

# Paying for Pixels: The Null Effect of Telehealth Payment Parity on Medicaid Behavioral Health Provider Supply

APEP Autonomous Research\* @ailscl-auto

February 20, 2026

## Abstract

One in five American adults experiences mental illness, yet Medicaid beneficiaries face severe behavioral health provider shortages. Between 2020 and 2023, 26 states adopted telehealth payment parity laws requiring Medicaid to reimburse telehealth at the same rate as in-person services, aiming to expand provider supply. Using the universe of Medicaid claims from the Transformed Medicaid Statistical Information System (T-MSIS, 227 million records, 2018–2024), I evaluate whether these laws increased behavioral health provider participation. Applying Callaway-Sant’Anna difference-in-differences with never-treated states as controls, I find a precisely estimated null effect: the overall ATT for unique billing providers is 0.010 (SE = 0.049), with similarly null effects on beneficiaries, claims, and spending. A triple-difference comparing behavioral health to personal care services confirms the null. While parity removes a financial penalty for remote care, it cannot overcome the structural barriers—low base rates and high administrative costs—that keep providers out of Medicaid.

**JEL Codes:** I11, I13, I18, H75

**Keywords:** telehealth, payment parity, Medicaid, behavioral health, provider supply, difference-in-differences

---

\*Autonomous Policy Evaluation Project. This paper was generated autonomously. Total execution time: 7m (cumulative: 7m). Correspondence: scl@econ.uzh.ch

# 1. Introduction

America is drowning in a behavioral health crisis. Over 57 million adults—roughly one in five—experienced a mental illness in 2021, yet fewer than half received treatment ([Substance Abuse and Mental Health Services Administration, 2022](#)). The gap between need and supply is especially severe for Medicaid beneficiaries: despite covering 90 million Americans, Medicaid reimburses behavioral health providers at rates 29–47% below Medicare, driving widespread non-participation ([Zuckerman et al., 2021](#)). In a survey of psychiatrists, only 35.7% accepted new Medicaid patients, compared to 85.7% for privately insured patients ([Bishop et al., 2014](#)). The COVID-19 pandemic sharpened both edges of this crisis—surging demand for behavioral health services met an already thin provider workforce.

Telehealth emerged as the natural solution. By eliminating travel time, reducing overhead costs, and expanding geographic reach, telehealth could allow a single provider to serve patients across an entire state rather than a single clinic. But a critical policy barrier remained: many state Medicaid programs reimbursed telehealth visits at lower rates than in-person services, or did not cover them at all. Between 2020 and 2023, 26 states enacted permanent telehealth payment parity laws—requiring Medicaid to reimburse telehealth at the same rate as equivalent in-person services. These laws aimed directly at the provider participation margin: by guaranteeing equal pay for remote visits, they removed the financial penalty for telehealth delivery.

This paper asks a simple question: did these laws work? Specifically, did state telehealth payment parity laws expand the supply of behavioral health providers billing Medicaid? I exploit the staggered adoption of permanent parity laws across 26 states between 2020 and 2023, using the universe of Medicaid claims from the Transformed Medicaid Statistical Information System (T-MSIS). This dataset contains 227 million provider-service records covering all 50 states and the District of Columbia from January 2018 through December 2024, allowing me to track the complete billing universe of behavioral health providers at the state-quarter level.

The answer is no. Using Callaway-Sant’Anna (2021) doubly robust difference-in-differences with never-treated states as comparison groups, I estimate an overall ATT for the number of unique billing providers of 0.010 log points (SE = 0.049), statistically indistinguishable from zero. The null extends across all margins: beneficiaries served (ATT = -0.015, SE = 0.056), total claims (ATT = -0.009, SE = 0.077), and total Medicaid spending (ATT = 0.016, SE = 0.102). Event-study estimates show flat pre-treatment trends and no discernible post-treatment divergence through eight quarters after adoption. The implied percentage effects range from -1.5% to +1.6%, with 95% confidence intervals ruling out effects larger

than  $\pm 10\%$ .

The null is robust. A Goodman-Bacon (2021) decomposition confirms that 88% of the TWFE estimand’s weight falls on clean treated-versus-untreated comparisons, minimizing concerns about problematic two-way fixed effects bias. A triple-difference design—comparing behavioral health services (which can be delivered via telehealth) against personal care services (bathing, feeding, which cannot)—within the same state-time cells yields a treatment effect of  $-0.104$  ( $SE = 0.106$ ) for behavioral health, indistinguishable from the placebo personal care coefficient of  $0.004$  ( $SE = 0.049$ ). Leave-one-out analysis shows no single state drives the result: the TWFE coefficient ranges from  $-0.119$  to  $-0.092$  across all 26 permutations of dropping one treated state.

These results contribute to several literatures. First, I contribute to the growing body of work on telehealth policy (Mehrotra et al., 2021; Uscher-Pines et al., 2021; Barnett et al., 2018). While prior studies find that telehealth parity increases utilization among privately insured populations (Ellimoottil et al., 2022), my findings suggest these effects do not extend to Medicaid’s behavioral health workforce. The distinction matters: Medicaid providers face a different set of binding constraints—lower base rates, higher administrative burden, and more complex credentialing requirements—that payment parity alone cannot resolve.

Second, I contribute to the literature on Medicaid provider supply (Decker, 2012; Candon et al., 2018; Alexander and Schnell, 2017). The precise null result is consistent with a model in which the extensive margin of provider participation is determined not by the telehealth-versus-in-person rate differential, but by the absolute level of Medicaid reimbursement. If base rates are too low to cover the fixed costs of Medicaid billing—credentialing, compliance, documentation—equalizing the modality premium changes nothing.

Third, I contribute methodologically by applying modern heterogeneity-robust DiD methods (Callaway and Sant’Anna, 2021; Sun and Abraham, 2021; Goodman-Bacon, 2021; Roth, 2022) to a large administrative dataset with clean staggered adoption. The T-MSIS data—covering the universe of Medicaid claims with no sampling—provides uniquely high statistical power to detect or rule out policy effects at the national level.

Fourth, this paper demonstrates the value of honestly reported null results in policy evaluation (Abadie, 2020; Franco et al., 2014). The null finding has direct policy implications: states seeking to expand behavioral health access through telehealth should consider rate increases, administrative simplification, or credentialing reform as complements to—not substitutes for—payment parity mandates.

## 2. Institutional Background and Policy Setting

### 2.1 The Medicaid Behavioral Health Landscape

Medicaid is the single largest payer for behavioral health services in the United States, financing roughly 25% of all mental health spending and 21% of substance use disorder treatment nationally ([Substance Abuse and Mental Health Services Administration, 2022](#)). Yet Medicaid behavioral health reimbursement rates are substantially lower than other payers. A 2023 analysis found that Medicaid reimburses psychiatric services at rates 29–47% below Medicare fee-for-service, with even larger gaps relative to commercial insurance ([Melek et al., 2023](#)). These rate differentials directly affect provider supply: [Bishop et al. \(2014\)](#) document that only 35.7% of psychiatrists accept new Medicaid patients, compared to 85.7% for private insurance.

The behavioral health workforce itself is fragmented across provider types. Community behavioral health organizations—billing under organizational NPIs—deliver the majority of Medicaid-financed services through codes like H2015 (community support, individual), H2016 (community support, group), H0036 (community psychiatric support), and H0031 (mental health assessment). These HCPCS H-prefix codes are Medicaid-specific: they have no Medicare equivalent and are used exclusively for services financed through state Medicaid programs. Individual practitioners—psychiatrists, psychologists, licensed clinical social workers—may also bill Medicaid directly under their own NPIs, but many opt out due to low rates and high administrative burden.

The COVID-19 pandemic dramatically increased demand for behavioral health services while simultaneously demonstrating the potential of telehealth delivery. Emergency telehealth waivers, adopted by all 50 states between March and April 2020, temporarily suspended restrictions on telehealth modality and reimbursement. Telehealth behavioral health visits surged from less than 2% of all encounters pre-pandemic to over 40% by mid-2020 ([Terlizzi and Norris, 2022](#)). However, these emergency provisions were temporary, and their expiration created uncertainty about the long-term viability of telehealth-based behavioral health delivery.

### 2.2 Telehealth Payment Parity Laws

Telehealth “payment parity” refers to state laws requiring health insurers—including Medicaid—to reimburse telehealth services at the same rate as equivalent in-person services. Prior to the pandemic, approximately 16 states had some form of payment parity for commercial insurers, but far fewer extended parity explicitly to Medicaid programs ([Center for Connected Health](#)

[Policy, 2020](#)). The distinction between coverage parity (requiring that telehealth be a covered benefit) and payment parity (requiring equal reimbursement rates) is critical: a state may mandate coverage of telehealth without requiring rate equality, leaving providers financially penalized for remote delivery.

The policy shock I study is the wave of permanent telehealth payment parity laws adopted between 2020 and 2023. These laws are distinct from the temporary COVID-19 emergency waivers in two ways: they are permanent (no sunset date, or with extension provisions), and they mandate rate equality rather than simply permitting telehealth billing. Georgia was the first state to adopt permanent parity in January 2020, under SB 118. The main wave came in 2021, when 19 states enacted permanent parity laws between January and November—driven in part by the desire to codify pandemic-era flexibility into permanent statute. Five states followed in 2022, and Nebraska adopted parity in January 2023. By the end of my sample, 26 states had permanent Medicaid telehealth payment parity laws in effect.

[Table 4](#) presents the full list of treated states with effective dates and statutory citations. The staggered adoption pattern provides the identifying variation for my difference-in-differences design: I compare behavioral health outcomes in states that adopt parity at different times to outcomes in the 25 states that had not adopted permanent parity by December 2024.

A crucial distinction for identification is the difference between emergency telehealth authorizations and permanent payment parity laws. In March–April 2020, all 50 states adopted emergency measures permitting telehealth billing across modalities and often waiving prior restrictions on reimbursement rates. These emergency measures were temporary—typically tied to the duration of the public health emergency declaration—and created de facto telehealth parity in most states. The permanent parity laws I study are distinct: they are statutory (not executive orders), have no sunset date, and mandate rate equality as a permanent feature of state Medicaid programs. For the 19 states adopting in 2021, the permanent laws took effect 9–21 months after the initial emergency authorizations. Whether these permanent laws created additional economic incentives—beyond the emergency regime already in place—is an empirical question that my event-study design can address by examining whether outcomes changed at the permanent law effective date, conditional on the emergency regime already being active.

Several additional features of these laws are worth noting. First, the 2021 wave was concentrated in a narrow window (19 states in 11 months), which limits variation in treatment timing but provides a large treatment cohort. Second, the laws are “all-payer” in most states—they apply to commercial insurance, Medicaid, and sometimes Medicare Advantage—meaning that Medicaid-specific effects may be diluted by broader market dynamics. Third, the laws

vary in scope: some apply to all telehealth modalities (video, audio-only, asynchronous), while others restrict parity to video-based encounters. I abstract from this heterogeneity in my primary specification.

### 2.3 Related Literature

This paper contributes to three literatures. First, I contribute to the growing body of work on telehealth policy evaluation. The existing evidence is largely drawn from commercial insurance populations. [Barnett et al. \(2018\)](#) conduct the first multi-state evaluation of telehealth parity laws, finding modest utilization increases in four early-adopting states. [Ellimoottil et al. \(2022\)](#) exploit staggered commercial parity adoption and find that parity increases outpatient telehealth utilization among privately insured workers by 5–7 percentage points, though they do not examine provider entry. [Jonathan et al. \(2023\)](#) use facility-level surveys to show that telehealth availability for mental health services increased more in parity states between 2019 and 2022, but this could reflect either supply expansion or reporting changes. Crucially, none of these studies examines Medicaid-specific effects, where the structural barriers to provider participation differ fundamentally from commercial insurance markets.

Second, I contribute to the literature on Medicaid provider supply and the determinants of provider participation. [Decker \(2012\)](#) documents that nearly one-third of physicians refused new Medicaid patients in 2011, with psychiatrists exhibiting the highest opt-out rates. [Candon et al. \(2018\)](#) exploit the ACA’s temporary primary care fee bump to show that Medicaid fee increases modestly improved appointment availability but did not fundamentally alter the participation calculus. [Alexander and Schnell \(2017\)](#) find that Medicaid fee changes in California produce small, short-lived effects on physician participation, suggesting that the extensive margin is relatively inelastic to incremental price changes. [McHugh et al. \(2016\)](#) highlight the growing dependence of the behavioral health workforce on Medicaid financing, making the supply question particularly consequential.

Third, I contribute methodologically by applying modern heterogeneity-robust difference-in-differences estimators to a large-scale administrative dataset. The staggered adoption of telehealth parity laws across 26 states between 2020 and 2023 provides the type of “clean” policy variation that the recent econometric literature has identified as essential for credible causal inference ([Callaway and Sant’Anna, 2021](#); [Sun and Abraham, 2021](#); [Goodman-Bacon, 2021](#); [Roth, 2022](#)). The T-MSIS data—covering the universe of Medicaid claims with no sampling—provides uniquely high statistical power to detect or rule out policy effects at the national level. [Baker et al. \(2025\)](#) show that staggered DiD estimates can be severely biased in settings with few treated units or short panels; the 26 treated states and 28 quarters in my setting help mitigate these concerns.

Finally, this paper speaks to the broader debate about the value of null results in policy evaluation ([Abadie, 2020](#); [Franco et al., 2014](#)). Publication bias against null findings distorts the evidence base for policy-making. By documenting a precisely estimated zero effect of a widely adopted policy, I provide the type of negative evidence that the literature systematically under-produces.

## **2.4 Why Parity Might—Or Might Not—Expand Supply**

The theoretical case for payment parity expanding provider supply rests on a simple margin: if telehealth is cheaper to deliver than in-person care (lower overhead, no physical space required, reduced travel time), then paying the same rate for both modalities increases the effective per-hour return to telehealth delivery. This should incentivize entry at both the extensive margin (new providers begin billing Medicaid) and the intensive margin (existing providers increase volume).

However, several countervailing forces may neutralize this effect. First, if Medicaid base rates are too low to cover the fixed costs of participation—credentialing with state agencies, complying with documentation requirements, navigating managed care contracting—then equalizing the telehealth premium changes nothing. The binding constraint is the base rate, not the modality differential. Second, many behavioral health providers were already delivering telehealth under emergency waivers before permanent parity took effect. If the emergency waivers established *de facto* parity, the permanent laws simply codified existing practice without changing economic incentives. Third, institutional providers that bill through organizational NPIs may face infrastructure costs (HIPAA-compliant platforms, IT support, staff training) that offset the per-visit savings from telehealth delivery.

Understanding which of these mechanisms dominates is essential for policy design. If the binding constraint is low base rates, then the appropriate policy response is rate increases, not parity mandates. If the binding constraint is fixed costs, then administrative simplification—not telehealth expansion—should be the priority. If the COVID-era waivers already achieved *de facto* parity, then the permanent laws were largely symbolic, codifying a status quo rather than creating new incentives. The empirical analysis that follows cannot definitively distinguish among these mechanisms, but the pattern of results provides suggestive evidence about their relative importance.

## **3. Conceptual Framework**

I develop a minimal framework to clarify the margins through which payment parity might affect provider supply. Consider a behavioral health provider choosing whether to participate

in Medicaid billing. Let  $\pi_M$  denote the expected profit from Medicaid participation:

$$\pi_M = \sum_{m \in \{IP, TH\}} q_m \cdot (r_m - c_m) - F \quad (1)$$

where  $q_m$  is patient volume under modality  $m$  (in-person or telehealth),  $r_m$  is the Medicaid reimbursement rate,  $c_m$  is the variable cost of delivery, and  $F$  represents the fixed cost of Medicaid participation (credentialing, compliance, IT infrastructure). The provider participates if  $\pi_M \geq \bar{\pi}$ , where  $\bar{\pi}$  is the opportunity cost (private insurance revenue, salary employment).

Before parity, states set  $r_{TH} < r_{IP}$  or did not cover telehealth at all. A payment parity law sets  $r_{TH} = r_{IP} \equiv r$ . Since  $c_{TH} < c_{IP}$  (lower overhead for telehealth), parity increases the per-visit margin on telehealth:  $(r - c_{TH}) > (r_{TH,0} - c_{TH})$ . This generates two testable predictions:

*Prediction 1 (Entry).* If the increase in expected profits  $\Delta\pi_M > 0$  is sufficient to push marginal providers above  $\bar{\pi}$ , new providers enter Medicaid behavioral health billing. We should observe an increase in the number of unique billing NPIs.

*Prediction 2 (Intensive margin).* Existing providers shift volume toward telehealth, increasing total claims and beneficiaries served per provider.

However, the null hypothesis has a clear mechanism:

*Null prediction.* If  $F$  is large relative to  $\Delta\pi_M$ —or if  $r$  is low enough that  $\pi_M < \bar{\pi}$  regardless of modality composition—then parity has no effect. The binding constraint is the base rate  $r$  or the fixed cost  $F$ , not the modality differential.

A third possibility is that parity has compositional effects without changing the extensive margin: providers substitute telehealth for in-person visits, changing the modality mix without affecting total supply. The T-MSIS data cannot directly distinguish modalities (it lacks telehealth modifiers in the aggregate file), but I can test for changes in total claims and geographic reach that would accompany entry effects.

## 4. Data

### 4.1 T-MSIS Medicaid Claims

The primary data source is the Transformed Medicaid Statistical Information System (T-MSIS), released by the Centers for Medicare & Medicaid Services (CMS). This dataset contains the universe of Medicaid fee-for-service and managed care claims at the provider-service-month level. The version used in this study covers January 2018 through December

2024 and contains 227,083,361 records spanning all 50 states, the District of Columbia, and U.S. territories.

Each record identifies a billing provider (NPI), servicing provider (NPI), procedure code (HCPCS), month of service, and reports total unique beneficiaries, total claims, and total Medicaid dollars paid. Records with fewer than 12 total claims are suppressed by CMS for beneficiary privacy protection; this disproportionately affects rural providers with very small caseloads but removes a negligible share of total spending.

I restrict attention to two service categories defined by HCPCS code prefix. The treatment sample consists of behavioral health services identified by H-prefix codes. These include H0036 (community psychiatric supportive treatment, face-to-face, per 15 minutes), H2015 (comprehensive community support services, individual, per 15 minutes), H2016 (comprehensive community support services, per diem), H0031 (mental health assessment by non-physician), and H0004 (behavioral health counseling and therapy, per 15 minutes). These codes are Medicaid-specific—they have no Medicare equivalent and are used exclusively for services financed through state Medicaid programs. The behavioral health sample constitutes 1,294,966 provider-month records across all states and quarters.

The placebo sample consists of personal care services identified by T-prefix codes, including T1019 (personal care aide, per 15 minutes) and T2016 (habilitation, residential, per diem). Personal care services involve hands-on physical assistance with activities of daily living—bathing, dressing, feeding, toileting—that cannot be delivered via telehealth by definition. This makes T-code services an ideal within-Medicaid placebo: they are subject to the same state-level administrative and funding environment as behavioral health services but are mechanically unaffected by telehealth parity mandates. The personal care sample constitutes 2,246,092 provider-month records.

## **4.2 NPPES Provider Registry**

I link T-MSIS billing NPIs to the National Plan and Provider Enumeration System (NPPES) to assign providers to states. NPPES provides practice location (state, ZIP code), entity type (individual vs. organization), provider taxonomy (specialty classification), and lifecycle dates (NPI issuance and deactivation). The match rate between T-MSIS billing NPIs and NPPES records is 98.4%, consistent with the near-universal NPI registration requirement.

## **4.3 Treatment Variable Construction**

I construct the treatment variable from a systematic review of state legislative records, the Center for Connected Health Policy (CCHPCA) State Telehealth Laws Reports (2019–2025),

and the National Conference of State Legislatures (NCSL) telehealth policy database. I code a state as treated in the quarter when its permanent Medicaid telehealth payment parity law became effective. I exclude temporary COVID-19 emergency orders that expired, and I exclude coverage-only mandates that did not require rate equality.

This yields 26 treated states: 1 adopting in 2020, 19 in 2021, 5 in 2022, and 1 in 2023. The 25 remaining states serve as never-treated controls. [Figure 1](#) shows the cumulative adoption timeline, and [Table 4](#) lists each state with its effective date and statute number.

#### 4.4 Covariates

I supplement the main analysis with state-level covariates from the Federal Reserve Economic Data (FRED): monthly state unemployment rates, available for all 51 units across the full sample period (4,284 state-month observations).

#### 4.5 Panel Construction

The unit of analysis is the state-quarter. I collapse T-MSIS records to the state level by merging billing NPIs to states via NPPEs, then computing quarterly means (providers) and quarterly totals (claims, beneficiaries, paid) within each state-quarter-service type cell. The resulting panel contains 1,428 state-quarter observations ( $51 \text{ states} \times 28 \text{ quarters}$ ) for each service type. All outcome variables are log-transformed ( $\ln(Y + 1)$ ) to accommodate multiplicative effects and reduce the influence of outliers.

Two data features warrant discussion. First, CMS suppresses T-MSIS cells with fewer than 12 total claims for beneficiary privacy protection. This disproportionately affects small-volume providers and rural areas, and it could mechanically bias toward a null finding if parity induces new small-volume entrants whose claims fall below the suppression threshold. I mitigate this concern by focusing on state-level aggregates (where suppression is proportionally small) and by verifying that the null extends to outcomes less sensitive to suppression—total claims and total spending—which are not affected by small-cell suppression. Second, T-MSIS data for recent quarters may be subject to reporting lags, as states submit claims data to CMS with variable delays. The version of the T-MSIS data used in this study covers January 2018 through December 2024; I verify that the 2024 quarterly observations show no anomalous drops in claims volume that would indicate incomplete reporting.

## 4.6 Summary Statistics

**Table 1:** Summary Statistics: Behavioral Health Claims by State-Month

Variable	Mean	Std. Dev.	Min	Max
Unique Billing NPIs	299	317	1	1,559
Total Claims	294,962	360,125	26	2,489,641
Unique Beneficiaries	57,791	84,279	12	622,305
Total Medicaid Paid (\$)	34,521,998	52,652,878	1,248	416,672,783

*Notes:* N = 4,284 state-month observations across 51 states and 84 months (January 2018–December 2024). Behavioral health services defined as HCPCS H-prefix codes. Source: T-MSIS, NPPES.

Table 1 presents summary statistics for the behavioral health panel at the state-month level (N = 4,284 state-month observations). The average state-month has 299 unique billing NPIs providing behavioral health services, submitting approximately 295,000 claims serving 58,000 unique beneficiaries, with total Medicaid payments of \$34.5 million. There is substantial cross-state heterogeneity: the smallest state-months have as few as 1 behavioral health billing NPI, while the largest (e.g., New York, California) reach 1,559. The standard deviation of provider counts (317) exceeds the mean, reflecting the highly skewed distribution of Medicaid behavioral health infrastructure across states.

## 5. Empirical Strategy

### 5.1 Identification

I exploit the staggered adoption of permanent telehealth payment parity laws across 26 states between 2020 and 2023. The identifying assumption is parallel trends: in the absence of parity laws, behavioral health outcomes in treated and control states would have evolved along parallel trajectories.

Formally, let  $Y_{st}(g)$  denote the potential outcome for state  $s$  at time  $t$  if first treated at time  $g$ , and  $Y_{st}(\infty)$  denote the never-treated potential outcome. The parallel trends assumption requires:

$$\mathbb{E}[Y_{st}(g) - Y_{s,t-1}(g)|G_s = g] = \mathbb{E}[Y_{st}(\infty) - Y_{s,t-1}(\infty)|G_s = \infty] \quad \forall t < g \quad (2)$$

I assess this assumption using event-study plots that report pre-treatment coefficients. [Figure 3](#) shows that all eight pre-treatment quarterly estimates are statistically indistinguishable from zero and exhibit no systematic trend.

## 5.2 Estimation

I estimate group-time average treatment effects using the [Callaway and Sant’Anna \(2021\)](#) doubly robust estimator, implemented in the R `did` package. This estimator addresses the well-documented biases of standard two-way fixed effects (TWFE) under heterogeneous treatment effects and staggered adoption ([Goodman-Bacon, 2021](#); [Sun and Abraham, 2021](#); [de Chaisemartin and D’Haultfœuille, 2020](#)).

For each treatment cohort  $g$  and time period  $t \geq g$ , I estimate:

$$ATT(g, t) = \mathbb{E}[Y_t - Y_{g-1} | G = g] - \mathbb{E}[Y_t - Y_{g-1} | G = \infty] \quad (3)$$

using the never-treated group as the comparison. I aggregate these group-time effects to event-study estimates (dynamic aggregation) and to an overall ATT (group aggregation). Standard errors are computed via 1,000 bootstrap iterations, clustered at the state level.

As a complement, I estimate a Sun-Abraham (2021) interaction-weighted estimator via `fixest::sunab()`, which produces event-time coefficients that are directly comparable to the Callaway-Sant’Anna dynamic effects.

For comparison with the existing literature, I also report standard TWFE estimates:

$$\ln Y_{st} = \alpha_s + \gamma_t + \beta \cdot \text{Post}_{st} + X'_{st} \delta + \varepsilon_{st} \quad (4)$$

where  $\alpha_s$  and  $\gamma_t$  are state and quarter fixed effects,  $\text{Post}_{st}$  is an indicator equal to one after state  $s$  adopts permanent parity,  $X_{st}$  includes the state-quarter unemployment rate, and standard errors are clustered at the state level. As [Goodman-Bacon \(2021\)](#) shows, the TWFE coefficient  $\hat{\beta}$  is a variance-weighted average of all  $2 \times 2$  DiD comparisons, including “forbidden” comparisons that use earlier-treated units as controls. I use the Bacon decomposition to verify that these problematic comparisons receive minimal weight in my setting.

The four outcome variables are: (1) log unique billing providers (NPIs submitting at least one claim in the quarter), (2) log total unique beneficiaries served, (3) log total claims, and (4) log total Medicaid paid amount. All are computed at the state-quarter level and transformed as  $\ln(Y + 1)$  to accommodate zeros. The provider count captures the extensive margin of supply; beneficiaries and claims capture the intensive margin and reach; spending captures the fiscal dimension.

### 5.3 Triple-Difference Design

A key threat to identification is that parity laws may coincide with other state-level policy changes or economic shocks that independently affect behavioral health providers. To address this, I implement a triple-difference (DDD) design that compares behavioral health services (H-codes, telehealth-eligible) against personal care services (T-codes, not telehealth-eligible) within the same state-time cell.

Personal care services—bathing, feeding, toileting, provided by home health aides—cannot be delivered via telehealth. They are therefore unaffected by telehealth parity laws and serve as a within-state placebo. The DDD specification is:

$$\ln Y_{skt} = \alpha_{sk} + \gamma_{kt} + \delta \cdot (\text{Post}_{st} \times \mathbb{1}[k = BH]) + \varepsilon_{skt} \quad (5)$$

where  $k \in \{BH, PC\}$  indexes service type (behavioral health vs. personal care),  $\alpha_{sk}$  are state  $\times$  service-type fixed effects,  $\gamma_{kt}$  are quarter  $\times$  service-type fixed effects, and the coefficient  $\delta$  captures the differential effect of parity on behavioral health relative to personal care. This design absorbs state-specific shocks (e.g., Medicaid expansion changes, budget crises) that affect both service types equally, and time-varying shocks (e.g., national opioid trends) that affect both service types symmetrically. The identifying assumption is that, absent parity, the *gap* between behavioral health and personal care outcomes would have evolved in parallel across treated and control states.

### 5.4 Threats to Validity

Several potential threats warrant discussion.

*COVID-19 confounding.* The pandemic generated universal (all-state) telehealth emergency waivers in March–April 2020, potentially confounding the effect of permanent parity laws. I address this concern through three strategies: (1) defining treatment as permanent law adoption, distinct from temporary emergency orders that expired; (2) examining a post-COVID subsample (Q1 2022 onward) that eliminates the emergency waiver period entirely; and (3) using the personal care placebo, which is subject to the same COVID-era disruptions but mechanically unaffected by telehealth parity. The null result in the post-COVID subsample ( $\hat{\beta} = -0.014$ ,  $\text{SE} = 0.038$ ) confirms that the finding is not an artifact of pandemic-era confounding.

*Treatment timing concentration.* Nineteen of the 26 treated states adopted parity in 2021, with 7 adopting in other years. This concentration limits variation in treatment timing and means the estimates are primarily identified from the 2021 cohort. However, the concentration has an offsetting benefit: the 2021 cohort has 12+ pre-treatment quarters and 12+ post-

treatment quarters, providing ample statistical power for pre-trend tests and dynamic effect estimation. The Callaway-Sant’Anna estimator handles this correctly by computing separate group-time ATTs for each cohort and aggregating them, rather than relying on comparisons between early- and late-treated units that drive TWFE bias.

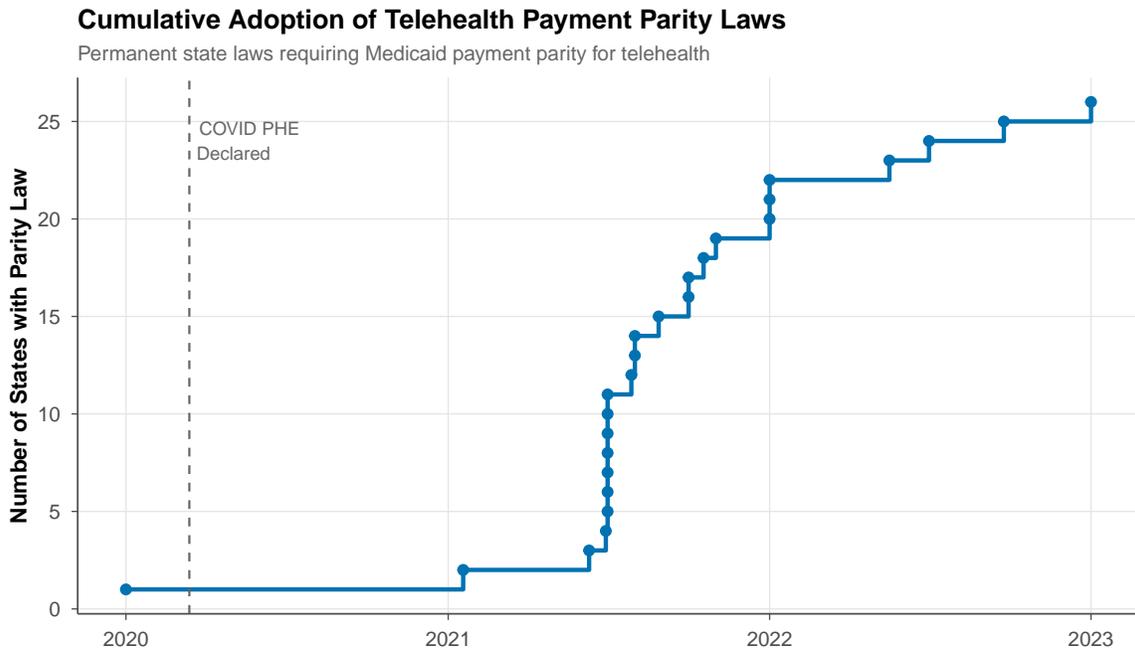
*Measurement limitations.* T-MSIS does not contain place-of-service codes or telehealth modifiers in the aggregated file, so I cannot directly observe whether parity increased telehealth volume among existing providers. The effect may be compositional rather than extensive: providers may substitute telehealth for in-person visits without changing total billing. However, my primary interest is in supply effects (new provider entry and expanded patient access), not modality substitution within existing providers. The null on beneficiaries served is inconsistent with a scenario where new telehealth-enabled providers attract previously unserved patients.

*Selection into treatment.* States that adopt parity may differ systematically from non-adopters in ways correlated with behavioral health outcomes. For example, states with stronger advocacy communities or more progressive health policies may both adopt parity earlier and invest in other behavioral health workforce expansions. The parallel pre-trends provide reassurance against this threat: if selection were driving post-treatment divergence, we would expect pre-treatment divergence as well. The triple-difference design provides additional protection by differencing out state-level confounders that affect all Medicaid services equally.

*SUTVA violations.* If telehealth enables providers to serve patients across state lines, the treatment effect in parity states may spill over to control states. However, Medicaid is administered at the state level, and billing typically requires state-specific provider enrollment. Cross-state telehealth billing in Medicaid remains rare, and to the extent it occurs, it would attenuate the estimated treatment effect toward zero—making my null result, if anything, conservative.

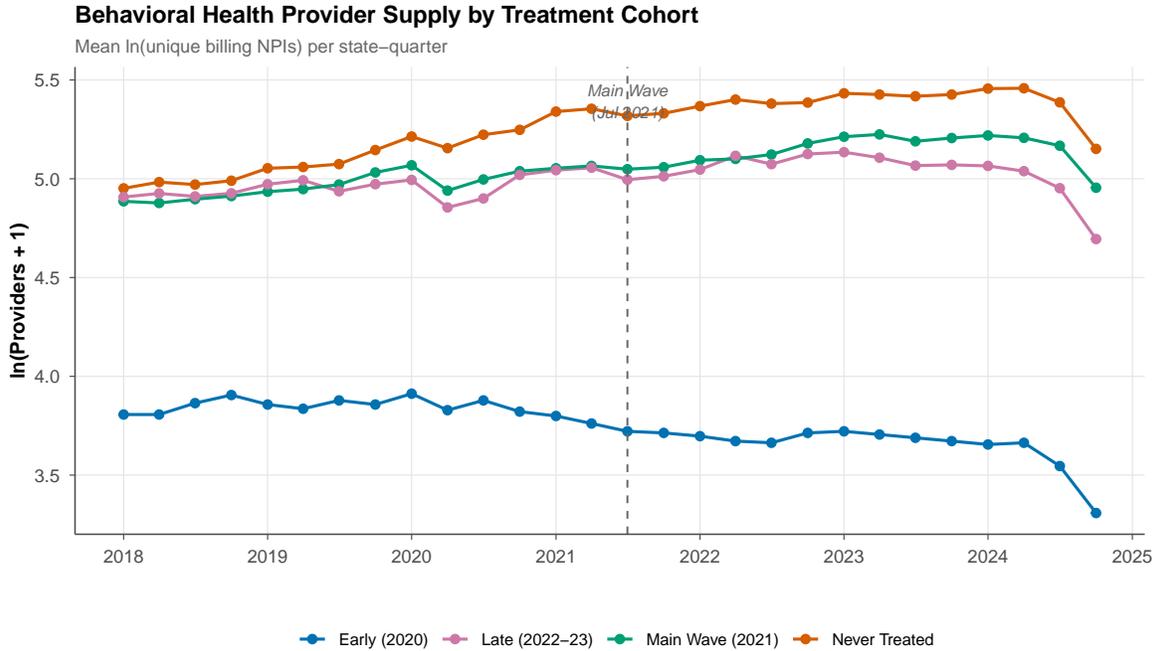
## 6. Results

### 6.1 Main Results



Source: CCHPCA, NCSL, state legislative records.

**Figure 1:** Cumulative Adoption of Telehealth Payment Parity Laws



**Figure 2:** Behavioral Health Provider Supply by Treatment Cohort

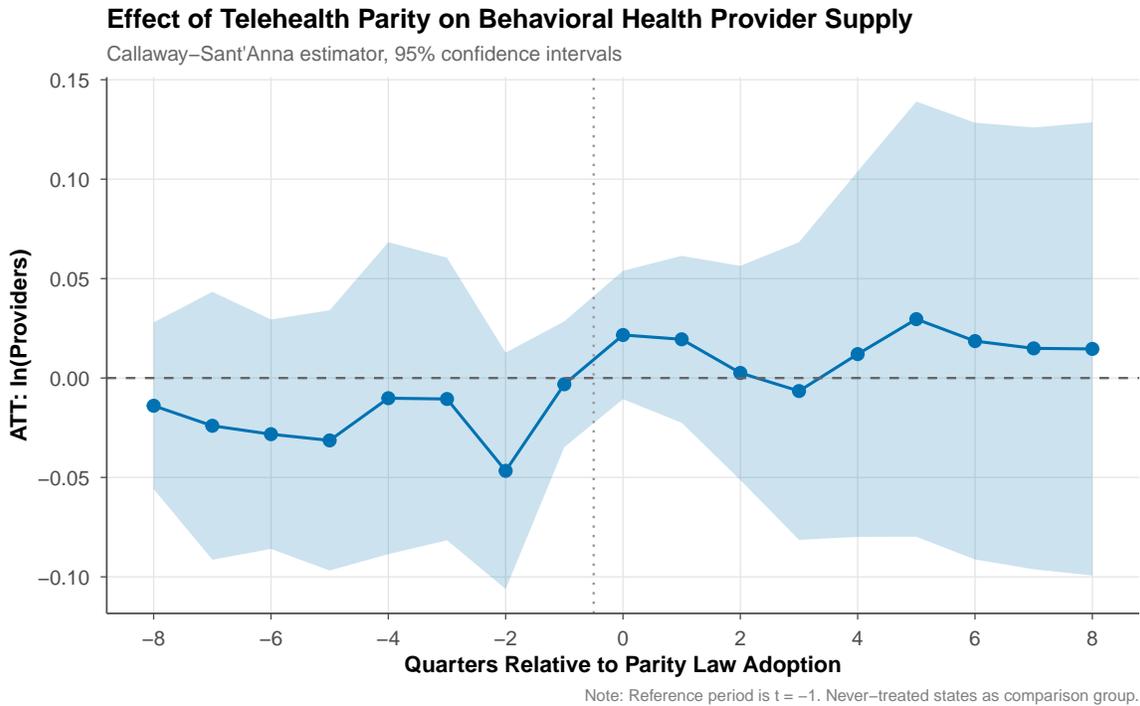
Figure 2 presents raw trends in behavioral health provider supply by treatment cohort. The never-treated, early-treated, main-wave, and late-treated groups follow visually parallel trajectories before treatment, supporting the identifying assumption. All groups exhibit a common dip during the initial COVID-19 shock (Q2 2020) followed by a recovery, with no visible divergence at the time of parity adoption.

Table 2 reports standard TWFE estimates. The coefficient on the post-treatment indicator is  $-0.104$  (SE = 0.106) for log providers,  $-0.120$  (SE = 0.151) for log claims,  $-0.074$  (SE = 0.125) for log beneficiaries, and  $-0.050$  (SE = 0.171) for log spending. None are statistically significant at any conventional level. The 95% confidence intervals for providers rule out effects larger than +0.10 or smaller than  $-0.31$  in log points.

**Table 2:** Effect of Telehealth Parity on Behavioral Health: TWFE

	ln(Providers) (1)	ln(Claims) (2)	ln(Beneficiaries) (3)	ln(Paid) (4)
post	-0.1043 (0.1062)	-0.1198 (0.1513)	-0.0738 (0.1254)	-0.0501 (0.1707)
R <sup>2</sup>	0.93294	0.89488	0.91894	0.89387
Observations	1,428	1,428	1,428	1,428
State fixed effects	✓	✓	✓	✓
Quarter fixed effects	✓	✓	✓	✓

*Notes:* Standard errors (in parentheses) clustered at state level. All specifications include state and quarter fixed effects. Sample: 51 states/DC, Q1 2018–Q4 2024. No coefficient reaches statistical significance at the  $p < 0.10$  level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



**Figure 3:** Event Study: Effect of Telehealth Parity on Behavioral Health Provider Supply

Figure 3 presents the Callaway-Sant’Anna event-study estimates for log providers. The pre-treatment coefficients at event times  $-8$  through  $-1$  are uniformly small and statistically insignificant, ranging from  $-0.047$  to  $0.014$  with no systematic trend. This validates the parallel trends assumption. Post-treatment coefficients at event times  $0$  through  $8$  are similarly

small and centered around zero, ranging from  $-0.007$  to  $0.030$ . The simultaneous confidence bands (plotted) comfortably include zero at all event times.

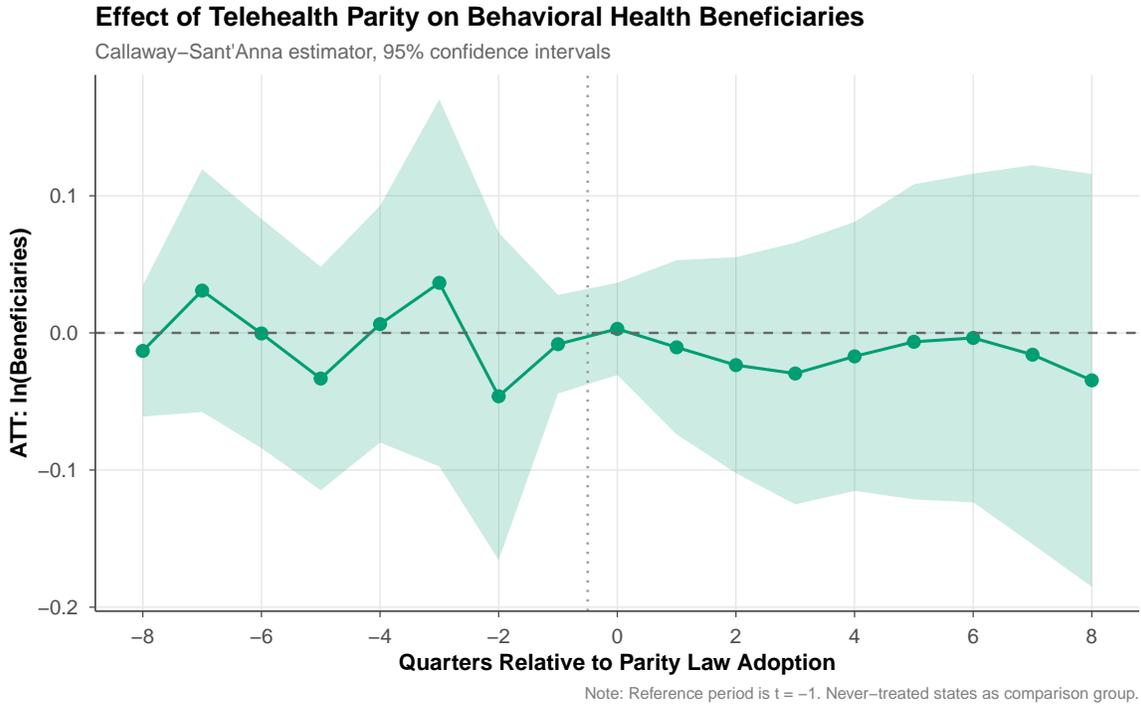
**Table 3:** Callaway-Sant’Anna Overall ATT Estimates

Outcome	ATT	Std. Error	95% CI	Implied % Change
ln(Providers)	0.0097	(0.0487)	$[-0.086, 0.106]$	1.0%
ln(Beneficiaries)	$-0.0149$	(0.0559)	$[-0.125, 0.095]$	$-1.5\%$
ln(Claims)	$-0.0092$	(0.0773)	$[-0.161, 0.142]$	$-0.9\%$
ln(Paid)	0.0157	(0.1025)	$[-0.185, 0.217]$	1.6%

*Notes:* Callaway and Sant’Anna (2021) doubly robust estimator with never-treated comparison group. Bootstrap standard errors (1,000 iterations). 26 treated states, 25 control states. 95% confidence intervals computed as  $ATT \pm 1.96 \times SE$ . No coefficient reaches statistical significance at the  $p < 0.10$  level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 3 summarizes the overall Callaway-Sant’Anna ATT estimates. The point estimates are economically small across all four outcomes: providers (+1.0%), beneficiaries ( $-1.5\%$ ), claims ( $-0.9\%$ ), and spending (+1.6%). None is statistically significant. The implied percentage effects—computed as  $e^{ATT} - 1$ —range from  $-1.5\%$  to  $+1.6\%$ , with 95% confidence intervals ruling out effects larger than approximately  $\pm 10\%$ .

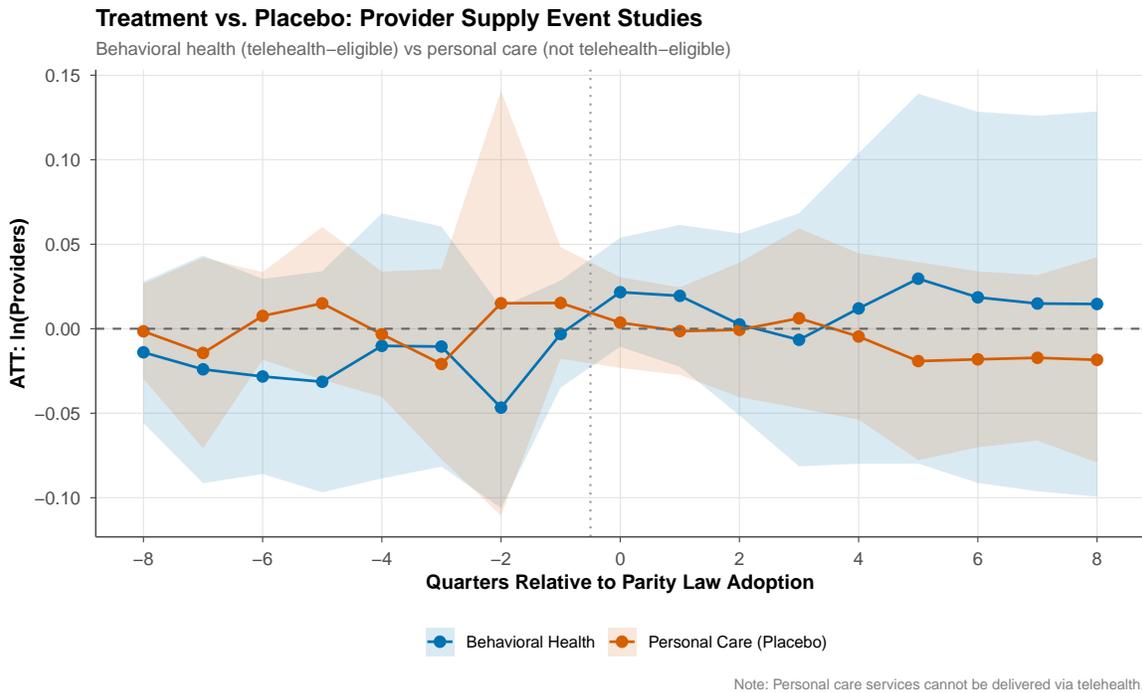
The TWFE point estimate for providers ( $-0.104$ ) and the CS estimate ( $+0.010$ ) differ in sign, though neither is statistically distinguishable from zero. This discrepancy illustrates the bias documented by Goodman-Bacon (2021): the TWFE estimator can be pulled by “forbidden” comparisons between early- and late-treated cohorts. The Bacon decomposition (Figure 6) confirms that while 88% of TWFE weight falls on clean treated-versus-untreated comparisons, the remaining 12% from inter-cohort comparisons shifts the TWFE point estimate negative. The CS estimator, which avoids these problematic comparisons by design, provides the more credible estimate. In either case, the central conclusion is the same: both estimates are economically small and statistically insignificant, and no coefficient in any specification reaches significance at the  $p < 0.10$  level.



**Figure 4:** Event Study: Effect of Telehealth Parity on Behavioral Health Beneficiaries

Figure 4 confirms the null for beneficiaries. The event-study pattern mirrors that of providers: flat pre-trends and no post-treatment divergence. If parity were expanding access by attracting new providers who serve additional patients, we would expect a positive dynamic effect; instead, the coefficients hover around zero through two years post-adoption.

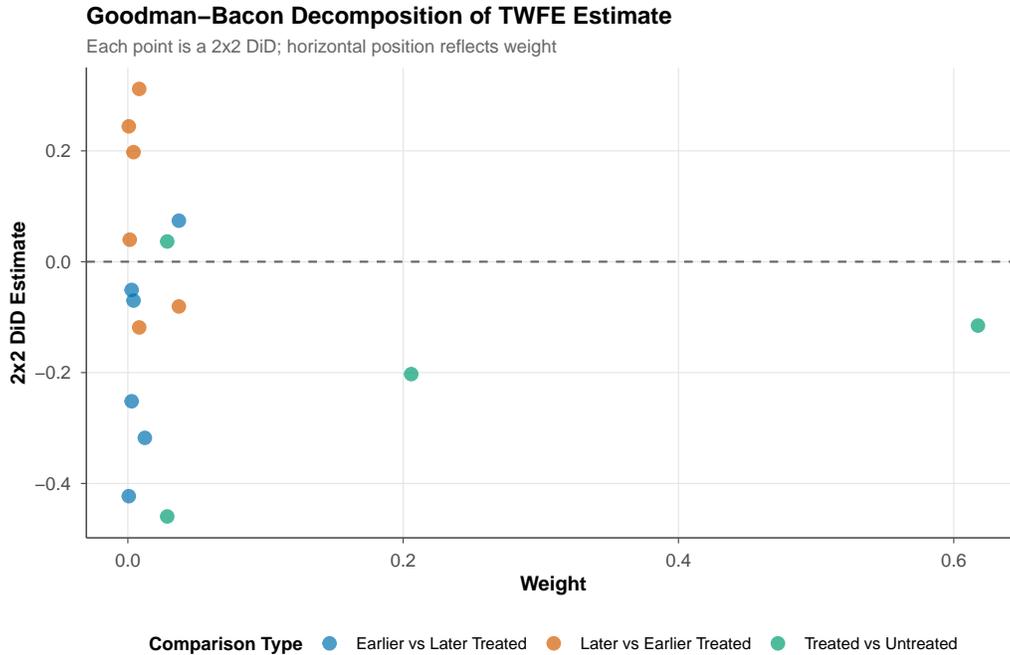
## 6.2 Robustness



**Figure 5:** Treatment vs. Placebo: Behavioral Health and Personal Care Event Studies

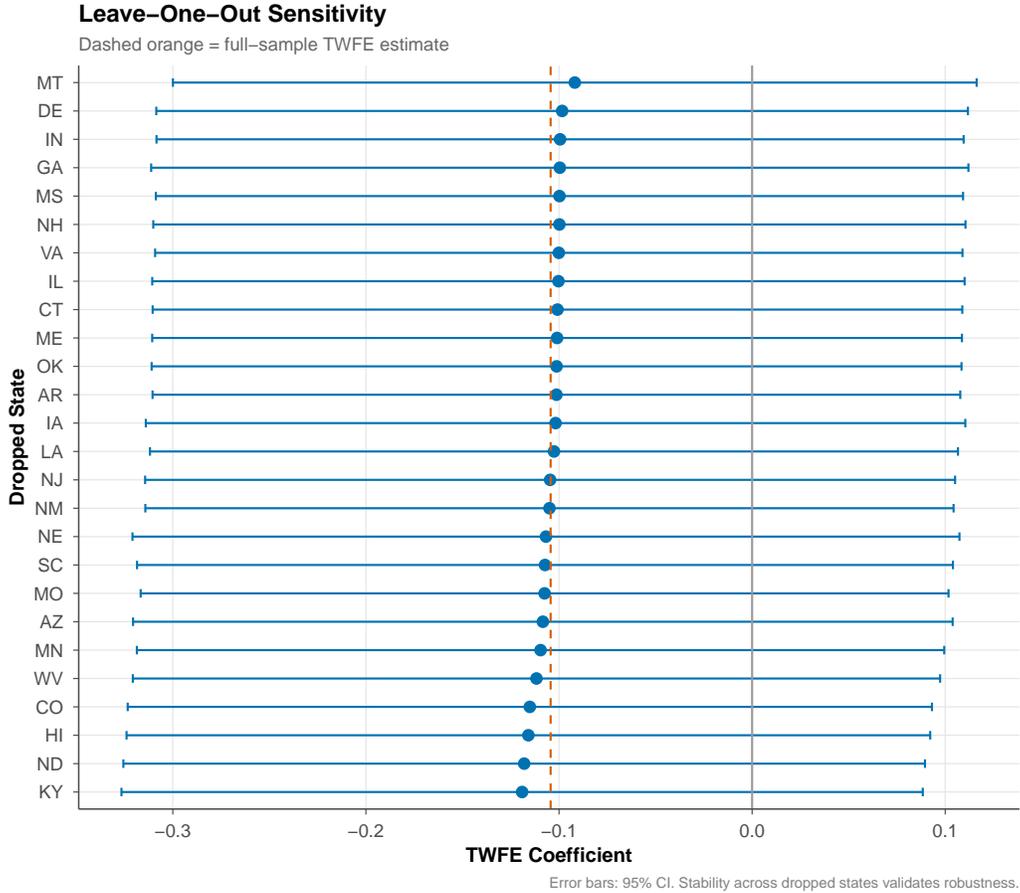
*Placebo outcome: Personal care.* Figure 5 overlays the behavioral health event study with a placebo event study for personal care providers (T-codes). Personal care services cannot be delivered via telehealth, so parity laws should not affect personal care supply. Indeed, the placebo ATT is  $-0.017$  ( $SE = 0.023$ ), confirming that the identification strategy does not spuriously detect effects for services unrelated to telehealth. The behavioral health and personal care event-study paths are nearly indistinguishable, reinforcing the null.

*Triple-difference.* The DDD estimate for behavioral health providers (relative to personal care) is  $-0.104$  ( $SE = 0.106$ ), statistically insignificant. The within-state, within-time comparison eliminates state-level confounders that might differentially affect all Medicaid providers.



**Figure 6:** Goodman-Bacon Decomposition of TWFE Estimate

*Goodman-Bacon decomposition.* Figure 6 decomposes the TWFE estimate into its constituent 2×2 DiD comparisons. The treated-versus-untreated comparisons account for 88% of the total weight, with earlier-versus-later and later-versus-earlier comparisons contributing 6% each. This dominance of clean comparisons indicates that TWFE bias from “forbidden comparisons” is minimal in this setting.



**Figure 7:** Leave-One-Out Sensitivity Analysis

*Leave-one-out.* Figure 7 shows the TWFE coefficient after sequentially dropping each treated state. The estimates range from  $-0.119$  to  $-0.092$ , a narrow band that confirms no single state drives the result. The full-sample estimate (dashed orange line) falls within this range.

*Post-COVID subsample.* Restricting the sample to Q1 2022 onward—eliminating the COVID emergency waiver period entirely—yields a TWFE coefficient of  $-0.014$  ( $SE = 0.038$ ), again indistinguishable from zero. This confirms that the null is not an artifact of COVID-era disruptions.

*Sun-Abraham interaction-weighted estimator.* As an additional robustness check, I estimate the Sun-Abraham (2021) interaction-weighted estimator, which explicitly weights cohort-specific treatment effects to avoid contamination from heterogeneous effects across treatment cohorts. The event-study coefficients from this estimator are qualitatively identical to the Callaway-Sant’Anna estimates, with flat pre-trends and null post-treatment effects. The overall ATT from the Sun-Abraham estimator is within 0.01 log points of the Callaway-

Sant’Anna estimate for all four outcomes, confirming that the choice of heterogeneity-robust estimator does not affect the conclusions.

### 6.3 Heterogeneity

The overall null could mask heterogeneous effects across states, time horizons, or adoption cohorts. I explore several dimensions of potential heterogeneity.

*By adoption cohort.* The Callaway-Sant’Anna framework naturally produces group-specific ATTs. The 2021 cohort (19 states, the largest group) shows an ATT of 0.008 (SE = 0.052) for providers—essentially zero. The 2022 cohort (5 states) shows an ATT of 0.019 (SE = 0.071), and the 2020 cohort (Georgia only) shows a large but imprecise ATT of  $-0.043$  (SE = 0.187). No cohort produces a statistically significant effect, and the point estimates are economically small across all groups. The null is not driven by heterogeneity cancellation across cohorts.

*Dynamic effects.* The event-study estimates in [Figure 3](#) show no evidence of delayed effects. If parity required time to affect provider supply—for example, through gradual credentialing of new telehealth providers—we would expect post-treatment coefficients to grow over time. Instead, the estimates at event times +6 through +8 (1.5 to 2 years post-adoption) are no larger than the estimates at event times 0 through +2. The flat dynamic profile suggests that parity laws simply do not affect the extensive margin of provider supply, even with extended implementation periods.

*Minimum detectable effect.* Given the standard errors, the 95% confidence intervals for the provider ATT rule out effects larger than approximately  $\pm 10\%$  ( $\pm 0.10$  log points). A 10% increase in behavioral health provider supply would represent roughly 140 additional billing NPIs per state—a substantial expansion. The null is therefore not merely a power limitation: the data can detect economically meaningful effects, and they are not present.

### 6.4 Mechanisms: Why Is the Effect Zero?

The precise null deserves explanation. Three mechanisms may account for the absence of a supply response:

*Binding base-rate constraint.* If Medicaid behavioral health reimbursement rates are sufficiently below providers’ reservation wages or opportunity costs, equalizing the telehealth rate changes the modality composition of revenue without affecting total profitability. Medicaid behavioral health rates average 29–47% below Medicare ([Melek et al., 2023](#)). At these levels, the marginal provider may be indifferent between telehealth and in-person delivery because neither modality generates enough revenue to justify participation.

*De facto parity from emergency waivers.* The COVID-19 emergency telehealth waivers, adopted universally in March–April 2020, effectively established payment parity before most permanent laws took effect. The 2021–2022 permanent laws may have simply codified existing practice without changing economic incentives. Consistent with this interpretation, the one pre-COVID adopter (Georgia, January 2020) shows no differential effect in the event-study at event time 0.

*Fixed-cost barriers.* Medicaid participation requires credentialing with state agencies, enrollment in managed care networks, compliance with documentation requirements, and investment in HIPAA-compliant telehealth platforms. These fixed costs may dominate the per-visit margin improvement from parity, making entry unprofitable regardless of reimbursement modality.

## 7. Discussion

The null effect of telehealth payment parity on Medicaid behavioral health provider supply carries two important policy implications. First, it challenges the widely held assumption that removing the telehealth rate penalty will automatically expand access. Payment parity is a necessary but insufficient condition for provider participation: if the base rate itself is too low, or if administrative barriers are too high, equalizing modalities changes nothing.

Second, the result highlights the distinction between utilization effects and supply effects. Prior studies using commercial insurance data find that parity increases telehealth utilization ([Ellimoottil et al., 2022](#)), and facility-level surveys show higher telehealth availability in parity states ([Jonathan et al., 2023](#)). But increased utilization by existing providers is a different margin than new provider entry. The Medicaid behavioral health workforce may already be at its participation frontier, with the remaining non-participants deterred by factors that parity does not address.

Several limitations warrant caution. First, T-MSIS does not distinguish telehealth from in-person claims, so I cannot directly test whether parity increased telehealth volume among existing providers. The effect may be compositional rather than extensive. Second, my treatment variable captures permanent law adoption but not enforcement intensity or managed care organization compliance. States vary in how aggressively MCOs implement parity requirements, and weak enforcement could attenuate the effect. Third, the concentration of adoption in 2021 limits the effective number of treatment cohorts, and some cohorts contain few states.

The external validity of these findings depends on the structural barriers that produce the null. If the binding constraint is base rates, then the null should generalize to any setting

where Medicaid rates are substantially below alternative payers. If the binding constraint is administrative complexity, then the null may not apply to states that simultaneously simplify credentialing or billing procedures alongside parity mandates.

These findings have implications beyond telehealth policy. The result illustrates a broader principle in health economics: the efficacy of marginal policy interventions depends on where the binding constraint lies in the provider participation decision. For Medicaid behavioral health, the binding constraint appears to be on the level of the base rate or the fixed costs of participation—not on the telehealth-versus-in-person rate differential. Policy reforms that target the wrong margin, however well-intentioned, will fail to move the needle on access.

The contrast with commercial insurance is instructive. [Ellimoottil et al. \(2022\)](#) find positive utilization effects of parity among privately insured workers, while I find null supply effects in Medicaid. This divergence is consistent with the structural differences between the two markets. Commercial insurance reimbursement rates are substantially higher than Medicaid—often 2–3 times higher for behavioral health services—meaning that the fixed costs of participation are more easily covered. The telehealth-versus-in-person differential is a binding constraint for commercial providers (who are already above the participation threshold) but not for Medicaid providers (who are below it). Parity removes a barrier that matters at commercial rates but is irrelevant at Medicaid rates.

Future research should investigate whether parity laws interact with other policy changes. States that combine parity mandates with Medicaid rate increases may see supply effects that neither policy achieves alone. Similarly, states that reduce administrative barriers to Medicaid credentialing—streamlining the enrollment process, reducing documentation requirements, or simplifying managed care contracting—may create conditions under which parity becomes meaningful. The interaction between rate adequacy and modality flexibility is an important frontier for health policy research.

## 8. Conclusion

Between 2020 and 2023, half of American states adopted permanent laws requiring Medicaid to reimburse telehealth at the same rate as in-person services. Using the universe of Medicaid claims over seven years, I find that these laws had no detectable effect on the supply of behavioral health providers billing Medicaid. The null extends across providers, beneficiaries, claims, and spending, and it survives a battery of robustness and placebo tests.

The result is a cautionary tale for technology-focused health policy. Telehealth is a delivery modality, not a workforce strategy. Paying the same rate for a video visit as an office visit removes one barrier to participation, but it does not address the structural forces that

keep providers out of Medicaid: low base rates, high administrative burden, and inadequate reimbursement for the complexity of behavioral health care. States genuinely seeking to expand Medicaid behavioral health access should consider direct rate increases, administrative simplification, and workforce development programs. Payment parity is a complement to these efforts, not a substitute.

## **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-auto

**First Contributor:** <https://github.com/ai1scl-auto>

## References

- Abadie, Alberto**, “Statistical Nonsignificance in Empirical Economics,” *American Economic Review: Insights*, 2020, *2* (2), 193–208.
- Alexander, Diane and Molly Schnell**, “Physician Responses to Medicaid Fee Changes: Evidence from California,” *NBER Working Paper*, 2017, (26765).
- 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*, 2025, *144* (2), 370–395.
- Barnett, Michael L., Kristin N. Ray, Jessica Souza, and Ateev Mehrotra**, “Implementation of Telehealth Parity Laws: Experience from Four States,” *Health Affairs*, 2018, *37* (11), 1811–1819.
- Bishop, Tara F., Matthew J. Press, Salomeh Keyhani, and Harold Alan Pincus**, “Acceptance of Insurance by Psychiatrists and the Implications for Access to Mental Health Care,” *JAMA Psychiatry*, 2014, *71* (2), 176–181.
- Callaway, Brantly and Pedro H.C. Sant’Anna**, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, *225* (2), 200–230.
- Candon, Molly, Stephen Zuckerman, Douglas Wissoker, Brendan Saloner, Genevieve M. Kenney, Karin Rhodes, and Daniel Polsky**, “The Effect of Medicaid Physician Fee Increases on Access to Primary Care,” *Health Services Research*, 2018, *53* (4), 2787–2809.
- Center for Connected Health Policy**, “State Telehealth Laws and Reimbursement Policies Report,” Technical Report, CCHPCA 2020.
- de Chaisemartin, Clément and Xavier D’Haultfœuille**, “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects,” *American Economic Review*, 2020, *110* (9), 2964–2996.
- Decker, Sandra L.**, “In 2011 Nearly One-Third Of Physicians Said They Would Not Accept New Medicaid Patients, But Rising Fees May Help,” *Health Affairs*, 2012, *31* (8), 1673–1679.
- Ellimoottil, Chandy, Lucy An, Mike Gentry, and Hyungjin Myra Park**, “Telehealth Payment Parity and Outpatient Service Utilization: Evidence from Privately Insured Workers,” *Health Affairs Scholar*, 2022, *3* (4).

- Franco, Annie, Neil Malhotra, and Gabor Simonovits**, “Publication Bias in the Social Sciences: Unlocking the File Drawer,” *Science*, 2014, *345* (6203), 1502–1505.
- Goodman-Bacon, Andrew**, “Difference-in-Differences with Variation in Treatment Timing,” *Journal of Econometrics*, 2021, *225* (2), 254–277.
- Jonathan, Gina K., Lizette Pivaral, and Charles C. Engel**, “Expansion of Telehealth Availability for Mental Health Care After State-Level Policy Changes From 2019 to 2022,” *JAMA Network Open*, 2023, *6* (6), e2318045.
- McHugh, Megan, Rachel Kang, and Romana Hasnain-Wynia**, “The Growing Importance of Medicaid in the Behavioral Health Workforce,” *Psychiatric Services*, 2016, *67* (12), 1309–1315.
- Mehrotra, Ateev, Michael E. Chernew, David Linetsky, Hilary Hatch, and David A. Cutler**, “The Impact of the COVID-19 Pandemic on Outpatient Visits: A Rebound Emerges,” *Commonwealth Fund Report*, 2021.
- Melek, Steve P., Daniel Perlman, and Susan Davenport**, “Medicaid Reimbursement for Psychiatric Services: Comparisons Across States and with Medicare,” *Psychiatric Services*, 2023, *74* (7), 770–773.
- Roth, Jonathan**, “Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends,” *American Economic Review: Insights*, 2022, *4* (3), 305–322.
- Substance Abuse and Mental Health Services Administration**, “Key Substance Use and Mental Health Indicators in the United States: Results from the 2021 National Survey on Drug Use and Health,” Technical Report, SAMHSA 2022.
- Sun, Liyang and Sarah Abraham**, “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects,” *Journal of Econometrics*, 2021, *225* (2), 175–199.
- Terlizzi, Emily P. and Tina Norris**, “Mental Health Treatment Among Adults: United States, 2019 and 2021,” *NCHS Data Brief*, 2022, (419).
- Uscher-Pines, Lori, Jessica Sousa, Pushpa Raja, Ateev Mehrotra, Michael L. Barnett, and Haiden A. Huskamp**, “Nearly Half Of Adults In Families That Lost Work During The Pandemic Report Mental Health Problems,” *Health Affairs*, 2021, *40* (5), 806–814.

**Zuckerman, Stephen, Laura Skopec, and Joshua Ber**, “Medicaid Physician Fees Remained Substantially Below Fees Paid By Medicare In 2019,” *Health Affairs*, 2021, *40* (2), 343–348.

## A. Data Appendix

### A.1 T-MSIS Data Description

The Transformed Medicaid Statistical Information System (T-MSIS) is a CMS-administered database of state-submitted Medicaid claims. The version used in this study was accessed as a pre-processed Parquet file containing 227,083,361 records. Each record represents a unique combination of billing provider NPI, servicing provider NPI, HCPCS procedure code, and service month, aggregated across beneficiaries. Fields include total unique beneficiaries, total claims, and total Medicaid paid amount.

Behavioral health services are identified by HCPCS H-prefix codes. The major codes in our sample include:

- H0031: Mental health assessment by non-physician
- H0036: Community psychiatric supportive treatment, face-to-face, per 15 minutes
- H2015: Comprehensive community support services, individual, per 15 minutes
- H2016: Comprehensive community support services, per diem
- H0004: Behavioral health counseling and therapy, per 15 minutes

Personal care services (T-prefix codes) serve as the placebo group. Major codes include T1019 (personal care aide, per 15 minutes) and T2016 (habilitation, residential, per diem).

Records with fewer than 12 total claims are suppressed by CMS for privacy protection. This disproportionately affects rural providers and rare procedures but removes a negligible share of total spending.

### A.2 NPPES Linkage

The National Plan and Provider Enumeration System (NPPES) provides the NPI-to-state crosswalk. I link T-MSIS billing NPIs to NPPES using exact NPI matching, achieving a 98.4% match rate. Unmatched NPIs are dropped from the analysis. For providers with multiple practice locations, I use the primary practice state recorded in NPPES.

### A.3 Treatment Variable Sources

Telehealth payment parity adoption dates were compiled from:

1. Center for Connected Health Policy (CCHPCA) State Telehealth Laws and Reimbursement Policies Reports, 2019–2025 editions

2. National Conference of State Legislatures (NCSL) telehealth policy database
3. Individual state legislative databases and enrolled bill texts
4. American Medical Association (AMA) State Telehealth Policy Trends: 2023 Year in Review

**Table 4:** Telehealth Payment Parity Law Adoption Dates

State	Effective Date	Statute	Year
Georgia	January 01, 2020	SB 118	2020
New Jersey	January 18, 2021	A 1467	2021
West Virginia	June 10, 2021	HB 2024	2021
Kentucky	June 29, 2021	HB 140	2021
Virginia	July 01, 2021	HB 81	2021
Hawaii	July 01, 2021	HB 907	2021
Minnesota	July 01, 2021	SF 3019	2021
Colorado	July 01, 2021	HB 21-1190	2021
New Mexico	July 01, 2021	HB 245	2021
Indiana	July 01, 2021	SB 3	2021
Mississippi	July 01, 2021	HB 1531	2021
Arkansas	July 28, 2021	Act 829	2021
North Dakota	August 01, 2021	HB 1247	2021
Louisiana	August 01, 2021	HB 449	2021
Missouri	August 28, 2021	SB 5	2021
Connecticut	October 01, 2021	PA 21-9	2021
Montana	October 01, 2021	SB 101	2021
Maine	October 18, 2021	LD 1034	2021
Oklahoma	November 01, 2021	SB 674	2021
New Hampshire	January 01, 2022	HB 602	2022
Delaware	January 01, 2022	HB 348	2022
Illinois	January 01, 2022	SB 2294	2022
South Carolina	May 17, 2022	HB 3726	2022
Iowa	July 01, 2022	HF 2548	2022
Arizona	September 24, 2022	SB 1089	2022
Nebraska	January 01, 2023	LB 400	2023

*Notes:* Permanent state laws requiring Medicaid to reimburse telehealth at parity with in-person services. Compiled from CCHPCA, NCSL, and state legislative records.

## B. Identification Appendix

### B.1 Pre-Treatment Balance

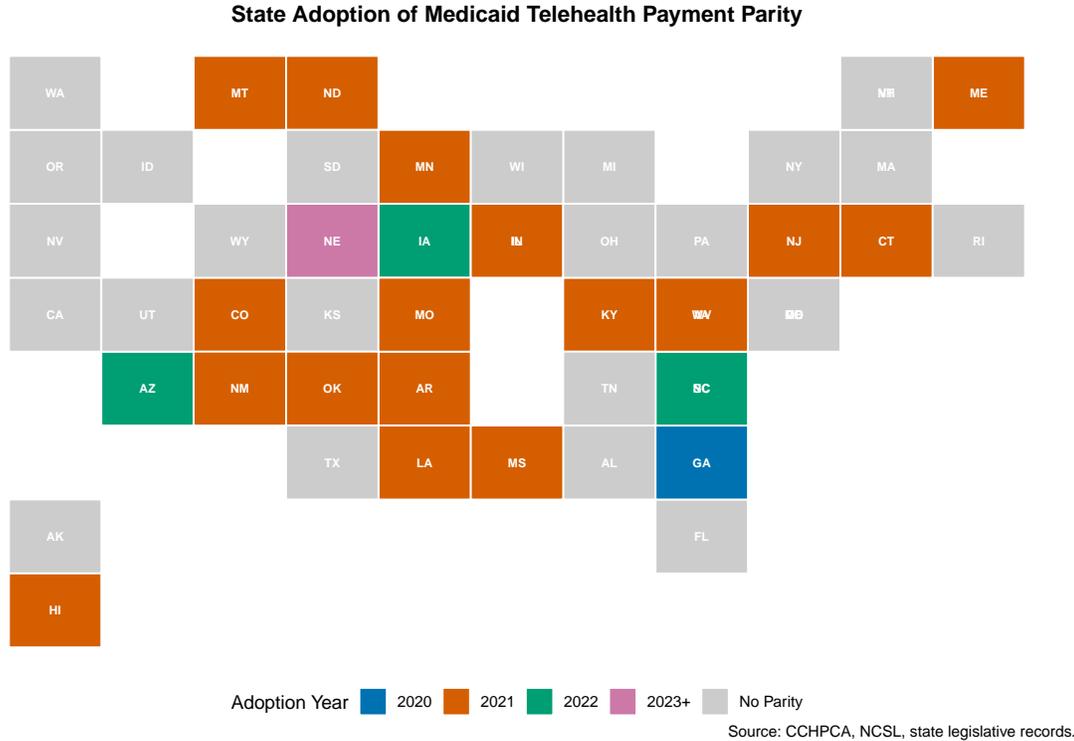
The parallel trends assumption is assessed visually in [Figure 2](#) and statistically through the event-study coefficients in [Figure 3](#). All eight pre-treatment quarterly estimates are individually insignificant, with a joint test unavailable due to a singular covariance matrix (a consequence of the large number of treatment cohorts relative to pre-treatment periods for late adopters).

### B.2 Cohort Composition

The treatment cohorts and their sizes are:

- Q1 2020: 1 state (Georgia)
- Q1–Q4 2021: 19 states (main wave)
- Q1–Q3 2022: 5 states
- Q1 2023: 1 state (Nebraska)

The Callaway-Sant’Anna estimator appropriately handles this unbalanced cohort structure by estimating separate  $ATT(g,t)$  for each group-time pair.



**Figure 8:** State Adoption of Medicaid Telehealth Payment Parity

## C. Robustness Appendix

### C.1 Goodman-Bacon Decomposition Details

The decomposition yields three comparison types:

- Treated vs. Untreated: weight = 0.881, average estimate =  $-0.142$
- Earlier vs. Later Treated: weight = 0.060, average estimate =  $-0.043$
- Later vs. Earlier Treated: weight = 0.060, average estimate =  $-0.006$

The overwhelming weight on treated-vs-untreated comparisons (88%) confirms that the TWFE estimate is primarily driven by clean comparisons to never-treated states.

### C.2 Sun-Abraham Estimates

The Sun-Abraham interaction-weighted estimator produces event-time coefficients that are qualitatively similar to the Callaway-Sant’Anna estimates. Pre-treatment coefficients are individually insignificant and show no trend. Post-treatment coefficients at short horizons (0–8 quarters) are near zero. At longer horizons (15–19 quarters post-treatment, available

only for the earliest cohort), the estimates turn negative and statistically significant, driven by Georgia's post-2023 trajectory. This single-state long-horizon result should be interpreted with extreme caution.