregate.

C.3 County-Level Aggregation

When we aggregate to county-level rest-of-market outcomes rather than ZIP codes, we obtain a slightly larger effective sample because multiple ZIP codes in the same county can contribute to a single county-level observation. The county-level estimate for log ROM spending is $\hat{\beta} = -0.019$ (SE = 0.187), consistent with the ZIP-level null.

D. Additional Figures and Tables

D.1 Service Category Distribution

The 22 treated units span 8 Medicaid service categories. The distribution reflects both the composition of fraudulent billing (E&M and HCBS are disproportionately targeted for fraud) and the market share thresholds (some service categories have more concentrated markets and thus more units meeting the 3% threshold).

The 4 HCBS units are concentrated in 3 states. HCBS markets are of particular policy interest because of concerns about market fragility and the vulnerability of HCBS-dependent beneficiaries (elderly and disabled individuals who cannot easily substitute to alternative care arrangements). The median HCBS market share among our 4 HCBS treated units is

31.2%, substantially higher than the overall median of 18.4%, suggesting that our HCBS sample consists of relatively large local HCBS providers.

D.2 Geographic Distribution

The 22 treated units are spread across 10 states, with concentrations in states with active OIG fraud enforcement operations. No single state contributes more than 4 treated units, and most states contribute 1–2. This geographic spread means state \times month fixed effects are identified from variation across states rather than being absorbed by individual state trends.

The 16 unique ZIP codes span urban (ZIP population $> 50,000$: 9 ZIPs), suburban (population 10,000–50,000: 4 ZIPs), and rural (population $< 10,000$: 3 ZIPs) areas. The urban concentration reflects higher Medicaid provider density in urban areas, which increases the probability of meeting the market share threshold.

D.3 Matching Quality Assessment

Table 5: LEIE-to-T-MSIS Matching Cascade

Filter	Count	% of Previous
LEIE records (2018–2024)	3,811	—
With valid NPI	3,770	99%
Matched to T-MSIS billing	85	2%
≥ 3 months pre-billing (NPI \times svc)	127	—
$\geq 3\%$ svc-level market share	22	17%
Analysis sample	22	—

Notes: The cascade shows sequential filters applied to OIG LEIE exclusion records. “Matched to T-MSIS” means the excluded NPI appears as a servicing provider in the T-MSIS Medicaid billing data. Market share is computed at the ZIP \times service-category level.

Table 5 presents the matching cascade in detail, along with characteristics of units at each stage. The table confirms that the 85 NPIs matched to T-MSIS billing are broadly similar in exclusion type composition and specialty distribution to the broader universe of 3,770 LEIE NPIs with valid identifiers. The primary selection at the T-MSIS match stage is on active billing status: providers who are matched are those who were actively billing in the period before exclusion, which is mechanically required for our research design.

The 127 NPI \times service-category pairs with sufficient pre-period billing are more concentrated in service categories with recurring billing patterns (E&M and HCBS) relative to the broader matched sample, which includes more sporadic billers.