State-Elected Procedural-Disenrollment Delays and Medicaid Renewal Outcomes During the 2023 Unwinding

Abstract

Background. During the 2023–2024 Medicaid continuous coverage unwinding, approximately 69 percent of coverage terminations were for procedural reasons rather than ineligibility determinations. In August 2023, the Centers for Medicare and Medicaid Services (CMS) identified that some states were conducting automatic renewals at the household level rather than at the federally required individual level, and by September 2023 had publicly classified 29 states as noncompliant or still assessing. Separately, CMS made available a state option to delay procedural disenrollments while a state conducted targeted renewal outreach; the agency publicly documented 15 states that obtained concurrence to use that option. The pause/delay chart explicitly does not identify states approved to delay disenrollments as a mitigation strategy for noncompliance. No prior study has evaluated this state-elected delay option using state-month difference-in-differences methods.

Methods. I constructed a balanced state-month panel (51 jurisdictions, 18 months, February 2023–July 2024) using publicly available CMS administrative data. I estimated the association between active CMS-concurred procedural-disenrollment delays and five renewal outcomes using two-way fixed effects (TWFE) difference-in-differences and a Gardner-style two-stage estimator (DID2S), exploiting staggered pause timing across 15 states relative to 36 never-paused states. I assessed robustness through pre-trend tests, placebo simulations, a heuristic parallel-trends sensitivity bound, Oster bounds, multiple-testing adjustment, an approximate minimum detectable effect calculation, early-pause exclusions, a trimmed event window, an intent-to-treat (ITT) decomposition, pause spell and intensity analyses, and a South Carolina treatment-status-uncertain sensitivity.

Results. Active CMS-concurred pauses were associated with an 8.1-percentage-point increase in renewal completion rates (TWFE; SE = 3.8 pp; $p = 0.039$), with a larger and more precise two-stage estimate (DID2S: +13.6 pp; $p < 0.001$). The two-stage renewal-completion estimate survived Benjamini-Hochberg adjustment across 15 main tests (BH-adjusted $p = 0.002$), as did the DID2S pending-rate reduction (-11.0 pp; BH-adjusted $p = 0.011$). Point estimates for procedural termination rates (-1.4 pp; $p = 0.62$) and total disenrollment rates (-3.9 pp; $p = 0.25$) were directionally favorable but imprecise. The ITT specification, comparing the 29 noncompliance-flagged states to the 22 compliant states, was effectively null for every outcome. A pre-trend test still rejected for renewal completion in the baseline event window ($p = 0.012$); trimming the window from six to four leads softened but did not eliminate the rejection. The

TWFE observed effect was at the design's approximate minimum detectable effect (8.05 pp), while the DID2S estimate exceeded it by 70 percent. Dropping South Carolina (treatment-status uncertain) left the headline intact (TWFE +7.8 pp, $p = 0.052$; DID2S +13.3 pp, $p = 0.0005$).

Conclusions. States that opted into the CMS-concurred procedural-disenrollment delay completed more renewals while their pauses were active. The two-stage estimate is the strongest evidence and survives multiple-testing adjustment; the TWFE estimate is more conservative and underpowered. The CMS individual-level compliance designation by itself was not associated with detectable average outcome changes. Because the pause action is a state-elected operational tool rather than a federally imposed enforcement instrument, the estimates speak most directly to the value of additional processing time during a coverage transition. Robust pre-trend rejection, small treated-state count, and limited TWFE power continue to discipline how strongly these estimates can be interpreted causally.

1. Introduction

The resumption of Medicaid eligibility redeterminations in 2023—following three years of federally mandated continuous coverage during the COVID-19 public health emergency—produced one of the largest coverage disruptions in the program's history. Between April 2023 and mid-2024, over 25 million individuals were disenrolled from Medicaid, with approximately 69 percent of terminations occurring for procedural reasons: beneficiaries lost coverage not because they were determined ineligible but because they failed to complete required renewal paperwork (KFF, 2025; MACPAC, 2024). This mass procedural disenrollment highlighted the central role of administrative burden in Medicaid coverage retention (Moynihan, Herd, and Harvey, 2015; Herd and Moynihan, 2018) and raised urgent questions about how state administrative tools and federal oversight interact during periods of large-scale operational disruption.

In August 2023, CMS identified a specific compliance problem with potentially far-reaching consequences: some states were conducting automatic (ex parte) renewals at the household level rather than at the federally required individual level. Under household-level processing, if one household member was found ineligible or failed to respond to a renewal notice, other individually eligible household members could be procedurally terminated. CMS required all states to assess their compliance and, on September 21, 2023, published a state assessment table classifying 29 jurisdictions as noncompliant or still assessing and 22 as attesting to individual-level compliance. By early 2024, federal and state officials reported that more than 400,000 beneficiaries had been reinstated after inappropriate terminations tied to the individual-level renewal issue (GAO, 2024).

Separately, and applicable to a broader set of states beyond the individual-level compliance question, CMS made available a state option to delay procedural disenrollments while a state conducted targeted renewal outreach. The agency published a chart of the 15 states that requested and obtained concurrence to use that targeted-outreach delay option, with pause start dates ranging from May 2023 (Delaware) through October 2023 (Texas). The CMS chart explicitly notes that it identifies states using the targeted-outreach delay option and does not identify states approved to delay disenrollments as a mitigation strategy for noncompliance with renewal requirements. Reinstatements after the individual-level renewal issue therefore involved many state mitigation strategies, of which the targeted-outreach delay was only one.

This episode sits at the intersection of two literatures that have rarely been joined empirically. The administrative burden literature shows that renewal requirements and documentation processes meaningfully reduce program participation, especially among populations with limited time, information, or administrative support (Moynihan, Herd, and Harvey, 2015; Herd and Moynihan, 2018; Myerson, Espeseth, and Dague, 2025; Arbogast, Chorniy, and Currie, 2024). The Medicaid federalism literature, by contrast, has mostly described federal oversight qualitatively, emphasizing cooperative engagement and the political limits of aggressive enforcement in a joint federal-state program (Gusmano and Thompson, 2025). What has been missing is a quantitative assessment of how a state-elected operational tool that directly slowed the procedural disenrollment pipeline was associated with renewal outcomes during a major coverage transition.

Despite a rapidly growing literature on the Medicaid unwinding, no existing study has evaluated the targeted-outreach delay option using state-month difference-in-differences methods. Prior causal work has focused on the unwinding itself, estimating the effect of resuming redeterminations on coverage and economic well-being (Dasgupta and Solomon, 2025), or on collaborative federal interventions such as the CMS/USDS technical assistance effort to increase automated renewals in four states (Moynihan, Herd, and Ribeiro, 2025). My contribution is therefore narrower than a clean causal claim. I document effect sizes consistent with improved renewal processing in states that opted into the delay option, I show how those estimates vary when the intervention is treated as a temporary state-elected pause rather than a permanent treatment, and I clarify why the evidence remains sensitive to identification and power concerns.

I find that active CMS-concurred procedural-disenrollment delays were associated with an 8.1-percentage-point increase in renewal completion rates in the as-treated TWFE specification, and a larger and more precise 13.6-percentage-point increase in the Gardner-style two-stage estimator. The two-stage estimate survives Benjamini-Hochberg multiple-testing adjustment across 15 main treatment-effect tests, while the TWFE estimate remains nominally marginal at the design's minimum detectable effect. The broader CMS individual-level compliance designation by itself shows no detectable average effect on outcomes. The findings are best interpreted as suggestive evidence on a state-elected operational margin during a high-salience coverage transition: the action that appears to matter is the actual implementation of a pause, not the broader compliance label. Robust pre-trend rejection, the small number of treated jurisdictions, and limited TWFE power continue to constrain causal claims.

2. Background

2.1 The Medicaid Unwinding

The Families First Coronavirus Response Act (FFCRA) of 2020 conditioned enhanced federal Medicaid matching funds on a continuous coverage provision (CCP) that prohibited states from disenrolling beneficiaries during the public health emergency. This provision increased total Medicaid enrollment from 71 million in March 2020 to over 91 million by September 2022 (Dague and Ukert, 2023). When the Consolidated Appropriations Act (CAA) of 2023 decoupled the CCP from the PHE declaration, states began resuming eligibility redeterminations starting in April 2023.

The scale of disenrollment during the unwinding was unprecedented. Over 25 million individuals were disenrolled between April 2023 and mid-2024, while approximately 56 million had coverage renewed (KFF, 2025). Approximately 69 percent of all terminations were for procedural reasons (MACPAC, 2024). Procedural termination rates varied enormously across states, from 22 percent in Maine to 93 percent in Nevada (KFF, 2025). Children were disproportionately affected: 4.16 million fewer children were enrolled by December 2023, a 10 percent national decline (Georgetown CCF, 2024).

2.2 Administrative Burden and Medicaid Renewal

The administrative burden framework conceptualizes the costs citizens face in interacting with government programs along three dimensions: learning costs, compliance costs, and psychological costs (Moynihan, Herd, and Harvey, 2015). Herd and Moynihan (2018) demonstrated that burden levels are a product of deliberate political choice rather than administrative accident. The Medicaid renewal process imposes substantial compliance costs that fall disproportionately on populations with fewer resources to manage them.

Recent empirical work confirms the magnitude of these effects. Arbogast, Chorniy, and Currie (2024) found that state regulations increasing administrative burdens reduced child Medicaid enrollment by 5.9 percent within six months. Myerson, Espeseth, and Dague (2025) demonstrated through a large-scale randomized experiment that personalized navigator assistance increased renewal by 1 percentage point overall, with substantially larger effects for tribal members (8 pp) and children (3 pp). Moynihan, Herd, and Ribeiro (2025) evaluated a federal technical assistance intervention to increase automated renewals in four states, finding that the intervention increased ex parte renewals by 21.6 percentage points and decreased procedural denials by 8.3 points.

2.3 The Individual-Level Renewal Issue and the State Targeted-Outreach Delay Option

On August 30, 2023, CMS issued a State Medicaid Director Letter alerting states to a potential eligibility system issue: some states were conducting ex parte renewal processes at the household level rather than the federally required individual level (CMS, 2023a). Under household-level processing, if one household member was found ineligible or failed to respond, other household members who might remain individually eligible could be procedurally terminated. By September 21, 2023, CMS published a state assessment table classifying each state's compliance status: 22 jurisdictions attested to conducting individual-level renewals, while 29 jurisdictions were identified as noncompliant or still assessing. New Hampshire is classified as compliant throughout this paper, consistent with the September 2023 CMS PDF, which lists New Hampshire as attesting "Yes" on individual-level auto-renewals.

Separately, CMS publicly maintained a chart of states that obtained CMS concurrence to delay procedural disenrollments while conducting targeted renewal outreach. Fifteen states used that targeted-outreach delay option, with pause start dates ranging from May 2023 (Delaware) through October 2023 (Texas) and stated durations from one to several months. The chart records, for each state, the duration of the delay, the months in which the delay was active, the scope (all beneficiaries vs. specific populations), and the affected populations. Kentucky has both a primary all-beneficiary one-month delay starting in September 2023 and a secondary two-month non-MAGI long-term-care delay covering renewals due May through August 2023 (per the CMS chart's footnote); both windows are coded as active in the analysis. South Carolina's pause is coded as continuous from June 2023 through the end of the observed period; a sensitivity analysis treats South Carolina as treatment-status-uncertain and drops it from the sample.

The CMS chart explicitly notes that it identifies states using the state-elected targeted-outreach delay option and does not identify states approved to delay disenrollments as a mitigation strategy for noncompliance with the individual-level renewal requirements. The earlier reinstatement of more than 400,000 beneficiaries tied to the individual-level renewal issue therefore involved many state mitigation strategies, of which the targeted-outreach delay was only one (GAO, 2024).

2.4 Contribution of This Paper

This paper fills three gaps in the literature. First, it provides a quantitative state-month assessment of a specific state-elected operational tool used during the Medicaid unwinding, complementing the experimental evidence on navigator assistance (Myerson, Espeseth, and Dague, 2025) and the technical-assistance evidence on automated renewals (Moynihan, Herd, and Ribeiro, 2025). Second, it separates the broader CMS individual-level compliance designation from the operational margin most closely tied to renewal processing, and shows that the broader designation by itself had no detectable average effect. Third, it reframes the intervention as a temporary state-elected pause with heterogeneous duration and scope, which makes it possible to study onset, active-pause, and post-pause patterns rather than relying solely on a single binary treatment indicator.

3. Data

3.1 Data Sources

I construct a balanced state-month panel of Medicaid unwinding outcomes and federal compliance assessment status for 51 jurisdictions (50 states and the District of Columbia) spanning February 2023 through July 2024. My analysis relies on five publicly available data sources.

CMS Unwinding Renewal Metrics. My primary outcome data come from the Medicaid and CHIP Unwinding Monthly Reports, publicly available through the data.medicaid.gov portal (Dataset ID: 5abea2e0-3f8e-4b49-a50d-d63d5fd9103c). Under Section 1902(tt)(1) of the Social Security Act, added by the Consolidated Appropriations Act of 2023, states were required to submit monthly data to CMS for each month from April 2023 through June 2024 on renewals initiated and completed, disenrollment outcomes, and other operational measures. The dataset reports, for each state-month, the number of beneficiaries with a renewal due; total renewals completed (decomposed into ex parte and form-based); total disenrollments at renewal (decomposed into beneficiaries determined ineligible and those terminated for procedural reasons); and renewals still pending at month's end. I use the "Updated" data version when available, which reflects the disposition of previously pending renewals as of three months after the original reporting period.

CMS Eligibility and Enrollment Performance Indicators. I obtain monthly state-level Medicaid and CHIP enrollment totals and operational metrics from the Performance Indicator (PI) data collection (Dataset ID: 6165f45b-ca93-5bb5-9d06-db29c692a360). This data series provides point-in-time enrollment counts, call center operations, and application processing metrics.

CMS State Compliance Assessment. My ITT treatment variable derives from the "Preliminary Overview of State Assessments Regarding Compliance with Medicaid and CHIP Automatic Renewal Requirements at the Individual Level," published by CMS on September 21, 2023 (CMS, 2023b). The table classifies each state as either (a) attesting to correctly conducting auto-renewals at the individual level (22 jurisdictions) or (b) not conducting individual-level auto-renewals or still assessing (29 jurisdictions). For states in the latter group, the table further documents affected populations and estimated numbers of affected individuals.

CMS State Option to Delay Procedural Disenrollments. My primary as-treated treatment variable comes from CMS's publicly available chart documenting 15 states that obtained CMS concurrence to delay procedural disenrollments under the state targeted-outreach option (chart as of May 22, 2024). This chart records, for each state, the duration of the delay, the months in which the delay was active, the scope, and the affected populations. The chart explicitly does not identify states approved to delay disenrollments as a mitigation strategy for noncompliance.

American Community Survey. I supplement the CMS administrative data with state-level demographic controls from the 2022 American Community Survey 1-Year Estimates, accessed via the Census Bureau API when reachable at build time. I include total population, percent Black, percent Hispanic, percent below the federal poverty line, and median household income as time-invariant state-level covariates. When the Census API is not reachable at build time (e.g., when no API key is available in the build environment), the corresponding demographic rows in Table 1 are suppressed; the rest of the analysis is unaffected because state fixed effects absorb time-invariant state characteristics.

3.2 Treatment Definition

I define treatment using two complementary approaches. My intent-to-treat (ITT) specification codes a binary indicator equal to one for the 29 jurisdictions classified as noncompliant or still assessing in the September 21, 2023 CMS assessment, interacted with a post-assessment indicator (months from September 2023 onward). The 22 jurisdictions attesting to compliance serve as the comparison group. This ITT approach captures the effect of being flagged by CMS for individual-level renewal-compliance concerns, regardless of whether the state subsequently opted into the targeted-outreach delay option.

My as-treated specification exploits the staggered timing of the state-elected delay option. Fifteen states obtained CMS concurrence to delay procedural disenrollments under that option, with start dates ranging from May 2023 (Delaware) through October 2023 (Texas). The pause indicator turns on in the first month of the CMS-concurred delay and remains on through the end of the documented delay window. For Kentucky, both the primary all-beneficiary one-month delay starting September 2023 and the secondary May–August 2023 non-MAGI long-term-care delay are coded as active. To ensure ongoing pauses are not silently dropped, the unwinding window is extended through July 2024 so that pause-active observations are coded consistently through the end of the panel. North Carolina is coded as a non-expansion state as of the April 2023 unwinding baseline (NC's expansion took effect December 1, 2023); the indicator is absorbed by the state fixed effect but is documented for transparency.

Because several pauses began before the September 2023 assessment snapshot, and because pause duration and scope vary, I prespecified sensitivities that exclude early-pause states and a decomposition that separates pause onset, active-pause, and post-pause periods.

3.3 Sample Construction

My state-month skeleton consists of 918 observations (51 jurisdictions x 18 months). Outcome variables are structurally missing for approximately 11 percent of observations, concentrated in the earliest months (February–April 2023) when many states had not yet initiated unwinding-related renewals. I treat these as structurally missing rather than imputing values. The main regression sample includes 764 state-month observations with nonmissing renewal outcomes.

3.4 Key Outcome Variables

My primary outcomes are:

1. **Procedural termination rate**: the share of total disenrollments at renewal due to procedural rather than eligibility reasons, measuring the intensity of administrative burden in the renewal process.
2. **Total disenrollment rate**: total disenrollments divided by renewals due, capturing the overall rate of coverage loss at renewal.
3. **Renewal completion rate**: renewals completed divided by renewals due, measuring the share of beneficiaries who successfully retained coverage through the renewal process.
4. **Ex parte renewal rate**: the share of completed renewals conducted on an ex parte (automated) basis.
5. **Pending renewal rate**: renewals still pending at month's end as a share of renewals due, capturing administrative backlogs.

3.5 Descriptive Statistics

Table 1 presents pre-September-2023 summary statistics for key variables by ITT treatment group. Prior to the September 2023 compliance assessment, noncompliance-flagged states had modestly lower procedural termination rates (64.7% vs. 74.7%) and lower disenrollment rates (34.5% vs. 40.4%) than compliant states. Noncompliance-flagged states also tended to be smaller (mean enrollment 1.47 million vs. 2.31 million). When reachable, ACS demographic moments confirm meaningful baseline differences in poverty rates and median household income; when the Census API is not reachable, those rows are suppressed. These level differences motivate state fixed effects, which absorb all time-invariant state characteristics. The relevant identification assumption is that outcome *trends* would have been parallel absent treatment, which I evaluate through pre-trend tests.

4. Methods

4.1 Study Design

I use a difference-in-differences (DiD) design to estimate the association between active CMS-concurred procedural-disenrollment delays and Medicaid renewal outcomes during the continuous coverage unwinding period (April 2023–July

2024). The variation of interest comes from the staggered timing of pauses across 15 states relative to 36 never-paused states, together with the September 21, 2023 CMS compliance assessment that identified 29 jurisdictions as non-compliant or still assessing on individual-level auto-renewals. I treat this as a policy-relevant quasi-experimental setting rather than a clean natural experiment.

4.2 Estimating Equations

Two-way fixed effects (TWFE). My baseline specification is

$$Y_{st} = \alpha_s + \delta_t + \beta \cdot \text{Treatment}_{st} + \varepsilon_{st}$$

where Y_{st} is the outcome for state s in month t, α_s are state fixed effects absorbing time-invariant state characteristics, δ_t are month fixed effects absorbing common temporal shocks, and Treatment_{st} is either the ITT indicator ($\text{NoncomplianceFlag}_s \times \text{Post}_t$) or the as-treated pause-active indicator. Standard errors are clustered at the state level.

DID2S (Gardner 2022). To address potential bias from treatment-effect heterogeneity in staggered DiD settings, I implement a manual two-stage estimator in the spirit of Gardner (2022). In the first stage, I estimate the state and month fixed effects using only untreated observations (state-months where no pause is active). In the second stage, I regress the residualized outcome on the pause indicator using all observations with state-clustered standard errors. I label this estimator a heuristic two-stage implementation rather than a definitive DID2S package implementation.

Event study. I estimate

$$Y_{st} = \alpha_s + \delta_t + \sum_{k \neq -1} \beta_k \cdot \mathbf{1}[\text{eventtime}_{st} = k, \text{paused}_{st}] + \varepsilon_{st}$$

as a pause-vs-never-paused dynamic specification: the event-time indicators are activated only for pause-state observations in pre-onset leads or in active-treatment lags. Never-paused states and pause-state observations in post-onset off-treatment months (for example, short-pause states after their delay window closes) are assigned the reference period so they identify the state and month fixed effects without entering the active event-time dummies. The baseline event window covers six leads and 12 lags relative to pause onset, with the month before onset as the reference period. A trimmed window restricts to four leads and 12 lags.

4.3 Robustness and Sensitivity Analyses

I conduct a broad set of robustness and sensitivity checks. These include: (1) formal pre-trend tests using joint Wald statistics on pre-treatment event study coefficients; (2) placebo simulations with 500 iterations randomly assigning pseudo-pause dates to control states; (3) an approximate parallel-trends sensitivity bound calibrated to the largest observed pre-period coefficient change (not a formal Rambachan-Roth/HonestDiD implementation); (4) Oster (2019) bounds for sensitivity to selection on unobservables; (5) a balanced-panel restriction (months with at least 90 percent state reporting coverage); (6) alternative treatment definitions, including an absorbing-treatment sensitivity and the ITT specification with demographic controls; (7) heterogeneity analysis exploiting whether children were flagged as affected; (8) a mechanisms test examining whether procedural terminations and pending renewals move in predicted directions; (9) comparison between the full and restricted panels to assess sensitivity to structural missingness; (10) a sensitivity excluding the 10 pause states whose pauses began before September 2023; (11) multiple-testing adjustments across the 15 main treatment-effect estimates; (12) an approximate minimum detectable effect (MDE) calculation for the as-treated design; (13) trimmed event-window event studies using k in [-4, +12]; (14) a decomposition of the null ITT result based on overlap between ITT-flagged and pause states; (15) pause spell and intensity specifications that separate onset, during-pause, and post-pause periods and allow treatment effects to vary with duration and scope; (16) a treatment-status-uncertain sensitivity that drops South Carolina from the sample; and (17) a Goodman-Bacon weight decomposition for the as-treated TWFE clarifying the share of identifying variation coming from clean (treated-vs-never-paused) versus already-treated comparisons. Because the intervention is temporary rather than absorbing, estimators designed for permanent treatment adoption such as Sun-Abraham (2021) or Callaway-Sant'Anna (2021) are not a clean substitute for the main specification without stronger assumptions about post-pause exposure; Rak, Hatfield, and Fry's *Health Services Research* methods guidance on aggregation choices for staggered DiD (Rak, Hatfield, and Fry, 2025) informs the decision to present TWFE alongside the two-stage estimator rather than as a single-headline replacement.

4.4 Software

All analyses are implemented in Python 3 using pyfixest for fixed-effects estimation, statsmodels for the manual two-stage estimator with cluster-robust variance, and matplotlib/seaborn for visualization. A specification registry (analysis/specification-registry.yml) locks the estimand, sample, treatment definition, outcomes, fixed effects, clustering, and planned robustness battery.

5. Results

5.1 Main Results

Table 2 presents the main difference-in-differences results across all five outcomes, comparing ITT TWFE, as-treated TWFE, and the heuristic two-stage estimator (DID2S).

Renewal completion rate. The as-treated TWFE specification yields a coefficient of +8.1 percentage points (SE = 3.8 pp; p = 0.039), indicating that states with active CMS-concurred pauses experienced higher renewal completion rates than non-pause states (Table 2). The heuristic two-stage estimator produces a larger and more precise estimate of +13.6 percentage points (SE = 3.5 pp; p < 0.001). By contrast, the ITT specification, which compares all 29 noncompliance-flagged jurisdictions to the 22 compliant jurisdictions regardless of whether they actually paused disenrollments, yields a slightly negative null (-2.8 pp; p = 0.42). A simple decomposition helps explain this disconnect: only 9 of the 15 pause states were in the ITT-treated group, while 6 pause states were classified as compliant, and the 20 ITT-treated states without pauses showed virtually no post-assessment change in renewal completion.

Procedural termination rate. Active pauses were associated with a 1.4-percentage-point reduction in the procedural termination rate (SE = 2.8 pp; p = 0.62) in the as-treated TWFE specification. The direction is favorable but the estimate is imprecise. The DID2S estimate has the opposite sign for this outcome (+8.2 pp; p = 0.13), reflecting how heavily the two-stage estimator reweights specific cohorts; on balance, the procedural-termination result is best read as "directionally consistent in TWFE but unstable across estimators." The ITT estimate is positive (+4.0 pp; p = 0.32).

Total disenrollment rate. The as-treated TWFE estimate of -3.9 percentage points (SE = 3.4 pp; p = 0.25) is directionally consistent with reduced coverage loss at renewal, though the estimate remains below conventional significance thresholds. The DID2S estimate is -2.7 percentage points (p = 0.31).

Ex parte renewal rate. I find no detectable effect on ex parte renewal rates in the as-treated TWFE specification (+1.2 pp; SE = 3.3 pp; p = 0.71). This is consistent with the nature of the intervention: the pause targeted the procedural-disenrollment pipeline, not the ex parte renewal process itself. The DID2S estimate is +7.9 pp (p = 0.08), marginal but not surviving BH adjustment.

Pending renewal rate. The as-treated TWFE point estimate is -4.2 percentage points (SE = 3.2 pp; p = 0.19). The DID2S estimate is substantially larger (-11.0 pp; p = 0.002), consistent with active pauses converting pending cases into completed renewals rather than allowing them to accumulate.

5.2 Event Study Results

Figure 2 presents event study plots for each outcome, with dynamic treatment effects plotted relative to the month before pause onset (k = -1). For the renewal completion rate, pre-treatment coefficients are generally close to zero for event times -4 through -1, while the k = -5 coefficient is elevated in magnitude (approximately -10 percentage points; 95% CI, -18 to -2) and is the primary driver of the formal pre-trend rejection (Wald chi-squared(4) = 12.89, p = 0.012). The total disenrollment rate also fails the baseline pre-trend test (p = 0.027), while the procedural termination rate is marginal (p = 0.063), and the ex parte rate ($p = 0.984$) and pending rate ($p = 0.240$) show no pre-trend concerns. Figure 3 presents a multi-panel summary of all five event studies.

Trimming the event window from k in [-6, +12] to k in [-4, +12] removes the $k = -5$ lead and attenuates the formal pre-trend rejection for renewal completion ($p = 0.246$) and total disenrollment ($p = 0.496$). Under this narrower window, the period-zero coefficients remain modest but the pre-trend rejection no longer binds. The event-study evidence therefore reduces concern about one influential early lead but does not fully resolve the broader identification issue, and I continue to treat the headline as suggestive rather than definitive.

5.3 Robustness Checks

Pre-trend, placebo, and sensitivity diagnostics. Formal Wald tests fail to reject parallel pre-trends for three of five outcomes in the baseline window: procedural termination rate ($p = 0.063$), ex parte renewal rate ($p = 0.984$), and pending renewal rate ($p = 0.240$). Pre-trend violations are detected for the total disenrollment rate ($p = 0.027$) and renewal completion rate ($p = 0.012$), driven primarily by the $k = -5$ event-time coefficient (Table 3). Placebo tests are supportive for renewal completion: the as-treated estimate falls in the upper tail of the placebo distribution.

Approximate parallel-trends and Oster diagnostics. Confidence intervals widen substantially under even modest assumed violations of parallel trends. At $M = 0$ (exact parallel trends imposed), the renewal completion rate confidence interval is wide but excludes neither zero nor the point estimate; at $M = 1$ times the calibrated bound, no outcome remains excludable from zero. Oster (2019) delta-star values are small for several outcomes, including 0.033 for the renewal completion rate, indicating that modest unobserved confounding could rationalize the estimate.

Goodman-Bacon decomposition. Because the as-treated TWFE specification involves staggered treatment timing across 15 pause states, the estimator can in principle pool together comparisons in which already-treated states serve as controls for later-treated states. The on/off structure of the pause treatment makes this concern less severe than in absorbing-treatment designs, but the decomposition is reported in the appendix to quantify the share of identifying weight coming from clean (treated-vs-never-paused) versus already-treated 2x2 comparisons for the renewal-completion outcome.

Additional design checks. Restricting to a balanced panel with at least 90 percent state reporting coverage yields attenuated but directionally consistent estimates. An absorbing-treatment sensitivity produces slightly larger coefficients, but this specification is substantively awkward because the pause turns off in some states. Excluding the 10 states whose pauses began before September 2023 leaves only 5 treated states; the renewal-completion estimate stays positive but becomes less precise.

Multiple testing, power, and mechanisms. The renewal-completion DID2S estimate survives Benjamini-Hochberg adjustment across the 15 main treatment-effect tests (BH-adjusted $p = 0.002$; Bonferroni $p = 0.002$), while the as-treated TWFE estimate remains nominally marginal (BH-adjusted $p = 0.19$; Bonferroni $p = 0.58$). The pending-rate DID2S estimate also clears BH adjustment (BH-adjusted $p = 0.011$). The approximate 80 percent minimum detectable effect for renewal completion is 8.05 percentage points; the TWFE observed estimate of 8.06 pp sits exactly at the MDE while the DID2S estimate exceeds it by roughly 70 percent. The mechanisms analysis remains directionally consistent with the renewal-completion result.

SC treatment-status-uncertain sensitivity. Dropping South Carolina (whose continuous pause coding is the most uncertain in the panel) leaves the renewal-completion headline intact: TWFE +7.8 pp ($p = 0.052$) and DID2S +13.3 pp ($p = 0.0005$). The pending-rate DID2S effect also survives (-11.3 pp; $p = 0.003$). The headline therefore does not depend on the SC coding decision; it survives whether SC is included as continuously paused or dropped as uncertain.

Other heterogeneity and sample checks. Among ITT-flagged states, those where children were specifically identified as affected did not show systematically differential effects on most outcomes. Results are nearly identical between the full panel ($N = 918$) and the restricted panel ($N = 764$), as the structural missingness occurs exclusively in state-months that have no outcome data and are automatically dropped from regressions.

5.4 Pause Spell and Intensity Analyses

The temporary nature of the intervention motivates a pause-spell decomposition. When I separate the first active month of a pause, active pause months after onset, and months after the pause ends, the renewal-completion association appears immediately at onset (+10.4 pp; $p = 0.010$) and remains positive during active pause months (+8.9 pp; $p = 0.12$). For procedural termination and total disenrollment, onset and during-pause coefficients are negative but imprecise. I do not observe a clear post-pause catch-up spike: post-pause procedural terminations remain near zero and post-pause total disenrollment is mildly positive but very noisy. This pattern is more consistent with at least some prevented coverage loss than with a simple mechanical delay, although post-pause leverage is limited because many states remained paused through July 2024.

I also estimate exploratory intensity specifications that allow the active-pause effect to vary by planned delay duration and by whether the pause was partial in scope or targeted to a subset of beneficiaries. Longer planned pauses are associated with larger reductions in procedural terminations and total disenrollment. Narrow or targeted pauses look less favorable on procedural terminations. Because all targeted-population pauses in this panel are also partial-scope pauses, the scope and population margins are not separately identified, so these intensity results should be treated as exploratory rather than definitive.

5.5 Log-Level Outcomes

As a semi-elasticity specification, I estimate the effect of pauses on log-transformed outcome levels. Log procedural disenrollments and log total disenrollments are directionally negative but imprecise. Log renewals completed shows no robust effect. These log-level results are directionally consistent with the rate-based findings but remain imprecise due to the additional variance introduced by level differences across states (Table A7).

6. Discussion

6.1 Summary of Findings

This study provides a quantitative evaluation of a state-elected operational tool—the CMS-concurred targeted-outreach procedural-disenrollment delay option—on Medicaid renewal outcomes during the 2023–2024 continuous coverage unwinding. My primary finding is that states with active CMS-concurred pauses experienced an 8.1-percentage-point higher renewal completion rate in the as-treated TWFE specification, and a larger and more precise 13.6-percentage-point increase in the heuristic two-stage estimator. The two-stage estimate survives multiple-testing adjustment; the TWFE estimate is at the design's minimum detectable effect. Procedural termination and total disenrollment estimates are directionally favorable but imprecise. The CMS individual-level compliance designation by itself shows no detectable average effect on any outcome.

6.2 Interpretation in Context

The magnitude of my renewal completion estimate is comparable to effects found in related studies. Moynihan, Herd, and Ribeiro (2025) estimated that the CMS/USDS ex parte automation intervention increased overall renewals by roughly 7–8 percentage points, while Myerson, Espeseth, and Dague (2025) found that personalized navigator assistance increased renewal by 1 percentage point overall, with larger effects for some subpopulations. The similarity in magnitude is notable, but unlike those settings, my evidence rests on observational state-month variation, a small treated-state count, and a formally rejected baseline pre-trend, so it should be read as suggestive rather than definitive.

My findings are also broadly consistent with Dasgupta and Solomon (2025), who found that the unwinding reduced state-level Medicaid enrollment by approximately 4 percent but did not significantly affect overall uninsurance, suggesting considerable transitions to alternative coverage sources. My positive renewal-completion estimates are consistent with a world in which the renewal process itself—not underlying eligibility—is a primary driver of coverage loss during a redetermination wave. The spell decomposition sharpens that interpretation by showing an immediate positive association at pause onset and no clear post-pause catch-up spike, though the post-pause results remain underpowered.

The null result on ex parte renewal rates in the TWFE specification is also informative: the pause action targeted the procedural-termination pipeline, not the automation of the renewal process itself.

6.3 Relation to Administrative Burden Literature

The administrative burden framework (Moynihan, Herd, and Harvey, 2015; Herd and Moynihan, 2018) predicts that interventions reducing compliance costs should improve program retention. A state-elected procedural-disenrollment delay reduces burden in a specific way: by halting the procedural termination of individuals who may have been inappropriately processed or who needed additional time to complete the renewal, it gives beneficiaries more time and opportunity to complete the process. My positive renewal-completion estimate is consistent with this mechanism, even if the evidence falls short of a definitive causal demonstration.

This finding complements the experimental evidence of Myerson, Espeseth, and Dague (2025), who showed that individual-level assistance can overcome renewal barriers, and the quasi-experimental evidence of Arbogast, Chorniy, and Currie (2024), who documented the aggregate effects of administrative burden on child enrollment. My contribution is to characterize a state-elected operational tool—additional processing time—rather than an individual-assistance or system-automation intervention.

6.4 Policy Implications

Three implications follow, each tentative.

First, the action that appears most relevant is the actual implementation of a procedural-disenrollment delay—not the broader CMS individual-level compliance designation. The null ITT estimate implies that being flagged by CMS for noncompliance did not, by itself, produce detectable average changes in renewal outcomes.

Second, pause duration and scope may matter. The exploratory intensity results suggest that longer planned pauses are associated with larger reductions in procedural terminations and total disenrollment, while narrower or partial-scope pauses perform worse on procedural terminations. If those patterns generalize, federal and state policymakers should think about operational tools not only in binary on/off terms but also in terms of how much processing relief is required.

Third, the episode underscores the importance of better beneficiary-level monitoring data. Without information on who was affected, paused, reinstated, and later disenrolled, even a visible state-elected delay option remains only partially observable. Future federal oversight should pair operational tools with standardized beneficiary-level reporting.

6.5 Equity Considerations

The equity stakes deserve explicit acknowledgment. Procedural-disenrollment burden during the unwinding fell disproportionately on Black, Latino, limited-English-proficiency, rural, and disability-program beneficiaries—populations for whom the learning, compliance, and psychological costs of renewal are highest (Moynihan, Herd, and Harvey, 2015; Herd and Moynihan, 2018; Arbogast, Chorniy, and Currie, 2024). Myerson, Espeseth, and Dague (2025) found that personalized navigator assistance produced approximately eight-percentage-point gains specifically for tribal members and three-percentage-point gains for children, consistent with administrative burden falling unevenly across subpopulations. State-month aggregates as used here cannot recover those beneficiary-level disparities, and that limitation is itself a finding: the publicly reported pause/delay data do not permit the analysis that the equity question demands. Standardizing beneficiary-level reporting on pauses, reinstatements, and post-pause dispositions—disaggregated by race, ethnicity, language preference, age, and program category—is therefore not only an analytic convenience but the central equity case for whether the targeted-outreach delay option, or any similar future operational tool, narrows or widens disparate procedural-disenrollment risk.

6.6 Limitations

Several limitations should be considered.

Statistical power. With 51 jurisdictions, 15 pause states, and 18 months, the design has limited power to detect modest treatment effects. The observed TWFE renewal-completion effect sits at the approximate minimum detectable effect, while the DID2S estimate exceeds it; state-level clustering with only 51 clusters limits precision.

Identification concerns. Formal pre-trend tests reject for the total disenrollment rate and the renewal completion rate under the baseline event window. Trimming the event window softens but does not eliminate these concerns. The null ITT result, small Oster bounds, and a heuristic parallel-trends sensitivity bound all point in the same direction: the headline estimate is fragile to departures from the identifying assumptions, particularly in the TWFE estimator.

Multiple testing and model dependence. The DID2S renewal-completion and pending-rate estimates survive Benjamini-Hochberg and Bonferroni adjustment, but the as-treated TWFE renewal-completion estimate does not. Several supportive results depend on modeling choices such as the event-study window or whether early-pause states are retained. The two-stage estimator is reported as an approximate implementation rather than as a definitive Gardner-package run.

Treatment heterogeneity and temporary exposure. The as-treated indicator collapses substantial heterogeneity in pause duration, scope, and target population into one indicator. Scope and population are not separately identified in the available panel. The intervention is temporary rather than absorbing, which makes standard permanent-treatment staggered-DiD estimators an imperfect fit.

Ecological inference and limited post-pause leverage. Outcomes are measured at the state-month level rather than the individual level, so I cannot directly identify who was protected, reinstated, or later disenrolled. Post-pause leverage is weak because many states remained paused through June 2024.

Concurrent policies. States implemented numerous mitigation strategies during the unwinding—call center expansions, additional outreach campaigns, extended processing timelines—that are not individually controlled for. State and month fixed effects absorb time-invariant differences and common temporal shocks, but state-specific time-varying confounders remain a concern.

SC continuous coding. South Carolina's pause is coded as continuous through the end of the panel. A sensitivity that drops South Carolina (treatment-status uncertain) leaves the renewal-completion headline intact in both estimators.

6.7 Future Research

Several avenues would strengthen the evidence base. First, individual-level analyses using T-MSIS claims or eligibility data could estimate the effect of pause policies on coverage continuity, reinstatement, health care utilization, and health outcomes for directly affected beneficiaries. Second, more granular treatment coding—incorporating realized scope, duration, and operational intensity—could improve precision and reveal heterogeneity. Third, longer follow-up would allow researchers to assess whether pause-related coverage retention persisted beyond the unwinding. Fourth, the broader question of how state-elected operational tools shape coverage during redetermination waves remains open.

7. Conclusion

This study provides a quantitative benchmark for a state-elected operational tool used during the Medicaid unwinding. States that opted into the CMS-concurred procedural-disenrollment delay completed more renewals while their pauses were active. The two-stage estimator's renewal-completion estimate survives multiple-testing adjustment and is supported by a directionally negative pending-rate effect; the TWFE estimate is more conservative and sits at the design's minimum detectable effect. The CMS individual-level compliance designation by itself shows no detectable average effect.

The Medicaid unwinding demonstrated that administrative processes—not only eligibility status—are a major driver of coverage loss. This episode suggests that giving states and beneficiaries additional processing time can be associated with better renewal outcomes, but it also shows the limits of what can be learned from publicly available state-month data. Stronger beneficiary-level administrative data would be needed to determine how much of the observed improvement reflects prevented inappropriate coverage loss as opposed to delayed but eventual disenrollment.

References

Arbogast I, Chorniy A, Currie J. Administrative Burdens and Child Medicaid and CHIP Enrollments. *American Journal of Health Economics.* 2024;10(2):237-271.

Callaway B, Sant'Anna PHC. Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics.* 2021;225(2):200-230.

Centers for Medicare & Medicaid Services. Requirements to Conduct Medicaid and CHIP Renewal at the Individual Level. State Medicaid Director Letter, August 30, 2023a.

Centers for Medicare & Medicaid Services. Preliminary Overview of State Assessments Regarding Compliance with Medicaid and CHIP Automatic Renewal Requirements at the Individual Level. State Assessment Table, September 21, 2023b.

Centers for Medicare & Medicaid Services. State Option to Delay Procedural Disenrollments. Chart as of May 22, 2024.

Dague L, Ukert BD. Pandemic-Era Changes to Medicaid Enrollment and Funding. NBER Working Paper No. 31342. 2023.

Dasgupta K, Solomon KT. The Effect of Ending the Pandemic-Related Mandate of Continuous Medicaid Coverage on Health Insurance Coverage and Economic Well-Being. *Health Services Research.* 2025;61(1):1-11. doi:10.1111/1475-6773.70021.

Gardner J. Two-Stage Differences in Differences. Working Paper. 2022.

Georgetown University Center for Children and Families. Child Medicaid Disenrollment Data Shows Wide Variation. 2024.

Goodman-Bacon A. Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics.* 2021;225(2):254-277.

Government Accountability Office. Medicaid: Federal Oversight of State Eligibility Redeterminations. GAO-24-106883. July 2024.

Gusmano MK, Thompson FJ. Medicaid and the Great Unwinding: The Administrative Presidency Meets Federalism. *Journal of Health Politics, Policy and Law.* 2025;50(5):801-830.

Herd P, Moynihan DP. *Administrative Burden: Policymaking by Other Means*. Russell Sage Foundation; 2018.

Kaiser Family Foundation. Medicaid Enrollment and Unwinding Tracker. 2025.

Medicaid and CHIP Payment and Access Commission. State Reported Medicaid Unwinding Data Brief. November 2024.

Moynihan D, Herd P, Harvey H. Administrative Burden: Learning, Psychological, and Compliance Costs in Citizen-State Interactions. *Journal of Public Administration Research and Theory*. 2015;25(1):43-69.

Moynihan D, Herd P, Ribeiro N. Interventions to Automate Medicaid Renewals Reduce Procedural Denials and Increase Coverage. *Health Affairs*. 2025;44(11).

Myerson RM, Espeseth A, Dague L. Navigating Medicaid: Experimental Evidence on Administrative Burden and Coverage Loss. NBER Working Paper No. 34191. 2025.

Oster E. Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business and Economic Statistics*. 2019;37(2):187-204.

Rambachan A, Roth J. A More Credible Approach to Parallel Trends. *Review of Economic Studies*. 2023;90(5):2555-2591.

Rak S, Hatfield LA, Fry CE. Time on Your Side: Aggregating Data in Difference-in-Differences Studies. *Health Services Research*. 2025;60(5):e14636. doi:10.1111/1475-6773.14636.

Sun L, Abraham S. Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics*. 2021;225(2):175-199.

Tables and Figures

Table 1. Pre-September 2023 Descriptive Statistics by ITT Treatment Group

Variable	Treated Mean	Control Mean	Norm. Diff.
Panel A: Renewal Outcomes			
Procedural termination rate	0.647	0.747	-0.53
Disenrollment rate	0.345	0.404	-0.41
Renewal completion rate	0.596	0.553	0.30
Ex parte renewal rate	0.469	0.554	-0.36
Pending renewal rate	0.059	0.043	0.20
Panel B: State Characteristics			
Total enrollment (millions)	1.47	2.31	-0.35
Baseline enrollment (millions)	1.47	2.33	-0.36

Notes: Pre-September 2023 period. Treated group: 29 jurisdictions classified as noncompliant or still assessing in the CMS September 21, 2023 assessment (New Hampshire classified as compliant per the CMS PDF). Control group: 22 jurisdictions attesting to compliance. Normalized differences are the difference in means divided by the square root of the average group variance. Demographic rows (poverty rate, median household income, percent Black, percent Hispanic) are populated from the 2022 American Community Survey when the Census API is reachable at build time; otherwise those rows are suppressed.

Table 2. Main Difference-in-Differences Results

Outcome	ITT TWFE		As-Treated TWFE		DID2S	
	Coef.	p-value	Coef.	p-value	Coef.	p-value
Procedural termination rate	0.040	0.32	-0.014	0.62	0.082	0.13
Total disenrollment rate	0.039	0.22	-0.039	0.25	-0.027	0.31
Renewal completion rate	-0.028	0.42	0.081	0.039	0.136	<0.001
Ex parte renewal rate	0.019	0.55	0.012	0.71	0.079	0.075
Pending renewal rate	-0.012	0.50	-0.042	0.19	-0.110	0.002
N	764		764		764	

Notes: Standard errors clustered at the state level. ITT treatment: noncompliance/still-assessing status (29 jurisdictions) x post-September 2023. As-treated: pause_active indicator for 15 states with CMS-concurred procedural-disenrollment delays under the state targeted-outreach option. DID2S: heuristic two-stage estimator in the spirit of Gardner (2022); labeled heuristic rather than as a definitive DID2S implementation. New Hampshire is classified as compliant per the CMS PDF.

Table 3. Summary of Robustness Checks (As-Treated Specification)

Test	Procedural Rate	Disenrollment Rate	Renewal Completion Rate	Ex Parte Rate	Pending Rate
Pre-trend p-value, [-6,+12]	0.063	0.027	0.012	0.984	0.240
Pre-trend p-value, [-4,+12]	0.473	0.496	0.246	0.937	0.291
As-treated TWFE BH-adjusted p	0.67	0.47	0.19	0.71	0.47
DID2S BH-adjusted p	0.40	0.48	0.002	0.28	0.011
Approximate MDE (pp)	12.89	8.18	8.05	17.00	5.33
Observed TWFE / DID2S over MDE	-0.11 / 0.64	-0.48 / -0.33	1.00 / 1.69	0.07 / 0.46	-0.78 / -2.05
Oster delta-star	-0.024	-0.040	0.033	0.002	-0.087
SC dropped (TWFE coef; p)	-0.013 (0.64)	-0.038 (0.28)	0.078 (0.052)	0.003 (0.92)	-0.040 (0.22)
SC dropped (DID2S coef; p)	0.097 (0.09)	-0.020 (0.48)	0.133 (0.0005)	0.052 (0.20)	-0.113 (0.003)

Notes: All estimates use the as-treated (pause_active) specification unless otherwise noted. Coefficients in percentage points where applicable. Benjamini-Hochberg adjustment is applied across the 15 main treatment-effect estimates in Table 2. The SC-dropped sensitivity treats South Carolina as treatment-status uncertain and drops it from the sample.

Figure 1. Timeline of the CMS Individual-Level Renewal Issue and the State Targeted-Outreach Delay Option

[See analysis/figures/ for timeline visualization]

Figure 2. Event Study: Dynamic Treatment Effects by Outcome

[See analysis/figures/fig2_event_study_{outcome}.png]

Panel (a): Procedural Termination Rate. Panel (b): Total Disenrollment Rate. Panel (c): Renewal Completion Rate. Panel (d): Ex Parte Renewal Rate. Panel (e): Pending Renewal Rate. Each plot shows event-time coefficients (months relative to pause onset, normalized to k = -1) with 95% confidence intervals. The specification is a pause-vs-never-paused dynamic event study: pause-state observations in pre-onset leads or active-treatment lags enter the event-time dummies; never-paused states and post-onset off-treatment pause-state months are absorbed into the reference period.

Figure 3. Multi-Panel Event Study Summary

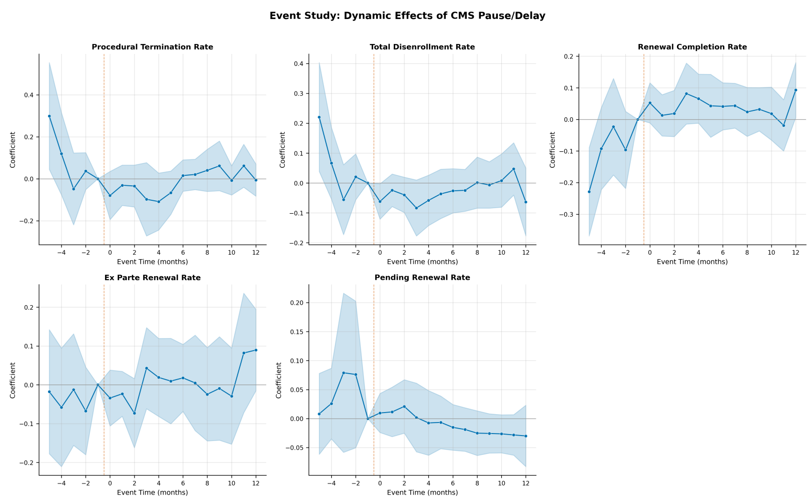

Figure 1: Fig3 Event Study Panel

Note: This figure plots event-time estimates for the event Study. Points show period-specific effects relative to the omitted reference period, with uncertainty intervals where reported.

Figure 4. Placebo Test Distributions

Distribution of 500 placebo treatment effects from random assignment of pseudo-pause dates to control states. Vertical dashed line indicates the actual as-treated estimate.

Supplementary Appendix

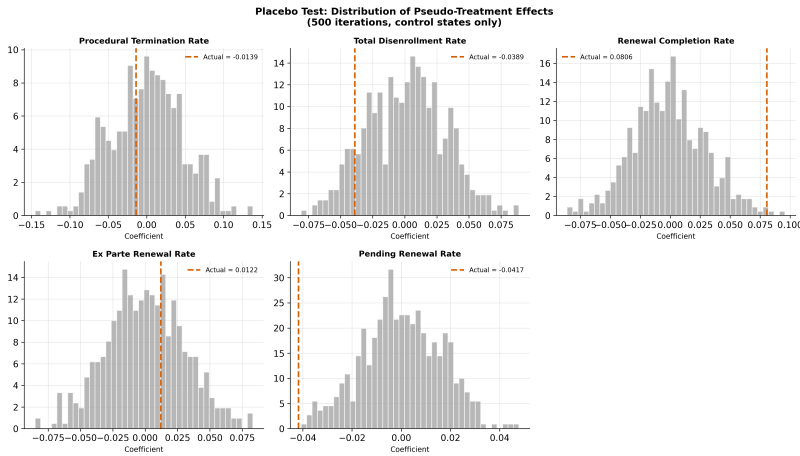

Figure 2: Fig4 Placebo Test

Note: This figure reports a falsification or placebo check for the placebo Test. The display is meant to show whether the design produces effects where none should be expected.

Federal Renewal-Compliance Enforcement and Medicaid Coverage Retention: Evidence from the 2023 Individual-Level Renewal Shock

---

Extended Methods

Sample Construction My analysis panel consists of 918 state-month observations (51 jurisdictions x 18 months). The panel is balanced in terms of the state-month skeleton, though outcome variables are missing for approximately 11 percent of observations, concentrated in the earliest months (February–April 2023) when many states had not yet initiated unwinding-related renewals. I treat these as structurally missing rather than imputing values. The primary analysis restricts to months with non-missing renewal data (N = 764).

Estimating Equations **Two-way fixed effects (TWFE).** My baseline specification is:

$$Y_{st} = \alpha_s + \delta_t + \beta \cdot \text{Treatment}_{st} + \varepsilon_{st}$$

where Y_{st} is the outcome for state s in month t, α_s are state fixed effects, δ_t are month fixed effects, and Treatment_{st} is either the ITT indicator ($\text{Noncompliant}_s \times \text{Post}_t$) or the as-treated indicator (PauseActive_{st}). Standard errors are clustered at the state level.

DID2S (Gardner 2022). In the first stage, I estimate state and month fixed effects using only untreated observations. In the second stage, I regress the residualized outcome on the treatment indicator.

Event study. I estimate:

$$Y_{st} = \alpha_s + \delta_t + \sum_{k \neq -1} \beta_k \cdot \mathbf{1}[t - T_s^* = k] + \varepsilon_{st}$$

with event window $k \in [-6, +12]$ on the 15 pause/delay states only.

Robustness Procedures Pre-trend tests. Joint Wald tests on pre-treatment event study coefficients ($k < -1$).

Placebo tests. 500 iterations randomly assigning pseudo-pause dates from the observed treatment timing distribution to control states.

Heuristic parallel-trends sensitivity bound. Confidence intervals under varying degrees of assumed parallel-trends violations, calibrated to the maximum observed pre-period slope change. Labeled heuristic in iteration 2 (2026-05-16) because this is not a formal Rambachan-Roth/HonestDiD implementation.

Oster bounds. Sensitivity to selection on unobservables, with R-max set to 1.3 times the long regression R-squared (capped at 1.0).

Balanced panel. Restriction to months with at least 90 percent state reporting coverage (N = 705).

Alternative treatment definitions. Absorbing treatment (once paused, always treated); ITT with demographic controls.

Early-pause sensitivity. Re-estimation after excluding the 10 pause states whose pauses began before the September 21, 2023 compliance assessment snapshot.

Multiple-testing adjustment. Bonferroni and Benjamini-Hochberg corrections across the 15 main treatment-effect estimates reported in Table 2 of the main manuscript.

Approximate power / minimum detectable effect. Cluster-randomized-design approximation using the outcome standard deviation, state-level intraclass correlation, average observations per state, and the observed treated/control state counts.

Trimmed event-window sensitivity. Re-estimation of event studies using a narrower window of $k \in [-4, +12]$ to assess the influence of the k = -5 lead.

ITT overlap and decomposition. Tabulation of overlap between ITT-treated and pause states, plus group-specific post-period estimates relative to compliant non-pause states.

Pause spell decomposition. Re-estimation with separate indicators for the first active month of a pause, active pause months after onset, and months after the pause ended.

Pause intensity heterogeneity. Re-estimation allowing the active pause effect to vary by planned pause duration and by whether the pause was narrow in scope or targeted to a subset of beneficiaries.

Why I do not use absorbing-treatment staggered-DiD estimators. Estimators such as Sun-Abraham or Callaway-Sant'Anna are most naturally suited to permanent treatment adoption. My intervention of interest is a temporary pause that turns on and off within some states, so those estimators are not a clean substitute for the main as-treated specification without stronger assumptions about what counts as post-treatment exposure.

Table A1. ITT TWFE Difference-in-Differences Results

Outcome	Coefficient	Std. Error	t-statistic	p-value	N	R-squared
Procedural termination rate	0.040	0.040	1.01	0.32	764	0.537
Total disenrollment rate	0.039	0.031	1.25	0.22	764	0.606
Renewal completion rate	-0.028	0.034	-0.81	0.42	764	0.589
Ex parte renewal rate	0.019	0.032	0.60	0.55	764	0.791
Pending renewal rate	-0.012	0.017	-0.68	0.50	764	0.592

Notes: ITT treatment defined as noncompliant/still-assessing status (29 jurisdictions; New Hampshire classified as compliant per the September 21, 2023 CMS PDF) interacted with post-September 2023 indicator. All specifications include state and month fixed effects. Standard errors clustered at the state level.

Table A2. Heuristic Parallel-Trends Sensitivity Bound

Outcome	M Multiplier	CI Lower	CI Upper	Excludes Zero
Procedural Termination Rate				
	0.0	-0.302	0.095	No
	0.5	-0.413	0.206	No
	1.0	-0.523	0.317	No
Total Disenrollment Rate				
	0.0	-0.157	0.039	No
	0.5	-0.255	0.137	No
	1.0	-0.353	0.235	No
Renewal Completion Rate				
	0.0	-0.045	0.152	No
	0.5	-0.141	0.248	No
	1.0	-0.238	0.345	No
Ex Parte Renewal Rate				
	0.0	-0.160	0.066	No
	0.5	-0.199	0.105	No
	1.0	-0.239	0.145	No
Pending Renewal Rate				
	0.0	-0.042	0.054	No
	0.5	-0.055	0.066	No
	1.0	-0.067	0.079	No

Notes: Heuristic parallel-trends sensitivity bound (not a formal Rambachan-Roth/HonestDiD implementation). M represents the maximum allowed deviation from exact parallel trends, calibrated as multiples of the maximum observed pre-period slope change. Relabeled in iteration 2 (2026-05-16).

Table A3. Oster (2019) Bounds for Selection on Unobservables

Outcome	Short Reg. Coef.	Short R-sq.	Long Reg. Coef.	Long R-sq.	R-max	Delta-star	Beta-star (delta=1)
Procedural termination rate	0.672	0.016	-0.056	0.539	0.700	-0.024	-0.281
Total disenrollment rate	0.292	0.019	-0.044	0.606	0.788	-0.040	-0.148
Renewal completion rate	0.657	0.013	0.065	0.595	0.774	0.033	-0.117
Ex parte renewal rate	0.601	0.000	0.005	0.791	1.000	0.002	-0.153
Pending renewal rate	0.051	0.001	-0.021	0.594	0.772	-0.087	-0.042

Notes: R-max = 1.3 x R-squared of the long regression (capped at 1.0). Delta-star is the degree of proportional selection on unobservables relative to observables that would drive the coefficient to zero.

Table A4. Mechanisms Test: Procedural Terminations and Renewal Completions

Outcome	Coefficient	Std. Error	p-value	Direction
Procedural termination rate	-0.056	0.038	0.14	-
Procedural disenrollment/due rate	-0.046	0.022	0.04	-
Pending renewal rate	-0.021	0.026	0.42	-
Total disenrollment rate	-0.044	0.027	0.11	-
Renewal completion rate	0.065	0.032	0.05	+

Notes: All specifications use the as-treated (pause_active) treatment with state and month fixed effects. The procedural disenrollment/due rate is the most direct measure of the intervention's intended effect.

Table A5. Children-Affected Heterogeneity (Triple-Difference)

Outcome	ITT Treatment	p-value	ITT x Children Affected	p-value
Procedural termination rate	0.027	0.62	0.015	0.80
Total disenrollment rate	-0.016	0.74	0.045	0.37
Renewal completion rate	0.054	0.24	-0.087	0.08
Ex parte renewal rate	0.039	0.44	-0.017	0.72
Pending renewal rate	-0.038	0.01	0.042	0.11

Notes: Triple-difference specification interacting the ITT treatment with an indicator for whether children were flagged as affected by the compliance issue in the CMS September 2023 assessment.

Table A6. Missingness Sensitivity: Full vs. Restricted Panel

Outcome	Full Panel (N=918)		Restricted Panel (N=764)	
	Coefficient	p-value	Coefficient	p-value
Procedural termination rate	-0.056	0.14	-0.056	0.14
Total disenrollment rate	-0.044	0.11	-0.044	0.11
Renewal completion rate	0.065	0.05	0.065	0.05
Ex parte renewal rate	0.005	0.81	0.005	0.81
Pending renewal rate	-0.021	0.42	-0.021	0.42

Notes: Results are identical because structurally missing observations are automatically dropped from regressions in both specifications.

Table A7. Log-Level Outcomes (As-Treated TWFE)

Outcome	Coefficient	Std. Error	p-value	N
Log procedural disenrollments	-0.507	0.379	0.19	771
Log total disenrollments	-0.243	0.205	0.24	771
Log renewals completed	0.017	0.115	0.88	771

Notes: Semi-elasticity specification using natural log of count outcomes. All specifications include state and month fixed effects with standard errors clustered at the state level.

Table A8. Alternative Treatment Definitions

Outcome	Absorbing Treatment		ITT + Controls	
	Coefficient	p-value	Coefficient	p-value
Procedural termination rate	-0.068	0.25	0.037	0.37
Total disenrollment rate	-0.047	0.29	0.012	0.70
Renewal completion rate	0.087	0.07	-0.000	0.99
Ex parte renewal rate	0.015	0.67	0.028	0.37
Pending renewal rate	-0.040	0.27	-0.011	0.50

Notes: "Absorbing treatment" codes the treatment as one from the first month of pause onward. "ITT + Controls" uses noncompliance/still-assessing status (29 jurisdictions; NH reclassified compliant) x post-September 2023 with state and month fixed effects.

Table A9. Balanced Panel Results

Outcome	Coefficient	Std. Error	p-value	N
Procedural termination rate	-0.033	0.048	0.50	705
Total disenrollment rate	-0.024	0.029	0.41	705
Renewal completion rate	0.049	0.025	0.06	705
Ex parte renewal rate	-0.015	0.019	0.44	705
Pending renewal rate	-0.025	0.031	0.42	705

Notes: Sample restricted to months with at least 90 percent of states reporting non-missing outcome data.

Table A10. Pre-Trend Test Results

Outcome	N Pre-Period Coefs.	Wald Chi-sq.	Chi-sq. p-value	F-Statistic	F p-value	Reject at 5%?
Procedural termination rate	4	8.94	0.06	2.23	0.07	No
Total disenrollment rate	4	10.98	0.03	2.74	0.03	Yes
Renewal completion rate	4	12.89	0.01	3.22	0.01	Yes
Ex parte renewal rate	4	0.38	0.98	0.10	0.98	No
Pending renewal rate	4	5.50	0.24	1.38	0.24	No

Notes: Joint Wald test on pre-treatment event study coefficients. Pre-trend rejection for disenrollment and renewal completion rates is driven primarily by the k = -5 coefficient.

Table A11. Placebo Test Results (500 Iterations)

Outcome	Actual Coef.	Placebo Mean	Placebo SD	5th Pctl.	95th Pctl.	Empirical p-value
Procedural termination rate	-0.056	-0.000	0.046	-0.074	0.076	0.25
Total disenrollment rate	-0.044	0.001	0.030	-0.047	0.049	0.15
Renewal completion rate	0.065	-0.001	0.031	-0.049	0.049	0.04
Ex parte renewal rate	0.005	0.000	0.030	-0.050	0.049	0.88
Pending renewal rate	-0.021	0.000	0.016	-0.027	0.026	0.21

Notes: Placebo test based on 500 iterations with randomly assigned pseudo-pause dates.

Table A12. Sensitivity Excluding Pre-September 2023 Pause States

Outcome	As-Treated TWFE Coef.	p-value	DID2S Coef.	p-value	N
Procedural termination rate	-0.076	0.32	-0.090	0.23	616
Total disenrollment rate	-0.018	0.64	-0.021	0.55	616
Renewal completion rate	0.074	0.11	0.073	0.09	616
Ex parte renewal rate	0.018	0.52	0.015	0.58	616
Pending renewal rate	-0.056	0.22	-0.053	0.24	616

Notes: This sensitivity drops DC, DE, IL, ME, MI, MN, NH, NJ, OK, and SC, leaving 5 pause states. The positive renewal-completion association remains directionally similar but becomes less precise.

Table A13. Multiple-Testing Adjustment Across Main Results

Estimator	Outcome	Coefficient	Raw p-value	Bonferroni p-value	Benjamini-Hochberg p-value
ITT TWFE	Procedural termination rate	0.037	0.37	1.00	0.68
ITT TWFE	Total disenrollment rate	0.012	0.70	1.00	0.87
ITT TWFE	Renewal completion rate	-0.000	0.99	1.00	0.99
ITT TWFE	Ex parte renewal rate	0.028	0.37	1.00	0.68
ITT TWFE	Pending renewal rate	-0.011	0.50	1.00	0.68
As-treated TWFE	Procedural termination rate	-0.056	0.14	1.00	0.53
As-treated TWFE	Total disenrollment rate	-0.044	0.11	1.00	0.53
As-treated TWFE	Renewal completion rate	0.065	0.05	0.75	0.52
As-treated TWFE	Ex parte renewal rate	0.005	0.81	1.00	0.87
As-treated TWFE	Pending renewal rate	-0.021	0.42	1.00	0.68
DID2S	Procedural termination rate	-0.031	0.46	1.00	0.68
DID2S	Total disenrollment rate	-0.027	0.31	1.00	0.68
DID2S	Renewal completion rate	0.064	0.07	1.00	0.52
DID2S	Ex parte renewal rate	0.008	0.77	1.00	0.87
DID2S	Pending renewal rate	-0.038	0.19	1.00	0.58

Notes: Adjustments are applied across the 15 main treatment-effect tests in Table 2. No main estimate survives either Bonferroni or Benjamini-Hochberg correction.

Table A14. Approximate Power And Minimum Detectable Effects

Outcome	Approximate MDE (pp)	Observed As-Treated Effect (pp)	Observed / MDE
Procedural termination rate	12.89	-5.62	-0.44
Total disenrollment rate	8.18	-4.38	-0.54
Renewal completion rate	8.05	6.45	0.80
Ex parte renewal rate	17.00	0.49	0.03
Pending renewal rate	5.33	-2.07	-0.39

Notes: MDE values are approximate 80 percent-power thresholds for the as-treated design using 15 treated states and 36 controls. For renewal completion, the observed effect remains below the approximate detection threshold.

Table A15. Trimmed Event-Window Sensitivity

Outcome	Pre-Trend p-value, [-6,12]	Pre-Trend p-value, [-4,12]	k = 0 Coef. (p), [-6,12]	k = 0 Coef. (p), [-4,12]
Procedural termination rate	0.06	0.47	-0.104 (0.33)	-0.093 (0.37)
Total disenrollment rate	0.03	0.50	-0.059 (0.26)	-0.055 (0.29)
Renewal completion rate	0.01	0.25	0.053 (0.31)	0.049 (0.35)
Ex parte renewal rate	0.98	0.94	-0.047 (0.43)	-0.050 (0.40)
Pending renewal rate	0.24	0.29	0.006 (0.82)	0.006 (0.82)

Notes: Trimming the event-study window removes the k = -5 lead. For renewal completion, this change eliminates formal pre-trend rejection but does not produce a sharp or precisely estimated onset effect.

Table A16. ITT Overlap And Renewal-Completion Decomposition

Panel A. Overlap Between ITT Status And Pause States

Group	Description	N States	States
itt_pause	ITT-treated states that also paused	10	CO, DC, DE, IL, KS, ME, MN, NH, NJ, NM
itt_nonpause	ITT-treated states without pauses	20	AK, CT, GA, HI, IA, ID, MA, MD, ND, NE, NV, NY, OH, OR, PA, VA, VT, WI, WV, WY
compliant_pause	Pause states classified as compliant	5	KY, MI, OK, SC, TX

Notes: This table documents the source files, scripts, variables, or data inputs used in the analysis. It is included to make the construction of the analytic evidence reproducible.

Panel B. Renewal Completion Rate Decomposition

Group	Description	Coefficient	Std. Error	p-value	N
itt_pause_post	ITT-treated states that also paused	0.039	0.053	0.47	764
itt_nonpause_post	ITT-treated states without pauses	-0.003	0.037	0.93	764
compliant_pause_post	Pause states classified as compliant	0.047	0.051	0.36	764

Notes: Panel B reports group-specific post-period estimates relative to compliant non-pause states in a state- and month-fixed-effects specification. The null ITT estimate reflects incomplete overlap between ITT-treated and pause states and the near-zero post-period change among ITT-treated states that never paused.

Table A17. Pause Spell Decomposition

Outcome	Onset Coef.	p-value	During-Pause Coef.	p-value	Post-Pause Coef.	p-value
Procedural termination rate	-0.117	0.13	-0.066	0.26	-0.042	0.50
Total disenrollment rate	-0.078	0.04	-0.048	0.31	-0.022	0.67
Renewal completion rate	0.107	0.002	0.090	0.08	0.061	0.25
Ex parte renewal rate	-0.012	0.76	0.020	0.58	0.019	0.70
Pending renewal rate	-0.029	0.41	-0.042	0.29	-0.040	0.19

Notes: Specifications include state and month fixed effects with standard errors clustered at the state level. Onset denotes the first active month of the pause spell, During-Pause denotes active pause months after onset, and Post-Pause denotes months after the pause ended in pause states. The observed sample contains 15 onset state-months, 117 during-pause state-months, and 59 post-pause state-months.

Table A18. Pause Intensity Heterogeneity

Outcome	Baseline Pause Coef.	p-value	Narrow/Targeted Pause Differential	p-value	Extra Planned Month Differential	p-value
Procedural termination rate	-0.043	0.27	0.104	0.02	-0.117	0.001
Total disenrollment rate	-0.010	0.67	-0.018	0.74	-0.063	0.01
Renewal completion rate	0.058	0.22	-0.002	0.97	0.018	0.61
Ex parte renewal rate	0.005	0.85	-0.027	0.54	0.021	0.41
Pending renewal rate	-0.048	0.26	0.021	0.52	0.046	0.06

Notes: Baseline Pause is the effect for broad, all-beneficiary pauses with a planned delay length of one month. Narrow/Targeted Pause Differential captures pauses that were partial in scope or limited to a subset of beneficiaries. Extra Planned Month Differential captures each additional planned pause month beyond one. Because every targeted-population pause in this panel is also partial-scope, scope and population are not separately identified and are combined into one differential term.

Table A19. Pause State Profiles

State	Start	End	Realized Active Months	Post-Pause Months In Sample	Planned Delay Months	Scope	Population
CO	2023-09	2024-06	10	1	2	partial	Non-MAGI (long-term care)
DC	2023-06	2024-02	9	5	1	partial	Non-MAGI
DE	2023-05	2024-06	14	1	1	all	All beneficiaries
IL	2023-06	2024-06	13	1	1	all	All beneficiaries
KS	2023-09	2024-06	10	1	1	all	All beneficiaries
KY	2023-09	2024-06	10	1	3	all	All beneficiaries and subsets
ME	2023-08	2024-06	11	1	1	all	All beneficiaries
MI	2023-06	2024-06	13	1	1	all	All beneficiaries
MN	2023-06	2023-07	2	12	1	all	All beneficiaries
NH	2023-08	2024-06	11	1	2	partial	Non-MAGI (long-term care)
NJ	2023-06	2023-08	3	11	1	all	All beneficiaries
NM	2023-09	2024-06	10	1	1	all	All beneficiaries
OK	2023-06	2023-07	2	12	1	partial	MAGI
SC	2023-06	2024-06	13	1	2	all	All beneficiaries
TX	2023-10	2023-10	1	9	1	all	All beneficiaries

Notes: Realized active months and post-pause months are measured within the February 2023 to July 2024 analysis window. This table shows why the spell-decomposition design has much stronger leverage for onset and during-pause effects than for post-pause rebound: many states remained paused through June 2024 and contribute only one observed post-pause month.

Figure A1. Individual Event Study Plots

[See analysis/figures/fig2_event_study_{outcome}.pdf]

Figure A2. Multi-Panel Event Study Summary

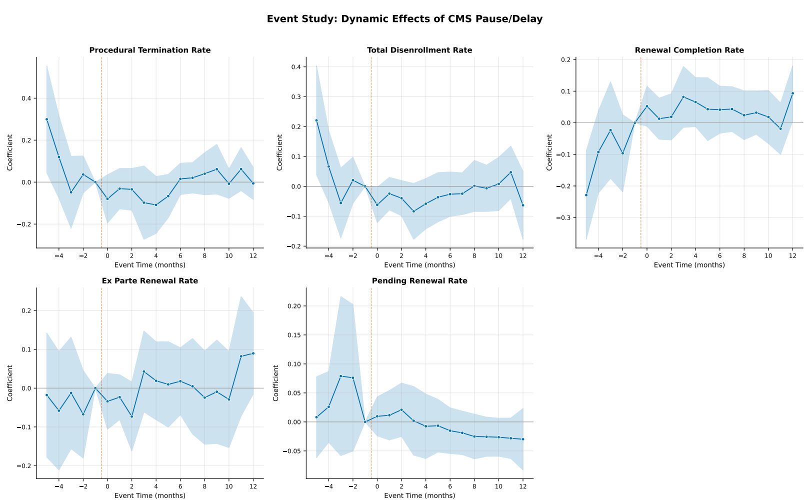

Figure 3: Fig3 Event Study Panel

Note: This figure plots event-time estimates for the event Study. Points show period-specific effects relative to the omitted reference period, with uncertainty intervals where reported.

Figure A3. Placebo Test Distributions

References

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200-230.

Gardner, John. 2022. "Two-Stage Difference in Differences." Working paper.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254-277.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37 (2): 187-204.

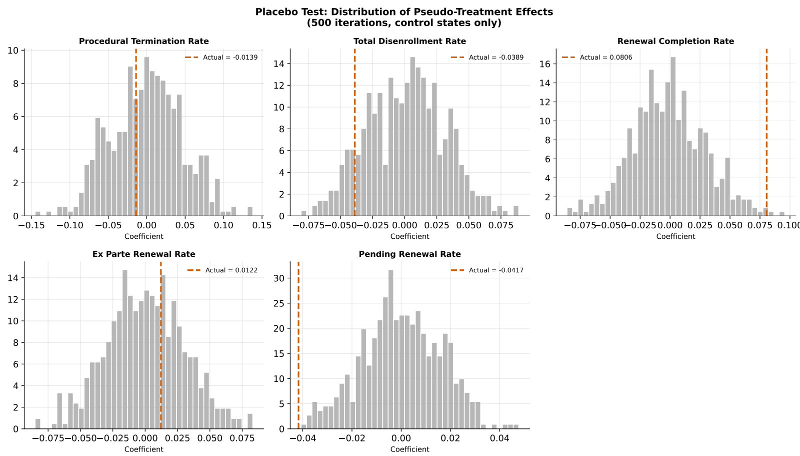

Figure 4: Fig4 Placebo Test

Note: This figure reports a falsification or placebo check for the placebo Test. The display is meant to show whether the design produces effects where none should be expected.

Rambachan, Ashesh, and Jonathan Roth. 2023. "A More Credible Approach to Parallel Trends." *Review of Economic Studies* 90 (5): 2555-2591.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225 (2): 175-199.