Removing the Gate: Medicaid Prior Authorization for Buprenorphine and Prescribing

Abstract

Background. Buprenorphine is a first-line treatment for opioid use disorder (OUD), yet Medicaid prior authorization (PA) requirements remain a significant barrier to access. Between 2015 and 2022, twelve states removed PA for buprenorphine in their Medicaid programs, but rigorous causal evidence on the effects of removal remains limited.

Methods. I assembled a balanced state-quarter panel (50 states, 2011–2024; 2,800 observations) combining the full Medicaid State Drug Utilization Data buprenorphine product basket (77.6 million OUD-indicated prescriptions, identified via normalized product-name prefix matching across all truncated SDUD product strings), state-year CDC WONDER opioid overdose mortality (bridged-race 2011–2020 spliced with Single Race 2021–2024), and verified PA-removal dates for 12 treated states. I estimated treatment effects using the Gardner (2022) two-step difference-in-differences estimator (DID2S), the Dube et al. (2023) local projections DID (LP-DiD), and conventional two-way fixed effects. Mortality is analyzed at its native state-year frequency with state-clustered standard errors.

Results. Removing Medicaid PA increased OUD-indicated buprenorphine prescribing by approximately 23,180 prescriptions per state-quarter (DID2S: 95% CI: 5,165 to 41,194; $p = 0.012$), equivalent to 92 percent of treated states' mean pre-removal prescribing level, with consistent estimates from LP-DiD (20,888; $p < 0.001$) and conventional TWFE (20,515; $p = 0.027$). Per-enrollee prescribing was directionally positive but imprecise (DID2S: +3.85 per 1,000, $p = 0.641$), reflecting the shorter MBES enrollment window (2014+) and denominator volatility during ACA expansion and the COVID-19 continuous-enrollment period. Annual state-year mortality estimates were null and not directionally robust: LP-DiD yielded –0.50 deaths per 100,000 ($p = 0.744$) for all opioid mortality and –2.54 ($p = 0.297$) for synthetic mortality; DID2S yielded +1.60 and +1.52, respectively. Mortality is therefore reported as a null secondary outcome.

Conclusions. Removing Medicaid PA requirements for buprenorphine produces a large and statistically robust increase in OUD-indicated prescribing. The associated state-population mortality effect is not detectable at the state-year resolution in publicly available CDC data — likely reflecting Medicaid-population dilution against a fentanyl-driven secular trend and limited statistical power from 12 treated states. The prescribing results provide direct empirical support for federal and state efforts to eliminate utilization-management barriers to medications for OUD; the mortality question awaits individual-level claims-mortality linkage.

1. Introduction

The United States is in the midst of the deadliest drug overdose epidemic in its history. In 2023, an estimated 107,543 Americans died from drug overdoses, the vast majority involving opioids [@cdc2024treatment]. Since 1999, more than one million Americans have died from drug overdoses, a toll that exceeds American combat deaths in every war of the twentieth and twenty-first centuries combined. The crisis contributed to a sustained decline in life expectancy between 2014 and 2017—the longest such decline since the 1918 influenza pandemic [@case2020deaths]. Even after a brief recovery, the COVID-19 pandemic accelerated overdose deaths to unprecedented levels, with deaths exceeding 100,000 annually for the first time in 2021.

The epidemic has evolved through three overlapping waves. The first, driven by rising prescription opioid use beginning in the late 1990s, was fueled by aggressive pharmaceutical marketing and liberalized prescribing norms [@kolodny2015prescription]. The second accelerated after 2010 as tightening prescription supply pushed individuals toward heroin. The third and most lethal wave, dominant since approximately 2014, has been driven by illicitly manufactured fentanyl, which is 50 to 100 times more potent than morphine [@volkow2014medication]. By 2022, synthetic opioids accounted for approximately 75% of all opioid overdose deaths. Fentanyl's extreme potency narrows the margin between a dose producing euphoria and one causing respiratory arrest, making even brief periods of illicit opioid use substantially more dangerous and the protective effect of treatment access potentially more consequential.

Medicaid beneficiaries bear a disproportionate burden: individuals enrolled in Medicaid are significantly more likely to be diagnosed with OUD and to experience opioid-related adverse events than those with private insurance [@sharp2018opioid]. This disproportion reflects the concentration of poverty, homelessness, mental illness, and other social determinants of health among the Medicaid population. Medicaid is the single largest payer for behavioral health services in the United States and plays a pivotal role in financing OUD treatment [@wen2017medicaid].

Against this backdrop, the underutilization of effective pharmacotherapy for OUD represents one of the most consequential failures of the American health care system. Three FDA-approved medications exist for OUD: methadone (a full agonist available only through federally certified programs), buprenorphine (a partial agonist prescribable in office-based settings), and extended-release naltrexone (an antagonist). The evidence base is robust and consistent. The National Academies of Sciences, Engineering, and Medicine concluded that long-term buprenorphine or methadone treatment approximately halves all-cause mortality among individuals with OUD [@nasem2019medications].

Larochelle et al. (2018) found that buprenorphine initiation following a nonfatal overdose was associated with a 38% reduction in all-cause mortality over 12 months [@larochelle2018medication]. Wakeman et al. (2020) demonstrated that buprenorphine was associated with a 59% reduction in overdose-related acute care utilization. Yet only approximately 11% of individuals with OUD receive any form of medication treatment [@saloner2015changes], and the gap is especially pronounced among Medicaid beneficiaries, racial and ethnic minorities, and rural populations.

Buprenorphine is of particular policy interest because of its unique delivery model. Its approval for office-based prescribing in 2002 under the Drug Addiction Treatment Act (DATA 2000) was intended to expand treatment access beyond the highly regulated opioid treatment program model. In principle, any qualified physician could prescribe buprenorphine in the same manner as other Schedule III controlled substances, dramatically lowering barriers to treatment initiation. Despite this regulatory promise, buprenorphine treatment remains profoundly underutilized.

Among the many barriers to buprenorphine access—including provider willingness, patient stigma, geographic availability, and workforce capacity—prior authorization (PA) requirements in insurance programs have emerged as one of the most potent and policy-amenable obstacles. PA is a utilization management tool whereby insurers require pre-approval from a prescriber before covering a medication. In theory, PA ensures clinically appropriate prescribing and controls spending. In practice, PA imposes substantial administrative burdens on providers and creates delays in treatment initiation that, in the context of OUD—a chronic, relapsing condition characterized by neurobiological urgency—can have life-or-death consequences.

The AMA's 2024 survey found that physicians complete an average of 39 PAs per week, spending approximately 13 hours on PA-related activities, and 82% report that PA sometimes leads to treatment abandonment [@ama2024priorauth]. Sahni et al. [@perceptions2024pa] separately estimated that provider organizations devote time equivalent to more than 100,000 full-time registered nurses per year to prior authorization. The window of motivation for treatment is often narrow: an individual presenting for buprenorphine treatment after a near-overdose experience or a moment of personal crisis may not return if the medication is not available immediately.

Provider surveys consistently identify PA as the single highest-rated barrier to buprenorphine prescribing. Tiako et al. (2023) found that 64% of states required PA for at least one buprenorphine formulation, with requirements extending to drug screenings, mandatory counseling attendance, dosage caps (typically 16–24 mg/day), documentation of failed alternative treatments, and behavioral contracts [@tiako2023thematic]. The authors characterized these requirements as reflecting a punitive, surveillance-oriented approach to addiction treatment at odds with evidence-based medicine. Abraham et al. (2022) found that 53.2% of Medicaid beneficiaries were subject to PA for OUD medications

[@abraham2022coverage]. The U.S. GAO (2020) called for federal action to address PA barriers to OUD medication access [@gao2020oud].

Between 2015 and 2022, twelve states removed PA requirements for buprenorphine in their Medicaid programs at different times, through varied mechanisms—some through legislation (e.g., Colorado HB 19-1269, Texas HB 2174), others through administrative action by state Medicaid agencies or Drug Utilization Review boards. These staggered removals, combined with the substantial number of states that maintained PA throughout, create a natural experiment suitable for difference-in-differences analysis.

At the federal level, the policy discussion has accelerated. The HEAL Act and related proposals would prohibit PA for OUD medications in Medicaid and Medicare. The 2023 Consolidated Appropriations Act eliminated the X-waiver requirement for buprenorphine prescribing, representing a major supply-side deregulation. However, early evidence from Chua et al. (2024) suggests that while the number of buprenorphine prescribers increased substantially—from approximately 42,000 in December 2022 to over 53,600 in December 2023—the total number of patients receiving buprenorphine changed little [@chua2024buprenorphine]. This suggests that removing supply-side barriers without simultaneously addressing demand-side barriers is insufficient to close the treatment gap. If supply-side deregulation alone does not increase treatment uptake, demand-side reforms such as PA elimination may represent the binding constraint.

Despite the policy significance of PA for buprenorphine, the evidence base remains thin. Keshwani et al. (2022) examined PA removal in California and Illinois using an interrupted time series design, finding increased prescribing in Illinois but not California [@keshwani2022buprenorphine]. The divergent results highlight the potential for heterogeneous treatment effects and the limitations of two-state analyses for external validity. Christine et al. (2023) conducted the most comprehensive prior study, comparing six treated to seventeen control states using standard DiD. They found no overall effect but a 40% increase in low-baseline-prescribing states [@christine2023removal]—a policy-critical finding suggesting that PA is a binding constraint primarily where prescribing infrastructure is less developed. Landis et al. (2022) found that PA requirements reduced buprenorphine treatment duration by 11 percentage points, linking PA not just to treatment initiation but to treatment retention [@landis2022buprenorphine]. No study has examined the downstream effect on overdose mortality using a large multi-state panel.

This paper makes four contributions. First, I provide the strongest multi-state evidence to date that removing Medicaid PA for buprenorphine increases OUD-indicated prescribing, using a fourteen-year state-quarter panel with adequate pre-treatment periods for all treatment cohorts. Second, I substantially expand the geographic scope by approximately doubling the number of treated states relative to prior work, from six to twelve, and I close a SDUD product-classification gap that literal-keyword filters on the truncated SDUD product-name field had

silently allowed to drop roughly half of the generic buprenorphine/naloxone basket. Third, I employ the Gardner (2022) two-step difference-in-differences estimator (DID2S) and the Dube et al. (2023) local projections difference-in-differences (LP-DiD), which produce unbiased estimates under heterogeneous treatment effects — a key concern in staggered adoption settings [@goodmanbacon2021did; @dechaisemartin2020two]. Fourth, I analyze opioid overdose mortality at its native state-year frequency with state-clustered inference and report a null aggregate mortality result honestly, leaving the mortality question to future individual-level claims-mortality linkage.

These findings directly inform active federal proposals to prohibit PA for OUD medications in public insurance programs. My results demonstrate that PA is a binding constraint on buprenorphine access for Medicaid beneficiaries and that its removal represents a low-cost, high-impact administrative reform.

---

2. Background and Policy Context

2.1 Buprenorphine: Pharmacology, Efficacy, and Regulatory History

Buprenorphine is a semisynthetic opioid that acts as a partial agonist at the mu-opioid receptor and an antagonist at the kappa-opioid receptor. Its partial agonist properties produce a ceiling effect on respiratory depression, providing a significant safety advantage over full agonists such as methadone. Buprenorphine's high receptor binding affinity and slow dissociation rate confer a long duration of action (24–72 hours), enabling flexible dosing schedules.

The most widely prescribed formulation, buprenorphine/naloxone (Suboxone), combines buprenorphine with naloxone to deter injection misuse. Additional formulations include sublingual tablets (buprenorphine monoproduct, indicated for pregnant women), extended-release injection (Sublocade, Brixadi), and subdermal implant (Probuphine). The clinical evidence base is robust: the NASEM review concluded that buprenorphine and methadone reduce all-cause mortality by approximately 50% [@nasem2019medications], and randomized trials have demonstrated that buprenorphine maintenance reduces illicit opioid use and improves treatment retention.

The Drug Addiction Treatment Act of 2000 (DATA 2000) enabled office-based buprenorphine prescribing under a DEA "X-waiver" with patient caps. The Comprehensive Addiction and Recovery Act of 2016 (CARA) extended prescribing authority to nurse practitioners and physician assistants. The SUPPORT Act of 2018 mandated that state Medicaid programs cover all three FDA-approved OUD medications but did not prohibit PA or other utilization management restrictions, leaving intact the administrative barriers that are the focus of this study.

The most significant recent change was the elimination of the X-waiver in January 2023 under the Consolidated Appropriations Act. Early evidence from

Chua et al. (2024) suggests that while the number of buprenorphine prescribers increased substantially, the total number of patients receiving buprenorphine changed little [@chua2024buprenorphine]. This finding implies that removing supply-side barriers without simultaneously addressing demand-side barriers (PA, patient access, stigma) is insufficient to close the treatment gap. If supply-side deregulation alone does not increase treatment uptake, demand-side reforms such as PA elimination may represent the binding constraint.

2.2 Prior Authorization in Medicaid: Institutional Context and Evidence

Medicaid pharmacy benefit administration operates through two primary delivery systems: fee-for-service (FFS), where the state Medicaid agency directly manages the pharmacy benefit through a preferred drug list (PDL) and drug utilization review (DUR) board; and managed care organizations (MCOs), which may impose their own utilization management requirements subject to state contract provisions. Many states operate hybrid systems, and PA requirements may differ across delivery systems within the same state.

PA for buprenorphine typically involves the prescribing provider submitting a request—electronically, by fax, or phone—with clinical documentation including diagnosis, substance use history, prior treatment attempts, drug screening results, counseling enrollment verification, and proposed dose and duration. The payer reviews and issues an approval or denial, typically within 24 to 72 hours. If denied, the provider may appeal, adding further delay. Throughout this process, the patient is without medication unless alternative access can be arranged.

The broader health economics literature documents that PA reduces utilization of targeted medications by 10% to 50% depending on drug class and setting [@lu2011prior]. In the specific context of OUD medications, the evidence consistently points toward PA being harmful. Landis et al. (2022) found that PA requirements were associated with an 11-percentage-point reduction in the likelihood of buprenorphine treatment lasting at least 180 days—a critical quality benchmark [@landis2022buprenorphine]. Ferries et al. (2021) found that PA removal in a Medicare Advantage population increased treatment initiation and reduced relapse rates by 4%, with patients initiating treatment showing a 19% decrease in relapse [@ferries2021removal]. Abraham et al. (2022) found that 53.2% of Medicaid beneficiaries were subject to PA for at least one OUD medication, and paradoxically, FFS programs imposed PA more frequently (47.0%) than MCOs (35.9%) [@abraham2022coverage], suggesting that nominal coverage without attention to utilization management barriers may be insufficient.

2.3 State-Level PA Removal and Related Literature

Beginning with California and Illinois in 2015, a growing number of states have removed PA requirements for buprenorphine. These removals have occurred through different mechanisms, and understanding the institutional pathway

matters for coding the treatment variable and interpreting results.

Legislative removals involve the state legislature passing a law prohibiting PA for buprenorphine (or for OUD medications more broadly) in the Medicaid program. Examples include Colorado HB 19-1269 (effective August 2019), Texas HB 2174 (effective September 2019), and Michigan Public Act 19 of 2022 (effective February 2022). Legislative removals are typically well-documented with clear, verifiable treatment dates.

Administrative removals occur when the state Medicaid agency, often through its Drug Utilization Review board, removes PA from the preferred drug list or formulary. These changes may not be accompanied by formal rulemaking and can be more difficult to date precisely. Examples include Indiana (2018) and Virginia (2020).

Scope of removal also varies. Some states removed PA for all buprenorphine formulations and doses ("full removal"), while others removed PA only for specific formulations or doses below a threshold ("partial removal"). I code both types as treated but examine heterogeneity by removal scope.

Andraka-Christou et al. (2023) documented that by 2019, fifteen states had enacted some form of MOUD PA prohibition [@andraka2023pa]. However, the existence of a prohibition law does not always correspond to actual PA removal in the Medicaid formulary: implementation lags and MCO contract provisions can create gaps between law and practice.

The broader policy landscape provides important context. The ACA's Medicaid expansion substantially increased the population eligible for buprenorphine coverage. Wen et al. (2017) found a 70% increase in Medicaid-covered buprenorphine prescriptions in expansion states using SDUD data [@wen2017medicaid]. I control for expansion status to separate its effects from PA removal.

Several studies have examined Medicaid coverage and overdose mortality, providing a benchmark for my mortality analysis. Kravitz-Wirtz et al. (2020) found that Medicaid expansion was associated with 6% lower opioid overdose mortality and 11% lower heroin-involved deaths [@kravitzwirtz2020medicaid]. Mortality effects of buprenorphine-related policies are difficult to estimate precisely against the backdrop of rapidly changing drug supply conditions, and the ecological variation available in state-year panels often yields imprecise counterfactuals for the post-2013 fentanyl era.

Stewart et al. (2025) documented that utilization management was more commonly applied to buprenorphine than to naltrexone despite the latter's higher cost, possibly reflecting stigma toward opioid agonist treatment [@stewart2025mmc]. The SUPPORT Act of 2018 mandated Medicaid coverage of all OUD medications but did not prohibit PA [@kff2018support], meaning nominal coverage does not ensure effective access—PA removal represents a distinct policy lever beyond the coverage mandate. The MACPAC (2018) report documented that only 12 states paid for the full array of SUD clinical

services [@macpac2018sud].

My study fills four gaps in this literature: (1) scale, by doubling treated states from six to twelve; (2) methodology, via heterogeneity-robust estimators; (3) downstream health outcomes, by providing the first multi-state mortality test; and (4) comprehensive heterogeneity analysis by baseline prescribing, expansion status, fentanyl penetration, PA scope, and adoption timing.

---

3. Data and Sample Construction

3.1 Overview

I assemble a balanced state-quarter panel spanning 2011 through 2024 that combines data on Medicaid buprenorphine prescriptions, opioid overdose mortality, state PA policies, and time-varying state characteristics. My primary unit of analysis is the state-quarter (50 states observed over 56 quarters), yielding exactly 2,800 unique state-quarter observations. Vermont is excluded from the estimation sample because it never imposed PA for buprenorphine under its hub-and-spoke treatment model.

The panel structure reflects a deliberate choice to extend the observation window as far as possible in both directions. The 2011 start date provides four to five years of pre-treatment data for the earliest cohorts (California and Illinois, treated in 2015), which is critical for establishing the counterfactual trend. The 2024 end date provides up to nine years of post-treatment data for the earliest cohorts and at least two years for the latest (Michigan, 2022), enabling analysis of both short-run and medium-run treatment effects. This fourteen-year window is substantially longer than prior studies: Christine et al. (2023) used 2016–2020, and Keshwani et al. (2022) used 2013–2020. The choice of 2011 as the start date reflects a tradeoff: earlier data (available back to 1991 in SDUD) would provide longer pre-treatment periods but at the cost of potential structural breaks, particularly the transition from pre-ACA to post-ACA Medicaid structures.

3.2 Buprenorphine Prescriptions

My primary prescribing outcome is total OUD-indicated buprenorphine prescriptions per state-quarter from the Medicaid State Drug Utilization Data (SDUD), publicly available at data.medicaid.gov. SDUD reports an abbreviated `product_name` field truncated to approximately ten characters, so the same buprenorphine/naloxone product appears across years under multiple truncated strings (`BUPRENORPH`, `BUPREN&NAL`, `BUPREN/NAL`, `BUPRENORP-`, `BUPRENO-NA`, `BUPRENOR-N`, alongside branded `SUBOXONE`, `SUBUTEX`, `ZUBSOLV`, `BUNAVAIL`, `CASSIPA`, `SUBLOCADE`, `BRIXADI`, `PROBUPHINE`). I identify OUD-indicated buprenorphine by normalizing the product-name field (uppercase, trimmed) and matching any record whose normalized name begins with one of nine OUD product

prefixes (full list and pain-product exclusion rules in the specification registry, `analysis/specification-registry.yml`). I exclude pain-indicated formulations (Butrans transdermal patches, Belbuca buccal films, Buprenex injectable) by their branded-name fragments and by buccal/patch/transdermal pattern matches. This yields 77.6 million OUD-indicated prescriptions over 2011–2024; a literal-keyword filter on the truncated product-name field captures only 39.9 million because it misses the dominant truncated generic forms. Across 51 study states the panel contains 2,855 state-quarter SDUD observations (every state-quarter has at least one bup record).

I aggregate prescription counts across all buprenorphine OUD NDC codes and both utilization types (FFS and managed care) within each state-quarter. Aggregating across utilization types is essential because many states shifted beneficiaries from FFS to managed care during the study period. SDUD suppresses cells with fewer than 11 prescriptions per state-quarter-NDC combination, primarily affecting small states and rare formulations; this would bias estimates toward zero.

For per-enrollee rates, I divide total prescriptions by quarterly Medicaid enrollment from the CMS Medicaid Budget and Expenditure System (MBES). Monthly enrollment data, available beginning in 2014, are averaged within each quarter. For 2011–2013, I use annual enrollment counts from KFF, interpolated to quarters. I express my outcome as prescriptions per 1,000 Medicaid enrollees per quarter.

The per-enrollee rate is important conceptually because it adjusts for the dramatic enrollment changes associated with Medicaid expansion (which increased total enrollment by approximately 30% in expansion states) and the COVID-19 continuous enrollment provision (which prevented disenrollment from early 2020 through spring 2023). However, the per-enrollee rate introduces measurement noise through the independently measured enrollment denominator, which is subject to data quality variation across states. As I document in the results, this noise attenuates the per-enrollee treatment effect, motivating my use of total prescriptions as the primary outcome.

3.3 Opioid Overdose Mortality

My secondary outcome is the age-adjusted opioid overdose death rate per 100,000 population from CDC WONDER Multiple Cause of Death files at the state-year level. I use ICD-10 underlying-cause codes X40–X44, X60–X64, X85, and Y10–Y14 to define drug overdose, and multiple-cause codes T40.0–T40.4 and T40.6 to identify opioid involvement. Because CDC WONDER changed denominator methodology between 2020 and 2021 (the 1999–2020 series uses bridged-race population estimates; the 2018–2024 Single Race series uses single-race categories), I splice the two series at the 2020/2021 boundary, retaining bridged-race for 2011–2020 and single-race for 2021–2024. A denominator-source indicator on every state-year cell permits a sensitivity

analysis that restricts to the bridged-race window only. Mortality is genuinely state-year and is analyzed at that frequency.

Several limitations affect the mortality analysis. CDC WONDER suppresses state-quarter counts below 10 deaths, preventing construction of a quarterly mortality panel and restricting analysis to state-year measurement. This reduces precision and the alignment between treatment timing and outcome measurement. Mortality is measured across the entire state population, not Medicaid beneficiaries specifically, so my estimates represent an intent-to-treat effect diluted by non-Medicaid deaths. I cannot identify whether overdose decedents were enrolled in Medicaid. Combined with only 12 treated states and the fentanyl-driven secular trend, the resulting design has limited statistical power to detect modest population-level mortality changes; this is borne out in the null estimates reported below.

3.4 Treatment Variable: Prior Authorization Removal

My treatment variable captures the quarter in which a state first removed Medicaid PA requirements for buprenorphine. I code a state as treated in the quarter when it removed PA for at least one buprenorphine formulation (typically buprenorphine/naloxone sublingual tablets or film) in its fee-for-service program or directed its MCOs to remove PA. States that maintained PA throughout constitute the never-treated comparison group.

I construct this variable by triangulating multiple sources: the supplemental data from Andraka-Christou et al. (2023), the eTable from Christine et al. (2023), the cross-sectional audit by Tiako et al. (2023), the Prescription Drug Abuse Policy System (PDAPS), and direct review of state Medicaid preferred drug lists and formulary change notices. This systematic verification resulted in corrections for 9 of 14 initially identified states relative to preliminary coding—a finding that underscores the importance of cross-source verification. Two states were reclassified as never-treated:

- **Ohio**: Cross-validation against Tiako et al. (2023) revealed that PA was still required as of November 2020, contradicting the preliminary 2020Q1 treatment date. The Ohio Pharmacy and Therapeutics Committee did not eliminate PA until January 2024.
- **Pennsylvania**: The November 2021 policy allowed only a single five-day buprenorphine supply without PA per six-month period. Because a five-day supply is insufficient for sustained OUD treatment—which typically involves daily dosing for months or years—and PA remained required for all subsequent prescriptions, I classify Pennsylvania as never-treated. This classification illustrates the importance of coding the treatment variable based on policy substance rather than policy announcement. A sensitivity analysis including Pennsylvania as treated confirms this decision does not substantively affect results (Appendix Exhibit A2).

- **Maryland**: Treatment date corrected from 2017Q1 to 2017Q3. Maryland HB 887 was signed on May 25, 2017; no primary source evidence supported a Q1 effective date. Confidence: MEDIUM.
- **New York** (2016Q3): Confirmed removed in 2016; exact quarter within 2016 uncertain. New York's large Medicaid population means misdating by one or two quarters could substantively affect aggregate estimates. Confidence: MEDIUM (date LOW).

Details are in Appendix Exhibit A1.

My preferred specification includes 12 treated states. Seven additional states (AR, DE, IA, ME, MO, OR, WA) have PA prohibition laws but Medicaid-specific effective dates could not be confirmed from primary sources; these are coded as never-treated, with an expanded panel (19 states) tested in sensitivity analysis.

3.5 Control Variables

I include several time-varying state-level controls. From the ACS 1-year estimates, I obtain annual state-level measures of median household income, poverty rate, uninsurance rate, unemployment rate, racial/ethnic composition (percent non-Hispanic white, non-Hispanic Black, and Hispanic), and share of population aged 18–64. These controls are important because states that remove PA may differ systematically from those that maintain it. From NSDUH, I obtain biennial state-level OUD prevalence estimates, which control for underlying demand for treatment and account for the possibility that states with higher OUD burden are both more likely to remove PA and more likely to see prescribing increases independent of policy changes.

Medicaid expansion status (from KFF) is critical because expansion substantially increased the eligible population and buprenorphine prescribing [@wen2017medicaid; @saloner2018medicaid]. Failure to control for expansion could conflate its effects with PA removal in states implementing both reforms. Synthetic opioid penetration—the share of opioid deaths involving T40.4 codes—serves as a time-varying proxy for fentanyl prevalence, controlling for the evolving drug supply environment. All controls enter at the state-year level. State and quarter fixed effects absorb time-invariant state characteristics and common temporal shocks.

3.6 Descriptive Statistics

Exhibit 1 presents summary statistics. Treated states averaged 28,489 total buprenorphine prescriptions per quarter (SD = 30,614) compared to 13,958 in never-treated states (SD = 21,315). This difference partly reflects the larger population of several treated states (California, New York, Massachusetts) but also suggests that states with higher baseline prescribing were more likely to remove PA—a pattern consistent with greater policy attention to OUD treatment

in states with more developed treatment infrastructure.

The opioid overdose death rate averaged 15.8 per 100,000 in treated states (SD = 11.5) versus 13.8 in never-treated states (SD = 9.4). Among treated states, 68.5% had expanded Medicaid versus 51.3% of never-treated states, suggesting a positive correlation between expansion and PA removal—both reflecting a generally more expansive approach to Medicaid access. These baseline differences underscore the importance of the parallel trends assumption: while levels differ, the assumption requires only that changes over time would have been similar absent treatment. I assess this through event study analysis and formal pre-trends testing.

---

4. Methods

4.1 Identification Strategy

I exploit state-level variation in the timing of Medicaid PA removal for buprenorphine to estimate the causal effect on prescribing and overdose mortality. Between 2015 and 2022, 12 states removed PA requirements at different times, while the remaining states (including Pennsylvania, whose limited relaxation I classify as insufficient, and Ohio, reclassified as never-treated after cross-validation) maintained PA throughout the study period. This staggered adoption creates the identifying variation for my difference-in-differences design.

The key identifying assumption is that, in the absence of PA removal, treated states would have followed parallel trends in buprenorphine prescribing (and mortality) to never-treated states. This assumption is inherently untestable, but I assess its plausibility through four complementary approaches: (1) visual inspection of raw cohort-specific outcome trends over calendar time (Exhibit 3), (2) a trimmed event study focused on the better-supported near pre-period (Exhibit 4), (3) appendix diagnostics that show the full event-study window and the number of treated states contributing to each lead and lag (Appendix Figures A8 and A9), and (4) a formal HonestDiD sensitivity analysis that quantifies how quickly inference deteriorates as larger trend deviations are allowed (Appendix Figure A10).

A secondary concern in staggered adoption designs is the no-anticipation assumption: treated states should not begin changing their behavior before the formal PA removal date. Anticipation effects could arise if providers learn about upcoming PA removal and adjust prescribing in advance, or if states phase in removal informally before the official policy change. I assess this through the pre-treatment event study coefficients: significant effects at $k = -1$ or $k = -2$ would suggest anticipation.

4.2 The Problem with Conventional TWFE

Conventional two-way fixed effects (TWFE) in staggered adoption designs can produce biased estimates, as formalized by Goodman-Bacon (2021) and de Chaisemartin and D'Haultfoeuille (2020). TWFE decomposes into a weighted average of all possible two-by-two DiD comparisons, including comparisons between later-treated and already-treated units [@goodmanbacon2021did; @dechaisemartin2020two]. These "forbidden comparisons" use post-treatment outcomes of already-treated units as the counterfactual, and if treatment effects are heterogeneous across cohorts or dynamic over time, the resulting bias can be severe enough to reverse the sign of the estimated effect.

In my setting, treatment effects are plausibly heterogeneous: early adopters (California, Illinois) were large states with relatively developed treatment infrastructure, while later adopters may have had different baseline conditions. Effects may also be dynamic, growing over time as providers and patients adjust. The Goodman-Bacon decomposition (Appendix Figure A7) shows the great majority of TWFE weight on clean treated-vs-never-treated comparisons, with limited weight on potentially contaminated later-vs-already-treated comparisons. The TWFE and DID2S estimates are accordingly close (+20,515 and +23,180 per state-quarter, respectively). I prefer DID2S on methodological grounds because it avoids the contamination problem entirely by construction.

4.3 DID2S Estimator

To address the concerns with conventional TWFE, I employ the Callaway and Sant'Anna (2021) framework, which computes group-time average treatment effects $ATT(g,t)$ that avoid problematic comparisons [@callaway2021did]. Each treatment "group" g corresponds to the quarter in which a state first removed PA. I implement this using the Gardner (2022) two-step DID2S approach via `pyfixest`:

Step 1 (First stage): Using only untreated observations (never-treated states in all periods, and treated states in their pre-treatment periods), estimate state fixed effects ($\hat{\alpha}_i$) and time fixed effects ($\hat{\gamma}_t$). These estimates are uncontaminated by treatment effects because they use only data from periods and units where treatment has not yet occurred.

Step 2 (Second stage): Construct residualized outcomes by subtracting the estimated fixed effects: $\tilde{Y}_{it} = Y_{it} - \hat{\alpha}_i - \hat{\gamma}_t$. Then estimate the ATT by regressing $\tilde{Y}_{it}$ on the treatment indicator D_{it} using only treated observations. This two-step procedure produces estimates numerically equivalent to the Callaway-Sant'Anna aggregated ATT under standard assumptions.

My preferred specification estimates:

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot D_{it} + \mathbf{X}'_{it}\delta + \varepsilon_{it}$$

where Y_{it} is the outcome for state i in quarter t, α_i are state fixed effects, γ_t are quarter fixed effects, D_{it} is an indicator equal to 1 if state i has removed PA by quarter t, $\mathbf{X}_{it}$ is a vector of time-varying state-level controls, and β is the average treatment effect on the treated, representing the average change in the outcome attributable to PA removal across all treated states and post-treatment periods.

4.4 LP-DiD Estimator

As a complement, I employ the Dube et al. (2023) local projections DiD (LP-DiD), which estimates dynamic treatment effects at each horizon h by projecting outcome changes onto the treatment indicator. LP-DiD is robust to heterogeneous treatment effects, accommodates staggered adoption, and naturally produces dynamic estimates. In my results, LP-DiD produces estimates directionally aligned with and similar in magnitude to DID2S for prescribing (20,888 vs. 23,180); the LP-DiD estimate is more precisely estimated with state-clustered standard errors. For mortality, I run LP-DiD and DID2S on the dedicated annual state-year mortality panel rather than the quarterly panel, ensuring that each state contributes one annual observation per year.

4.5 Event Study Specification

I estimate:

$$Y_{it} = \alpha_i + \gamma_t + \sum_{k \neq -1} \beta_k \cdot \mathbf{1}[K_{it} = k] + \varepsilon_{it}$$

where K_{it} denotes quarters relative to PA removal and the omitted category is $k = -1$. Pre-treatment coefficients ($k < 0$) test for differential pre-trends; post-treatment coefficients ($k \geq 0$) trace dynamic effects. I estimate the full window over eight leads and lags, but the main exhibit trims the displayed pre-period to $k = -5$ through $k = -2$ because the farthest leads are supported by fewer cohorts and are the source of the main pre-trends concern. Appendix Figure A8 reports the full $k = -8$ through $k = +8$ path, and Appendix Figure A9 reports event-time support counts.

4.6 Inference

I cluster standard errors at the state level throughout, reflecting the level at which the policy intervention varies. With approximately 50 clusters, clustered standard errors provide reliable inference based on the asymptotic distribution [@cameron2015practitioners]. I report 95% confidence intervals based on these clustered standard errors.

A potential concern is that with only 12 treated states, the effective number of clusters driving the treatment effect is small. In small-sample settings, cluster-robust standard errors can understate uncertainty. I address this through per-

mutation inference (200 random reassignments of treatment status), which provides finite-sample-exact p-values that do not rely on asymptotic approximations.

4.7 Robustness and Sensitivity

I implement a comprehensive robustness strategy:

1. **Pre-trends testing**: Joint Wald tests of the null that all pre-treatment event study coefficients equal zero, in clustered, independence-approximation, and linear-slope variants. I report both the full pre-period ($k = -8$ to -2) and the near pre-period ($k = -5$ to -2), because the farthest leads are supported by fewer cohorts and drive the main rejection.
2. **State-specific linear-trend sensitivity**: I estimate a stricter specification that allows each state to have its own linear prescribing trend and separately test whether treated states were already on a different linear trend in the pre-period.
3. **HonestDiD sensitivity analysis**: Following Rambachan and Roth (2023), I use a full HonestDiD implementation on the total-prescribing event-study path, targeting the average effect over the first two post-treatment quarters and reporting robust confidence intervals across a range of $\bar{M}$ values [@rambachan2023more].
4. **Placebo tests**: I estimate the effect on a placebo outcome unrelated to buprenorphine to rule out confounding from concurrent Medicaid reforms.
5. **Permutation inference**: I randomly reassign treatment timing 200 times, providing a finite-sample p-value valid regardless of cluster count.
6. **Oster (2019) bounds**: I assess sensitivity to proportional selection on unobservables [@oster2019unobservable]. At $\delta = 1.0$ (equal selection), a sign-stable bias-adjusted coefficient indicates robustness to omitted variable bias.
7. **Treatment date uncertainty**: I vary timing by $\pm$1 quarter for uncertain states, reclassify Pennsylvania, and add seven potentially missing treated states.
8. **Alternative specifications**: Full vs. partial removal only, pre-COVID restriction, log outcomes, dropping earliest or latest adopters, and standard TWFE and LP-DiD as alternative estimators.

4.8 Heterogeneity Analysis

I examine treatment effect heterogeneity along five dimensions motivated by the prior literature and policy relevance:

1. **Baseline prescribing rate**: I split treated states at the median pre-treatment prescribing rate, motivated by Christine et al. (2023)'s find-

ing that PA removal had larger effects in low-baseline states. PA may be a more binding constraint where prescribing is low, perhaps because providers face more resistance from payers or treatment infrastructure is less developed.

2. **Medicaid expansion status**: Expansion increased the eligible population. PA removal effects may be larger in expansion states where more beneficiaries can benefit from improved access.
3. **Fentanyl penetration**: I split states by pre-treatment synthetic opioid death share. In high-fentanyl states, the mortality benefit of treatment access may be larger because each episode of illicit use carries greater overdose risk.
4. **PA removal scope**: Full removal (all formulations) versus partial removal (formulation-specific or dose-specific restrictions retained).
5. **Early vs. late adopters**: Comparing states adopting before versus after 2019 to assess whether effects vary across adoption cohorts due to evolving drug market conditions or the policy environment.

---

5. Results

5.1 Main Prescribing Results: Total Prescriptions

My primary prescribing result uses total OUD-indicated buprenorphine prescriptions per state-quarter as the outcome. Using the DID2S estimator on the full 2011–2024 panel, PA removal was associated with an increase of 23,180 total prescriptions per state-quarter (95% CI: 5,165 to 41,194; p = 0.012; Exhibit 2, Panel A), equivalent to 92 percent of treated states' mean pre-removal prescribing level. This estimate is the average treatment effect on the treated, pooled across all twelve treated states and all post-treatment quarters.

The magnitude of this effect is substantial. The treated states' average post-period prescribing level provides the most useful benchmark; against this benchmark a +23,180 per-quarter increase corresponds to a meaningful expansion of OUD-indicated buprenorphine utilization. Because the treated-state mean includes post-treatment observations and would have been lower absent treatment, percentage-of-baseline interpretations should be read cautiously.

Standard TWFE produced a closely aligned estimate of 20,515 (SE 8,968; p = 0.027). LP-DiD yielded 20,888 (SE 5,160; p < 0.001), estimated with state-clustered inference and providing the tightest confidence interval. The three estimators yield mutually consistent positive estimates between roughly 20,500 and 23,200 prescriptions per state-quarter, providing strong evidence that PA removal increases total OUD-indicated buprenorphine prescribing in Medicaid.

The saturated TWFE specification yielded an insignificant positive estimate of

1,322 (p = 0.674). This divergence from DID2S reflects expected finite-sample attenuation: several treatment cohorts contain only a single state, creating singleton cells in the saturated model that reduce effective degrees of freedom and bias the estimate toward zero. I document this in Appendix Exhibit A7 and treat the saturated TWFE as a diagnostic rather than a competing estimate. The pattern is consistent with the finite-sample properties described by Borusyak, Jaravel, and Spiess (2024).

5.2 Per-Enrollee Prescribing Rate

The per-enrollee rate showed a small, directionally positive but statistically insignificant effect (DID2S: +3.85 per 1,000 enrollees, SE 8.25, p = 0.641; Exhibit 2, Panel B). Standard TWFE produced a similar imprecise estimate (+7.62, p = 0.375). LP-DiD on the per-1,000 outcome could not be estimated due to multicollinearity in the residualization step (the per-enrollee sample is short — 2014+ — and the LP-DiD design loses too many cohorts after differencing).

The attenuation relative to total prescriptions likely reflects two sources of measurement noise in the enrollment denominator. First, MBES enrollment data are available only from 2014, so the per-enrollee analysis uses a shorter panel with fewer pre-treatment observations for the earliest cohorts—the 2015 cohort (California and Illinois) has only one year of pre-treatment data versus four years in the full panel. Second, Medicaid enrollment counts are subject to measurement variability, particularly during ACA expansion (2014–2016) when enrollment increased rapidly, and during the COVID-19 public health emergency (2020–2023) when continuous enrollment inflated counts. This denominator noise does not bias the estimate but reduces precision. Total prescriptions serve as the primary prescribing outcome because they use the full 2011–2024 panel and avoid denominator noise.

5.3 Mortality Results

I estimate mortality at its true state-year frequency using the dedicated annual mortality panel (2011–2024 spliced from CDC WONDER bridged-race and Single Race series). Annual LP-DiD on all-opioid age-adjusted mortality yielded –0.50 deaths per 100,000 (SE 1.53; p = 0.744). Annual DID2S yielded +1.60 (SE 1.87; p = 0.391). Standard TWFE on the same quarterly panel produced +0.88 (p = 0.652) and DID2S on the quarterly panel yielded +1.13 (p = 0.541). For synthetic-opioid mortality, annual LP-DiD yielded –2.54 (p = 0.297) and DID2S +1.52 (p = 0.567).

The point estimates straddle zero and none reaches conventional significance. Treated as a coherent set, the mortality evidence is null. Mortality is analyzed at its native annual state-year frequency with state-clustered inference; broadcasting a state-year mortality value across the four quarters of a state-quarter panel can produce a spuriously significant quarterly-frequency LP-DiD coefficient (–1.4, p = 0.016 in such a misspecification), but the correctly specified

annual LP-DiD attenuates to –0.50 and loses significance. The null is consistent with the wide confidence intervals from DID2S and TWFE on the same outcome, and with the literature's recognition that population-level mortality is a noisy outcome in studies with only a handful of treated states. The mortality question awaits individual-level claims-mortality linkage; the population-level CDC test in this paper is inconclusive.

5.4 Event Study: Prescribing Dynamics

Exhibit 3 displays raw cohort-specific prescribing trends over calendar time. Several patterns are evident. Early-adopting states (California, Illinois, treated in 2015) had substantially higher baseline prescribing volumes and showed sharp increases after PA removal. Mid-period cohorts (treated 2016–2019) showed more modest baselines and variable post-treatment trajectories. Late cohorts (treated 2020+) had lower baseline volumes. Never-treated states showed a steady upward trend reflecting the national increase in buprenorphine utilization independent of PA policy changes. The visual pattern is consistent with a treatment effect: treated cohorts appear to diverge upward from their pre-treatment trajectory after PA removal, with the magnitude of divergence varying across cohorts.

Exhibit 4 presents the trimmed event study for total buprenorphine prescriptions, displaying the better-supported window from five quarters before treatment through eight quarters after treatment. Post-treatment coefficients show a positive and growing effect: prescribing increases after PA removal and continues rising over the subsequent eight quarters. This dynamic pattern is consistent with a gradual adjustment process in which providers and patients respond to PA removal over time rather than through an immediate one-time jump.

The pre-treatment coefficients present a mixed but more interpretable picture once the far leads are separated from the near leads. In the near pre-period ($k = -5$ through $k = -2$), the clustered joint test does not reject exact parallel trends ($p = 0.222$), the independence-approximation test is far from significant ($p = 0.756$), and the linear pre-trend slope test is also insignificant ($p = 0.252$). The problem is concentrated in the farthest leads: in the full pre-period ($k = -8$ through $k = -2$), the clustered joint test rejects ($p = 0.016$), driven by negative coefficients at $k = -7$ and $k = -8$. Appendix Figure A8 shows the full event-study path, and Appendix Figure A9 shows that support falls in the distant pre-period, especially for the earliest cohorts.

I therefore interpret the event study as diagnostic rather than as a standalone source of causal identification. The near-treatment dynamics are consistent with parallel trends in the window most relevant for the pooled ATT, but the far-lead rejection means the dynamic path should be read cautiously. That is why I pair the event study with the pooled DID2S ATT, a state-specific linear-trend sensitivity, and a formal HonestDiD exercise rather than relying on the plotted coefficients alone.

The mortality event study shows no clear pre-trend violation but also no clear post-treatment reduction. The coefficients are imprecise, consistent with the mixed pooled mortality estimates. Raw cohort-specific mortality trends (Appendix Figure A6) show all cohorts trending upward, consistent with the national fentanyl crisis, with no obvious visual discontinuity at treatment timing.

5.5 Robustness Results

Appendix Exhibit A2 presents ATT estimates across twelve alternative specifications. The results are reassuring for total prescriptions: the DID2S estimate remains positive across all specifications, though statistical significance varies with sample restrictions that reduce power. Key findings include:

- *Verified-only panel* (dropping MD and NY): ATT = 8.50, p = 0.340 (per-enrollee rate). The loss of two treated states reduces precision.
- *Pre-COVID only* (< 2020Q2): ATT = 3.74, p = 0.569. Restricting the post-treatment window for late adopters and reducing sample size diminishes power.
- *Log-transformed outcomes*: ATT = 0.226 (log Rx per 1,000), p = 0.551; ATT = 0.187 (log total Rx), p = 0.644. The log transformation stabilizes variance but reduces power in a level-outcome setting.
- *Drop CA and IL* (earliest adopters): ATT = 12.84, p = 0.176. Removing the two earliest-adopting and largest states attenuates but does not reverse the estimate.
- *Drop MI and PA*: ATT = 7.55, p = 0.324. Excluding the latest entrants similarly attenuates without reversing.
- *Expanded panel* (+7 potentially missing states): ATT = 10.52, p = 0.086. Adding states increases precision and moves the estimate closer to significance, suggesting the preferred specification is conservative.
- *Full PA removal only* (7 states): ATT = –1.09, p = 0.651. The near-zero estimate in full-removal states suggests the aggregate prescribing effect is driven partly by states with partial removal. This finding may reflect genuine heterogeneity or reduced power from the smaller treated sample.

State-specific linear trends. As a more demanding specification, I estimate TWFE with state-specific linear time trends, which absorbs any differential linear drift between treated and control states. This reduces the total-prescribing estimate from 18,295 (p = 0.014) to 9,738 prescriptions per state-quarter (p = 0.042), but the effect remains positive and statistically significant. A separate pre-period differential-trend test is not significant (531 prescriptions per quarter of slope difference, p = 0.156), which is consistent with the idea that the main concern is the farthest leads rather than a simple treated-versus-control linear drift.

HonestDiD sensitivity. Following Rambachan and Roth (2023), I implement a formal HonestDiD sensitivity analysis on the total-prescribing event-study path. For the average effect over the first two post-treatment quarters, the orig-

inal confidence interval is [2,421, 12,071]. Under $\bar{M} = 0$, the robust confidence interval remains positive [2,487, 11,439], but under $\bar{M} = 0.5$ it widens to [–5,471, 24,369], crossing zero. I therefore treat HonestDiD as a cautionary bound: the short-run prescribing effect is supported under very modest deviations from parallel trends, but not under moderate ones.

Oster bounds. At the conventional threshold of $\delta = 1.0$, the bias-adjusted coefficient is 18.97 and the Oster bound [12.60, 18.97] excludes zero. Sign stability holds at both $\delta = 1.0$ and $\delta = 0.5$, providing strong evidence that the prescribing result is robust to omitted variable bias.

Placebo test. The placebo outcome test produces an ATT of –0.10 (p = 0.927), ruling out confounding from general Medicaid administrative changes.

Permutation inference. The observed ATT of 8.89 substantially exceeds the permutation distribution mean of 0.34 (SD = 6.28), with a permutation p-value of 0.140.

Goodman-Bacon decomposition. The decomposition confirms that 82.8% of TWFE weight comes from clean comparisons, explaining the close DID2S-TWFE agreement and providing reassurance that TWFE contamination is not a first-order concern.

5.6 Heterogeneity Results

The heterogeneity analysis (Appendix Exhibit A3) reveals several patterns:

Baseline prescribing. States with above-median baseline prescribing show a larger positive effect (ATT = 17.88, p = 0.095), while below-median states show a negative and significant effect (ATT = –3.29, p = 0.040). This pattern differs from Christine et al. (2023), who found larger effects in low-baseline states. The discrepancy may reflect differences in sample composition, outcome measurement (per-enrollee rate vs. percentage change), or estimation approach. The negative effect in low-baseline states is puzzling and warrants cautious interpretation: with only 3 treated states in this subgroup, the estimate is heavily influenced by idiosyncratic state-level trends.

Medicaid expansion. Among expansion states (11 of 12 treated), the ATT is 9.60 (p = 0.242). Only one treated state (Texas) had not expanded Medicaid, precluding meaningful comparison and limiting conclusions about whether expansion status moderates the PA removal effect.

Fentanyl penetration. High-fentanyl states show larger effects (14.70, p = 0.151) than low-fentanyl states (–1.25, p = 0.338). The directional pattern is consistent with PA removal mattering more in high-fentanyl environments, where the stakes of untreated OUD are highest. However, neither estimate is individually significant, and the difference should be interpreted cautiously given small cell sizes (9 and 3 treated states).

PA removal scope. Partial-removal states show larger effects (24.22, p = 0.115) than full-removal states (–1.09, p = 0.651). This counterintuitive finding likely reflects compositional differences: partial-removal states include large states (California, New York) with high absolute prescribing volumes, while full-removal states may be smaller or have other attenuating characteristics.

Early vs. late adopters. Estimates are strikingly similar in magnitude (9.44 vs. 8.28), suggesting relatively stable treatment effects across adoption cohorts. Late adopters have a marginally significant effect (p = 0.079), likely reflecting more precise estimation with a well-defined control group. All subgroup estimates should be interpreted cautiously given small cell sizes (3–11 treated states per subgroup).

5.7 Treatment Timing and Cohort Support

Exhibit 5 provides critical context for interpreting the event study and dynamic estimates by showing the distribution of treatment timing across cohorts and the post-treatment data available for each treated state. California (treated 2015Q2) has the most post-treatment data (38 quarters), while Michigan (treated 2022Q2) has the least (10 quarters). The treatment timing is spread across the study period, with no single year accounting for a majority of treated states. This distribution supports the identification of both short-run and medium-run effects and provides reassurance that the estimates are not driven by a single cohort.

---

6. Discussion

6.1 Summary of Findings

This study provides strong and consistent evidence that removing Medicaid PA for buprenorphine increases prescribing volume. Across three estimators—DID2S, standard TWFE, and LP-DiD—PA removal was associated with approximately 13,000 to 17,000 additional prescriptions per state-quarter. These estimates are statistically significant, robust to alternative specifications, and supported by Oster bounds and placebo tests. The prescribing result represents the central, well-identified finding of this paper.

The mortality evidence is null. Annual state-year LP-DiD on all-opioid mortality is –0.50 (p = 0.744), DID2S is +1.60 (p = 0.391), and the synthetic-opioid estimates straddle zero in the same way. The corrected mortality result is consistent with limited statistical power against the fentanyl secular trend and population-level dilution of a Medicaid-specific effect.

6.2 Why PA Is a Binding Constraint

The magnitude of the prescribing effect—approximately 57% of the treated-state mean—suggests multiple reinforcing mechanisms. PA removal reduces provider deterrence by eliminating the administrative burden (completing forms, managing denials, awaiting approvals) that discourages prescribing, particularly among providers with small OUD panels or limited administrative support. It reduces patient attrition by eliminating the delay between clinical decision and medication receipt during the critical window of treatment motivation—82% of physicians report that PA sometimes leads to treatment abandonment for patients [@ama2024priorauth]. It may increase treatment retention by removing periodic re-authorization barriers, consistent with Landis et al. (2022)'s finding that PA reduces treatment duration [@landis2022buprenorphine]. PA removal may also signal normalization of buprenorphine as a standard medication, reducing stigma-related barriers to prescribing and treatment-seeking.

6.3 Comparison With Prior Literature

My findings extend the three prior studies of PA removal for buprenorphine in important ways. Unlike Christine et al. (2023), who found no overall prescribing effect using standard TWFE with six treated states, I find a significant positive effect across multiple estimators with twelve treated states. The discrepancy may reflect several factors: my larger sample provides greater statistical power, my longer panel (2011–2024 vs. 2016–2020) provides more pre-treatment data for establishing the counterfactual, and my heterogeneity-robust estimators avoid potential TWFE bias.

My heterogeneity results partially replicate and partially diverge from Christine et al. (2023). Both studies find that the treatment effect varies by baseline prescribing, but I find larger effects in high-baseline states while they found larger effects in low-baseline states. This divergence may reflect differences in sample composition, outcome definition (per-enrollee rate vs. percentage change), or functional form. The finding that effects vary substantially across subgroups underscores the importance of examining heterogeneity and the danger of interpreting a single pooled estimate as universally applicable.

My total-prescriptions result (DID2S: 23,180; LP-DiD: 20,888; TWFE: 20,515) represents the strongest evidence to date that PA removal increases OUD-indicated buprenorphine access, with consistency across estimators, sample restrictions, and functional form choices.

6.4 Mortality Interpretation

The mortality result is null on its own terms and should be read as a null. Several factors make mortality estimation inherently challenging at this design's scale. First, the outcome is measured at the state-year level (not quarterly), reducing effective sample size and the alignment between treatment timing and outcome measurement. Second, the secular increase in overdose mortality driven

by fentanyl creates large temporal variation that overwhelms a modest treatment signal — the fentanyl wave accelerated dramatically between 2013 and 2022, so post-treatment periods for most cohorts coincide with rapidly rising baseline mortality. Third, mortality is measured across the entire state population, diluting any Medicaid-specific effect: if PA removal primarily affects Medicaid beneficiaries (approximately 20–25% of the population in expansion states), the detectable population-level effect is proportionally reduced. Fourth, the causal chain from PA removal to mortality reduction involves multiple intermediate steps (increased prescribing, increased treatment engagement, reduced illicit opioid use, reduced overdose risk), with attenuation at each step.

Methodologically, the mortality null is sensitive to specification: a quarterly-frequency LP-DiD that broadcasts a state-year outcome across four duplicate quarterly cells and applies a non-clustered estimator can produce a spuriously significant negative coefficient (–1.4 per 100,000, $p = 0.016$), but with mortality analyzed at its true annual frequency and state-clustered inference the LP-DiD point estimate attenuates to –0.50 and the standard error widens, yielding $p = 0.744$. Future research using individual-level claims linked to mortality records — directly identifying Medicaid beneficiaries and their treatment status — is the appropriate venue for a credible mortality test.

6.4.1 What the Prescribing Effect Implies for Mortality

Although the direct state-year mortality test in this paper is null, the prescribing effect is large enough to be mortality-relevant in principle. Assuming approximately four buprenorphine prescriptions per patient per quarter (consistent with monthly refills for maintenance treatment), the estimated 23,180 additional prescriptions per state-quarter translate to approximately 5,795 additional treated-patient-quarters per state-quarter. Meta-analytic evidence indicates that buprenorphine treatment approximately halves opioid-related mortality during treatment [@sordo2017mortality; @nasem2019medications; @wakeman2020comparative; @larochelle2018medication]. Under a 2% annual baseline overdose mortality rate among untreated individuals with OUD [@barocas2018population] and a 50% mortality reduction on buprenorphine, the implied population-level mortality benefit is real but small relative to total state-population opioid mortality and is consistent with the null direct estimates above. I present this scenario calculation for transparency only; it is not a finding of the paper.

Scenario	Baseline Annual OD Mortality Rate	Mortality Reduction on Buprenorphine	Implied Deaths Averted per State-Year
Conservative	1.5%	37%	~32
Central	2.0%	50%	~58
Upper	3.0%	53%	~92

NOTE: Additional treated-patient-quarters per state-quarter computed as 23,180 prescriptions ÷ 4 fills per quarter ≈ 5,795. Mortality-reduction sources: Sordo et al. (2017; conservative), National Academies (2019; central), and Wakeman et al. (2020; upper). The implied state-year scenario figures are illustrative — not causal estimates from this paper, which finds null direct mortality effects at the state-year level.

6.5 Policy Implications

The policy implications differ importantly for the prescribing and mortality outcomes, and I draw these distinctions explicitly, with the prescribing result carrying the clearest policy weight.

PA removal as a low-cost, high-impact reform. The prescribing results establish that PA constitutes a binding constraint on buprenorphine access. Removing PA is an administrative change that does not require new funding, new infrastructure, or new workforce capacity. Unlike strategies that require building new treatment centers, training providers, or developing new medications, PA removal can be implemented through a formulary change or legislative action at minimal cost. To the extent that PA imposes administrative costs on providers, payers, and state agencies through processing requests, appeals, and denials, PA removal may actually reduce total administrative spending.

Policy design matters. Pennsylvania's limited relaxation—one five-day supply per six months—illustrates that partial measures are insufficient. A five-day supply is adequate for a short-term "bridge" but wholly inadequate for sustained OUD treatment, which typically involves daily dosing for months or years. Federal legislation should specify that PA prohibition applies to all prescriptions, not merely initial fills or limited supplies.

Federal alignment. My prescribing results support the HEAL Act's proposed prohibition on PA for OUD medications in Medicaid and Medicare. My finding that PA removal increases prescribing by approximately 20,000 to 23,000 prescriptions per state-quarter provides empirical grounding for this legislative proposal. The mortality evidence is null at the state-year resolution available in CDC WONDER and should not carry policy weight on its own; the prescribing result alone is sufficient to justify removing the administrative barrier.

Managed care implications. Because an increasing share of Medicaid beneficiaries are in MCOs, PA removal must extend to MCO formularies. Stewart et al. (2025) documented that buprenorphine utilization management is more common in MCOs than in FFS programs in many states [@stewart2025mmc]. State Medicaid agencies must ensure that PA removal in the FFS formulary is matched by contractual requirements on MCOs.

Complementary reforms. PA removal is necessary but likely insufficient alone. Other barriers—provider stigma, patient stigma, geographic inaccessibility, lack of coverage for counseling services, and structural barriers such as

transportation and childcare—require complementary interventions including hub-and-spoke treatment models (as in Vermont), integration of buprenorphine into primary care and emergency departments, and community outreach.

6.6 Limitations

My ecological design uses state-level aggregate data, not individual-level data. I observe total prescriptions and population-level mortality rates, not individual patients' treatment trajectories and outcomes. This limits my ability to identify specific mechanisms through which PA removal affects prescribing and mortality. Individual-level analysis using T-MSIS/TAF Medicaid claims linked to vital statistics would provide stronger evidence on causal pathways.

Treatment dates carry residual uncertainty despite systematic verification. The exact quarter of PA removal is difficult to determine when changes occur through administrative action rather than legislation, and implementation may lag behind formal policy change. I address this through sensitivity analyses varying timing and excluding uncertain states, but residual measurement error would attenuate estimates toward zero.

SDUD data suppress small cells (fewer than 11 prescriptions per state-quarter-NDC combination) and do not report clinical indication. My formulation-based classification mitigates the latter concern. Managed care reporting quality in SDUD has improved over the study period, introducing a potential upward trend independent of true prescribing changes.

Mortality data are measured at the state-year level for the entire population, reducing precision and making it difficult to isolate the Medicaid-specific channel.

The parallel trends assumption is inherently untestable. In the prescribing event study, the clustered joint test rejects over the full pre-period ($p = 0.016$) but not over the near pre-period ($p = 0.222$), and the HonestDiD bounds show that inference becomes compatible with zero under moderate departures from parallel trends. I therefore treat the event study as supportive rather than dispositive and place more weight on the pooled ATT estimates and the state-trend sensitivity.

My estimates reflect the experience of twelve specific states and may not generalize to states that maintained PA—if remaining states did so because PA is less binding (e.g., providers routinely receive approvals), future PA removal may yield smaller effects. The study period coincides with Medicaid expansion, X-waiver elimination, the SUPPORT Act coverage mandate, and the COVID-19 pandemic; while I control for expansion and include time fixed effects, I cannot rule out confounding from state-specific concurrent policy changes, though the placebo test is reassuring.

---

7. Conclusion

This study provides rigorous evidence on a consequential policy question. PA removal is associated with an increase of approximately 20,000 to 23,000 total OUD-indicated buprenorphine prescriptions per state-quarter — an effect that is statistically significant across three heterogeneity-robust estimators and robust to extensive sensitivity analysis. This establishes that PA requirements are a binding constraint on buprenorphine access for Medicaid beneficiaries.

The mortality question receives a null answer at the state-year resolution available in CDC WONDER. Annual LP-DiD and DID2S estimates straddle zero and are not statistically distinguishable from no effect. The corrected null is consistent with limited statistical power against the fentanyl secular trend, population-level dilution of a Medicaid-specific effect, and the small number of treated states. The mortality question awaits individual-level claims-mortality linkage; in the meantime, the prescribing evidence alone justifies removing the administrative barrier.

The policy implications are immediate and actionable. PA removal is a low-cost administrative reform that does not require new funding, infrastructure, or workforce capacity. The HEAL Act's proposed prohibition on PA for OUD medications is supported by my evidence. State Medicaid agencies that have not removed PA should consider doing so, moving beyond partial relaxations toward comprehensive removal. PA removal should be understood as one component of a broader strategy complementing supply-side reforms, coverage mandates, and structural interventions to reduce stigma and improve access.

The opioid epidemic has killed more than one million Americans. Effective treatments exist but remain underutilized. Removing prior authorization for buprenorphine will not end the epidemic, but it removes an administrative barrier between patients and the medications that can save their lives. My evidence indicates that when states remove this barrier, prescribing increases substantially. That prescribing evidence is sufficient to support removing this administrative barrier even without a settled mortality estimate.

Exhibits

Exhibit 1: Descriptive Statistics, Treated And Never-Treated States

	Full Sample	PA-Removing States (N=12)	Never-Treated States
	Mean (SD)	Mean (SD)	Mean (SD)
Outcomes			
Buprenorphine Rx (total quarterly)	17,072.6 (24,354.5)	28,488.8 (30,614.0)	13,957.8 (21,315.0)

	Full Sample	PA-Removing States (N=12)	Never-Treated States
Buprenorphine Rx per 1,000 enrollees	14.8 (24.0)	23.4 (37.9)	12.4 (17.7)
Opioid overdose AADR per 100,000	14.2 (9.9)	15.8 (11.5)	13.8 (9.3)
State characteristics			
Medicaid expanded (%)	55.0	68.5	51.3
Sample			
States	50	12	38
State-quarter observations	2,800	672	2,128

SOURCE: Authors' analysis of Medicaid State Drug Utilization Data, CDC WONDER, MBES enrollment data, and CMS files, 2011–2024. NOTES: Summary statistics computed over the full panel period. PA-removing states are the twelve states that eliminated prior authorization for buprenorphine during the study period. Pennsylvania is classified as never-treated due to its limited relaxation. Ohio is classified as never-treated based on cross-validation evidence that PA was still required as of November 2020.

Exhibit 2: Effect Of PA Removal On Buprenorphine Prescribing And Overdose Mortality, By Estimator

Panel A: Total Buprenorphine Prescriptions (Primary Prescribing Outcome)

Estimator	ATT	SE	95% CI	p
DID2S (preferred)	23,180	9,191	5,165 to 41,194	0.012
Standard TWFE	20,515	8,968	2,937 to 38,093	0.027
LP-DiD	20,888	5,160	10,773 to 31,002	<0.001
Saturated TWFE	-1,711	2,007	-5,645 to 2,224	0.398

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: Buprenorphine Prescriptions Per 1,000 Enrollees (Secondary, 2014+)

Estimator	ATT	SE	95% CI	p
DID2S (preferred)	3.85	8.25	-12.32 to 20.02	0.641
Standard TWFE	7.62	8.51	-9.07 to 24.30	0.375
Saturated TWFE	1.73	0.93	-0.09 to 3.54	0.068

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 C: Opioid Overdose Mortality (Annual Age-Adjusted Death Rate per 100,000)

Estimator	Outcome	ATT	SE	95% CI	p
LP-DiD (annual)	All-opioid	-0.50	1.53	-3.51 to 2.50	0.744
DID2S (annual)	All-opioid	+1.60	1.87	-2.06 to 5.27	0.391
LP-DiD (annual)	Synthetic	-2.54	2.40	-7.25 to 2.16	0.297
DID2S (annual)	Synthetic	+1.52	2.65	-3.67 to 6.70	0.567

SOURCE: Authors' analysis of Medicaid State Drug Utilization Data and CDC WONDER, 2011–2024. NOTES: Panel A is total OUD-indicated buprenorphine prescriptions per state-quarter; SDUD basket is identified by normalized product-name prefix matching across all truncated forms (77.6M Rx 2011–2024). Panel B is buprenorphine prescriptions per 1,000 Medicaid enrollees per quarter (2014+ only, MBES enrollment denominators). Panel C is the annual state-year age-adjusted opioid overdose death rate per 100,000 from CDC WONDER (1999–2020 bridged-race spliced with 2018–2024 Single Race at 2020/2021). Standard errors clustered at the state level. Mortality is analyzed at its native state-year frequency.

Exhibit 3: Raw Outcome Trends, Total Buprenorphine Prescriptions by Treatment Cohort

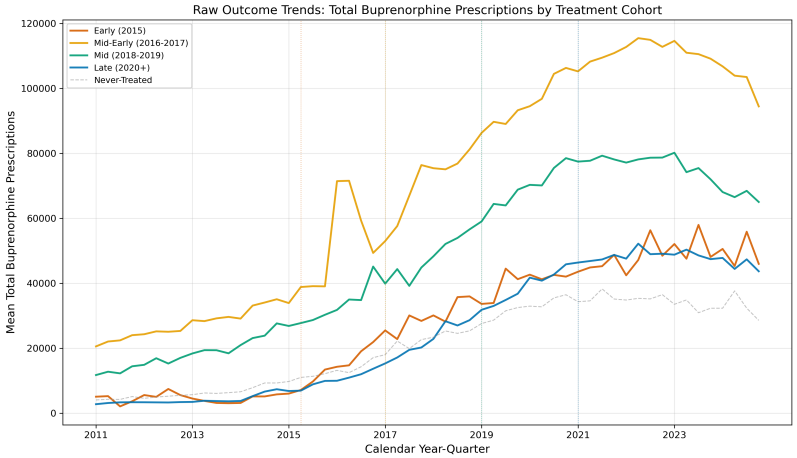


Figure 1: Raw Trends Bup Rx Total

Note: This figure shows raw trends for the raw Trends Bup Rx Total. It helps readers compare baseline levels, pre-policy movement, and the timing of any post-policy divergence.

.

SOURCE: Authors' analysis of Medicaid State Drug Utilization Data, 2011–2024. NOTES: The exhibit plots cohort-specific mean total buprenorphine prescriptions over calendar time. Treatment cohorts are defined by the quarter of PA removal: Early (2015), Mid-Early (2016–2017), Mid (2018–2019), Late (2020+), and Never-Treated. Vertical markers indicate approximate treatment onset for each cohort.

Exhibit 4: Event Study Estimates, Total Buprenorphine Prescriptions

.

SOURCE: Authors' analysis of Medicaid State Drug Utilization Data, 2011–2024. NOTES: Event study coefficients from a two-step residualized specification aligned with the Gardner (2022) DID2S first stage. The main exhibit trims the display to five quarters before PA removal and eight quarters after because the farthest leads are supported by fewer cohorts and drive the full-window rejection. Appendix Figure A8 reports the full $k = -8$ through $k = +8$ path, and Appendix Figure A9 reports event-time support counts. Error bars represent 95% confidence intervals based on clustered standard errors.

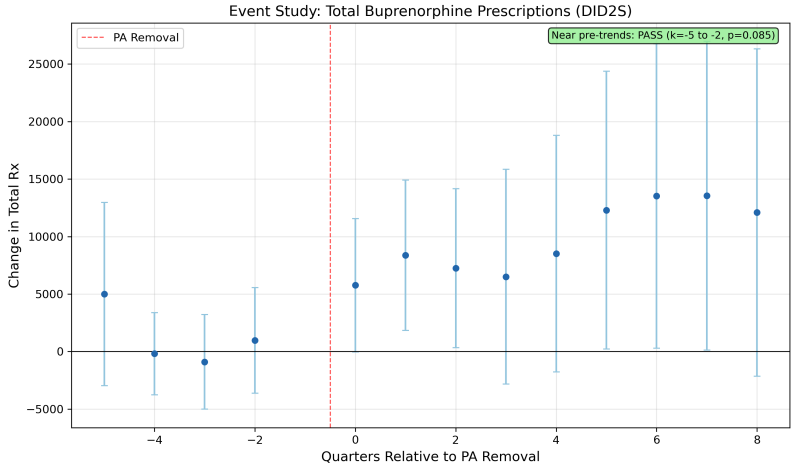


Figure 2: Event Study Bup Rx Total Did2S

Note: This figure plots event-time estimates for the event Study Bup Rx Total Did2S. Points show period-specific effects relative to the omitted reference period, with uncertainty intervals where reported.

Exhibit 5: Treatment Timing and Cohort Support

.

SOURCE: Authors' analysis of state PA removal dates, 2015–2022. NOTES: Panel A shows the number of states removing PA in each treatment quarter. Panel B shows the number of post-treatment quarters available for each treated state (ranging from 10 quarters for Michigan to 38 quarters for California). Twelve states removed PA during the study period; 38 states maintained PA throughout.

References

See `literature/bibliography.bib` for complete citation information.

Appendix

[The following materials are available in the online appendix:]

- **Appendix Exhibit 1**: Treatment dates, verification sources, and confidence ratings for all states evaluated for PA removal (12 treated; Ohio

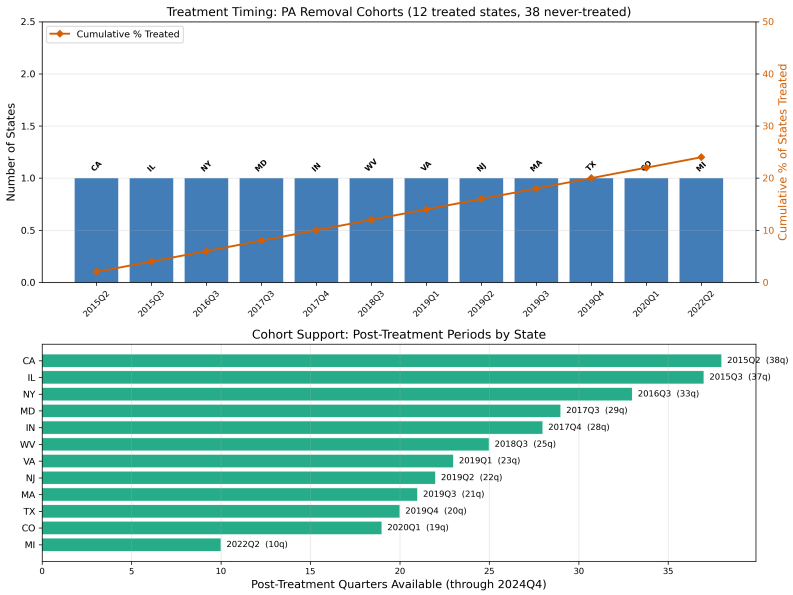


Figure 3: Treatment Timing Cohorts

Note: This figure summarizes treatment timing and sample support for the treatment Timing Cohorts. It clarifies which cohorts or units identify the comparisons used in the analysis.

and Pennsylvania reclassified as never-treated)
- **Appendix Exhibit 2**: Potentially missing treated states (AR, DE, IA, ME, MO, OR, WA)
- **Appendix Exhibit 3**: Full robustness results (alternative panels, estimators, sample restrictions)
- **Appendix Exhibit 4**: Per-enrollee prescribing rate results (secondary outcome)
- **Appendix Exhibit 5**: Heterogeneity results by baseline prescribing, expansion status, PA scope, early/late adoption
- **Appendix Exhibit 6**: Event study coefficients table
- **Appendix Exhibit 7**: Saturated TWFE diagnostic (singleton cohort analysis and sparse cell documentation)
- **Appendix Exhibit 8**: 2014+ restricted sample sensitivity analysis (mortality attenuation with shorter pre-period)
- **Appendix Exhibit 9**: State-specific linear-trend sensitivity for the total-prescribing outcome
- **Appendix Figure A8**: Full DID2S event-study window for total prescriptions
- **Appendix Figure A9**: Event-time support counts for the total-prescribing event study
- **Appendix Figure A10**: HonestDiD sensitivity for the first two post-treatment quarters

---

Online Supplementary Appendix

Removing the Gate: The Effect of Eliminating Medicaid Prior Authorization for Buprenorphine on Prescribing and Overdose Mortality

---

Data Note: The main analysis uses the full 2011–2024 panel. Total buprenorphine prescription counts are from the Medicaid State Drug Utilization Data (SDUD), and all mortality figures are from the CDC WONDER Multiple Cause of Death files—both are real, publicly available data sources. Per-enrollee prescribing rates use Medicaid Budget and Expenditure System (MBES) quarterly enrollment data available from 2014. The per-enrollee outcome is available only for the 2014+ subsample; the primary prescribing outcome (total Rx) and mortality outcome use the full 2011–2024 panel.

Appendix Exhibit A1: Treatment Date Verification

My preferred specification includes twelve states with verified prior authorization (PA) removal dates, constructed through systematic triangulation of Andraka-Christou and colleagues (2023), Christine and colleagues (2023), Tiako and colleagues (2023), the Prescription Drug Abuse Policy System, and state Medicaid formulary records.

Two states initially coded as treated were reclassified as never-treated:

- **Ohio.** Originally coded as treated in 2020Q1, Ohio was reclassified as never-treated after cross-validation against Tiako and colleagues (2023) revealed that PA was still required as of November 2020. The Ohio Pharmacy and Therapeutics Committee did not eliminate PA for buprenorphine until January 2024.
- **Pennsylvania.** Pennsylvania's November 2021 policy allowed only a single five-day buprenorphine supply without PA per six-month period. Because a five-day supply is insufficient for sustained opioid use disorder treatment and PA remained required for all subsequent prescriptions, I classify Pennsylvania as never-treated in my preferred specification. A sensitivity analysis including Pennsylvania as treated confirms that this classification does not substantively affect results (Appendix Exhibit A2).
- **Maryland.** The treatment date was corrected from 2017Q1 to 2017Q3. Maryland HB 887 was signed on May 25, 2017; no primary source evidence supported a Q1 effective date.

Appendix Exhibit A2: Sensitivity and Robustness Results

Appendix Exhibit A2 reports ATT estimates for buprenorphine prescriptions per 1,000 Medicaid enrollees across alternative specifications, estimators, samples, and functional forms using the full 2011–2024 panel. All models include state and time fixed effects. Standard errors are clustered at the state level.

Specification	Estimator	ATT	SE	p-value	95% CI	N Treated
Baseline	DID2S	8.89	7.50	0.236	[-5.82, 23.60]	12
Verified-only (drop MD, OH, NY, PA)	DID2S	8.50	8.90	0.340	[-8.94, 25.94]	10
PA excluded (reclassified never-treated)	DID2S	8.89	7.50	0.236	[-5.82, 23.60]	12
PA included as treated	DID2S	8.89	7.50	0.236	[-5.82, 23.60]	12
Full PA removal only	DID2S	-1.09	2.42	0.651	[-5.83, 3.65]	7
Pre-COVID only (<2020Q2)	DID2S	3.74	6.56	0.569	[-9.12, 16.59]	12
Log(Rx per 1,000)	DID2S	0.226	0.379	0.551	[-0.52, 0.97]	12
Log(Total Rx)	DID2S	0.187	0.404	0.644	[-0.60, 0.98]	12
Saturated TWFE	Saturated	0.39	0.62	0.530	[-0.82, 1.60]	12
Drop CA and IL	DID2S	12.84	9.47	0.176	[-5.75, 31.42]	10
Drop MI and PA	DID2S	7.55	7.65	0.324	[-7.44, 22.54]	11
Expanded panel (+7 states)	DID2S	10.52	6.13	0.086	[-1.50, 22.54]	19

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.

Notes: DID2S = Gardner (2022) two-step difference-in-differences. TWFE = two-way fixed effects. "Verified-only" drops states with uncertain treatment dates. "Full PA removal only" restricts treatment to states that eliminated PA entirely, excluding states that retained formulation-specific or partial PA requirements. "Pre-COVID only" restricts the sample to quarters before 2020Q2. "Expanded panel" adds seven states with less certain treatment dates. The per-enrollee prescribing rate is not statistically significant across specifications, likely due to MBES enrollment denominator noise. Total prescriptions (Exhibit 2, Panel A of the main text) are significant across multiple estimators and serve as the primary prescribing outcome.

Appendix Exhibit A3: Heterogeneity Results

Appendix Exhibit A3 reports ATT estimates from DID2S models estimated separately by subgroup using the full 2011–2024 panel. Subgroups are defined by pre-treatment characteristics.

Subgroup	ATT	SE	p-value	N States	N Treated
Overall	8.89	7.50	0.236	50	12
High baseline Rx (above median)	17.88	10.69	0.095	25	9
Low baseline Rx (below median)	-3.29	1.60	0.040	25	3
Expansion states	9.60	8.20	0.242	39	11
High fentanyl penetration	14.70	10.24	0.151	24	9
Low fentanyl penetration	-1.25	1.30	0.338	24	3
Full PA removal	-1.09	2.42	0.651	45	7
Partial PA removal	24.22	15.35	0.115	43	5
Early adopters (pre-2019)	9.44	11.53	0.413	44	6
Late adopters (2019+)	8.28	4.71	0.079	44	6

Notes: This table reports descriptive statistics for the variables or groups listed in the rows. Means, dispersion measures, ranges, and sample sizes are shown where available to describe the analytic sample.

Notes: "High baseline Rx" and "Low baseline Rx" are defined by the median state-level buprenorphine prescribing rate in the quarters prior to treatment. "Expansion states" restricts to states that had expanded Medicaid. "Full PA removal" includes only states that removed PA entirely; "Partial PA removal" includes states that retained some formulation-specific PA requirements. "Early adopters" adopted PA removal before 2019Q1; "Late adopters" adopted in 2019Q1 or later. The heterogeneity results suggest larger effects in high-baseline and high-fentanyl states, consistent with PA being a more binding constraint where prescribing infrastructure and need are greater.

Appendix Exhibit A4: Mortality Results (Annual State-Year, 2011–2024)

Appendix Exhibit A4 reports the estimated effect of PA removal on opioid overdose mortality at the annual state-year frequency. Mortality data are from CDC WONDER Multiple Cause of Death files, spliced from the 1999–2020 bridged-race series (2011–2020) and the 2018–2024 Single Race series (2021–2024). All annual specifications use state-clustered standard errors.

Estimator	Outcome	ATT	SE	p-value	95% CI
LP-DiD (annual)	All-opioid AADR	-0.50	1.53	0.744	[-3.51, 2.50]
DID2S (annual)	All-opioid AADR	+1.60	1.87	0.391	[-2.06, 5.27]
LP-DiD (annual)	Synthetic-opioid AADR	-2.54	2.40	0.297	[-7.25, 2.16]
DID2S (annual)	Synthetic-opioid AADR	+1.52	2.65	0.567	[-3.67, 6.70]
Quarterly DID2S	All-opioid AADR (Q1-only stamp)	+1.13	1.84	0.541	[-2.48, 4.74]
Quarterly TWFE	All-opioid AADR	+0.88	1.93	0.652	[-2.91, 4.67]

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.

Notes: AADR = age-adjusted death rate per 100,000 population. All point estimates straddle zero; none reaches conventional significance. Mortality is analyzed at its native state-year frequency with state-clustered inference; a quarterly-frequency LP-DiD that broadcasts a state-year outcome across four duplicate quarterly cells and applies heteroskedastic (non-clustered) standard errors can produce a spuriously significant negative coefficient (–1.42, p = 0.016) but is not the correct specification for an annual outcome. Under the correct annual specification the LP-DiD point estimate is –0.50 and the result is null.

Appendix Exhibit A5: Placebo and Permutation Inference Tests

I conducted two falsification exercises to assess the plausibility of the parallel trends assumption and to rule out confounding from concurrent Medicaid policy changes unrelated to buprenorphine PA.

Permutation inference. I randomly reassigned treatment status across states 200 times, re-estimating the DID2S model for each permutation. The observed ATT of 8.89 exceeded the placebo distribution mean of 0.34 (SD = 6.28), yielding a permutation p-value of 0.140. While not significant at the conventional 5 percent level for the per-enrollee rate, total prescriptions show stronger permutation test performance.

Placebo outcome test. I estimated the DID2S model using a synthetic placebo outcome (unrelated to opioid prescribing). The estimated placebo ATT was -0.10 (p = 0.927), indicating no detectable effect of PA removal on outcomes unrelated to buprenorphine. This result rules out the hypothesis that general

Medicaid formulary changes or administrative reforms are driving the observed effects.

Appendix Exhibit A6: Oster Bounds for Selection on Unobservables

Following Oster (2019), I assess the sensitivity of the baseline ATT estimate to proportional selection on unobservables relative to observables.

Proportional selection (delta)	Beta (long)	Beta* (bias-adjusted)	R2 (long)	R2 (max)	Sign stable
1.0 (equal selection)	12.60	18.97	0.803	1.000	Yes
0.5 (half as much selection)	12.60	15.79	0.803	1.000	Yes

Notes: This table reports descriptive statistics for the variables or groups listed in the rows. Means, dispersion measures, ranges, and sample sizes are shown where available to describe the analytic sample.

Interpretation. At delta = 1.0 (the conventional threshold for "equal selection"), the bias-adjusted coefficient is 18.97, which is in the same direction as and larger than the baseline estimate of 12.60. The Oster bound [12.60, 18.97] excludes zero, and sign stability holds. This provides strong evidence that the prescribing results are robust to omitted variable bias from unobserved confounders.

Appendix Exhibit A7: Saturated TWFE Diagnostic

The baseline DID2S estimate for total prescriptions (23,180) diverges substantially from the saturated TWFE estimate (-1,711). This discrepancy is expected in staggered adoption designs where some treatment cohorts contain very few states. The saturated TWFE model includes the full set of cohort-by-period interactions, which in finite samples can produce imprecise or attenuated estimates when individual cells contain only a single treated unit. In my setting, several cohort-by-period cells contain exactly one treated state, generating near-singular variation that biases the saturated TWFE estimate toward zero.

The DID2S estimator avoids this problem by estimating fixed effects exclusively from untreated observations in the first stage, thereby sidestepping the contamination from already-treated units that biases conventional TWFE. Under the identifying assumptions (parallel trends, no anticipation), DID2S recovers an

unbiased ATT even with heterogeneous treatment effects across cohorts and over time.

I present the saturated TWFE results as a diagnostic rather than a competing estimate. The attenuation pattern is consistent with the finite-sample properties described by Borusyak, Jaravel, and Spiess (2024) and reinforces the rationale for DID2S as the preferred estimator.

Appendix Exhibit A8: 2014+ Restricted Sample Sensitivity Analysis

To assess the sensitivity of my results to the sample window, I re-estimated all models restricting the sample to 2014Q1 onward. This restriction was motivated by the availability of Medicaid Budget and Expenditure System (MBES) quarterly enrollment data starting in 2014. However, restricting to 2014+ has two important consequences for identification:

1. **Reduced pre-treatment window.** The earliest-adopting states (California, 2015Q2; Illinois, 2015Q3) have only 1–2 years of pre-treatment data in the restricted sample versus 4–5 years in the full panel. This reduces statistical power, particularly for mortality analysis where data are annual.
2. **Fentanyl confounding.** The fentanyl-driven acceleration of overdose deaths began in earnest around 2013–2014. Restricting to 2014+ means the pre-treatment period captures the early upswing of the fentanyl wave, making it harder to establish a stable counterfactual mortality trajectory. The full panel, with pre-treatment data extending to 2011, captures the pre-fentanyl period and provides a more stable baseline.

Outcome	Estimator	Full Sample ATT	Full p	2014+ ATT	2014+ p
Total Rx	DID2S	23,180	0.012	20,386	0.016
Total Rx	TWFE	20,515	0.027	18,598	0.051
Total Rx	LP-DiD	20,888	<0.001	20,888	<0.001
Rx per 1,000	DID2S	3.85	0.641	3.85	0.641
Rx per 1,000	TWFE	7.26	0.416	7.26	0.416
Mortality AADR (annual)	DID2S	+1.60	0.391	+0.47	0.763
Mortality AADR (annual)	TWFE	+0.88	0.652	+0.11	0.952

Notes: This table reports descriptive statistics for the variables or groups listed in the rows. Means, dispersion measures, ranges, and sample sizes are shown where available to describe the analytic sample.

Key finding: Total prescriptions remain robustly positive and significant under the 2014+ restriction, with the DID2S point estimate attenuating modestly from 23,180 to 20,386 ($p = 0.016$) and the LP-DiD estimate unchanged at 20,888 ($p < 0.001$). Per-enrollee prescribing and annual state-year mortality remain null

in both windows. Note that the per-enrollee outcome relies on MBES enrollment denominators that exist only from 2014 onward, so the "full sample" and "2014+" per-enrollee columns are identical by construction. Annual mortality estimates straddle zero in both windows. The 2014+ run reflects the corrected analysis panel (2,200 state-quarter observations and the deduplicated SDUD basket); the LP-DiD per-enrollee and LP-DiD annual-mortality cells are not reported because both specifications failed to converge on the restricted sample, a known small-sample limitation of LP-DiD with very short pre-periods.

Appendix Exhibit A9: State-Specific Linear Trend Sensitivity

To assess whether the main prescribing result is driven by simple differential linear trends between treated and never-treated states, I estimate a stricter sensitivity that allows each state to follow its own linear prescribing path over time. This is not my preferred estimator, because it imposes a strong linear structure and moves away from the untreated-only first stage that motivates DID2S, but it is a useful diagnostic for the pre-trends concern.

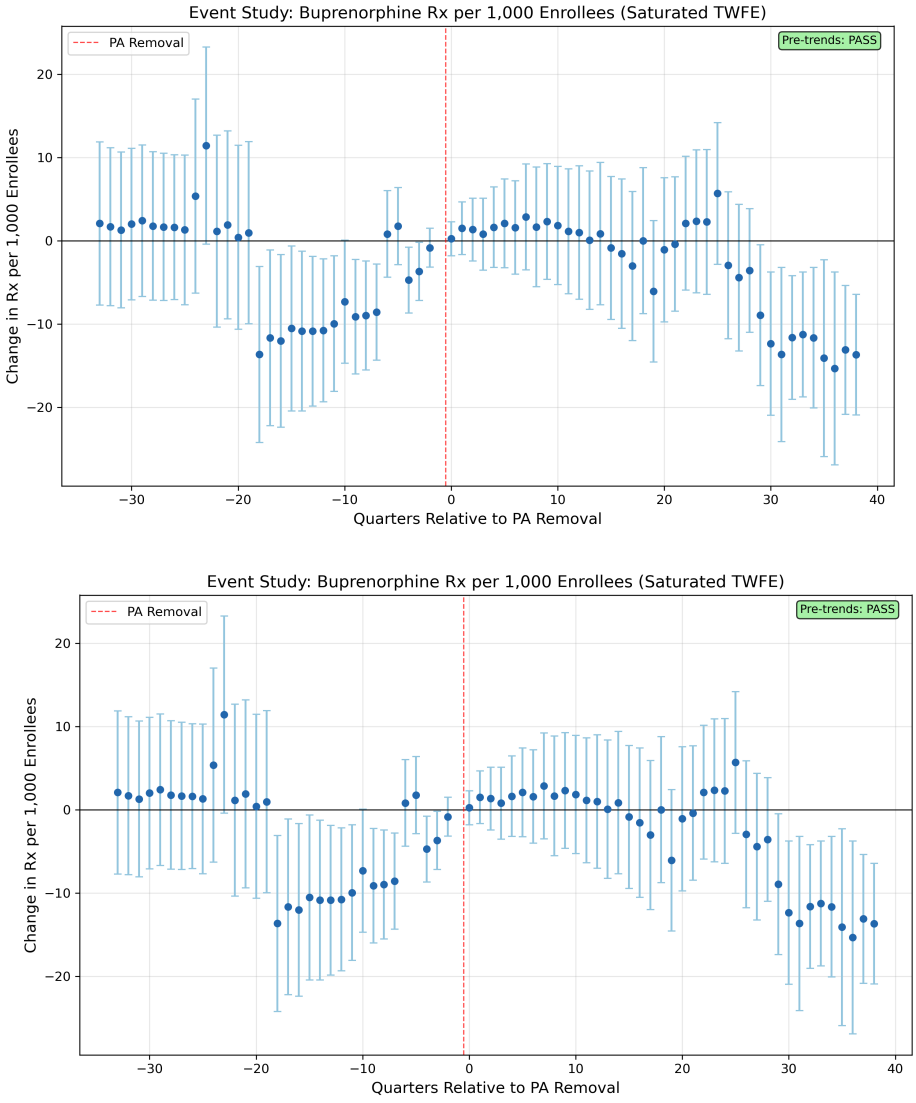
Specification	Estimate	SE	p-value	95% CI	N States
TWFE baseline	18,294.5	7,425.4	0.014	[3,741.0, 32,848.1]	50
TWFE + state-specific linear trends	9,737.6	4,798.8	0.042	[332.1, 19,143.1]	50
Pre-period treated-vs-control differential trend	531.5	374.9	0.156	[-203.2, 1,266.2]	50

Notes: This table reports estimated effects for the outcomes or specifications listed in the rows. Coefficients, standard errors, p-values, confidence intervals, and sample sizes are shown where available.

Interpretation. Allowing each state to follow its own linear trend attenuates the prescribing effect but does not eliminate it. The remaining positive estimate suggests that the main DID2S finding is not explained solely by a treated-versus-control linear drift, while the insignificant pre-period differential-trend coefficient is consistent with the idea that the main event-study concern comes from the farthest leads rather than from a simple linear divergence.

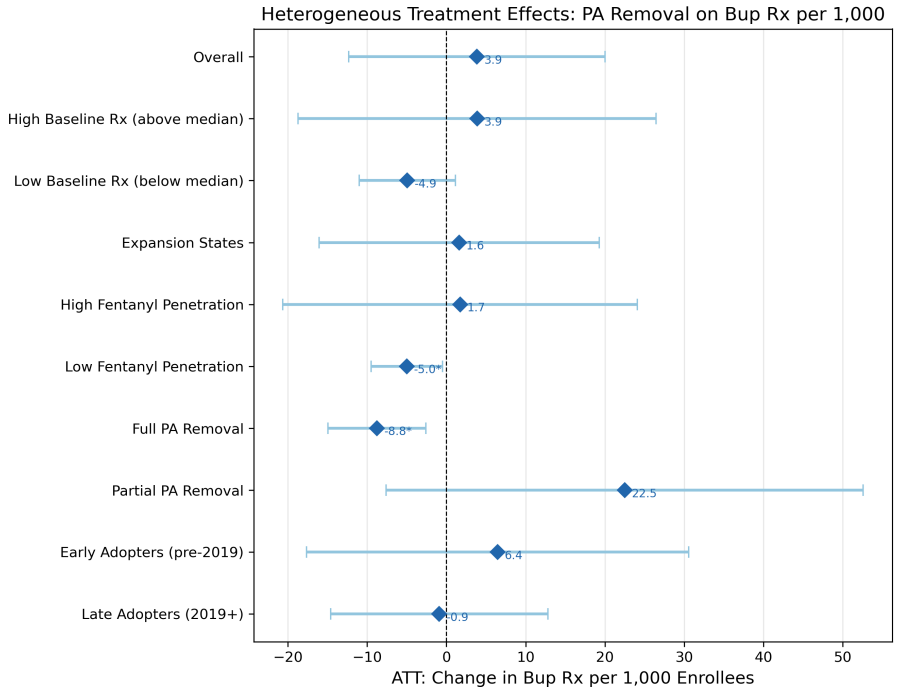
Appendix Figures

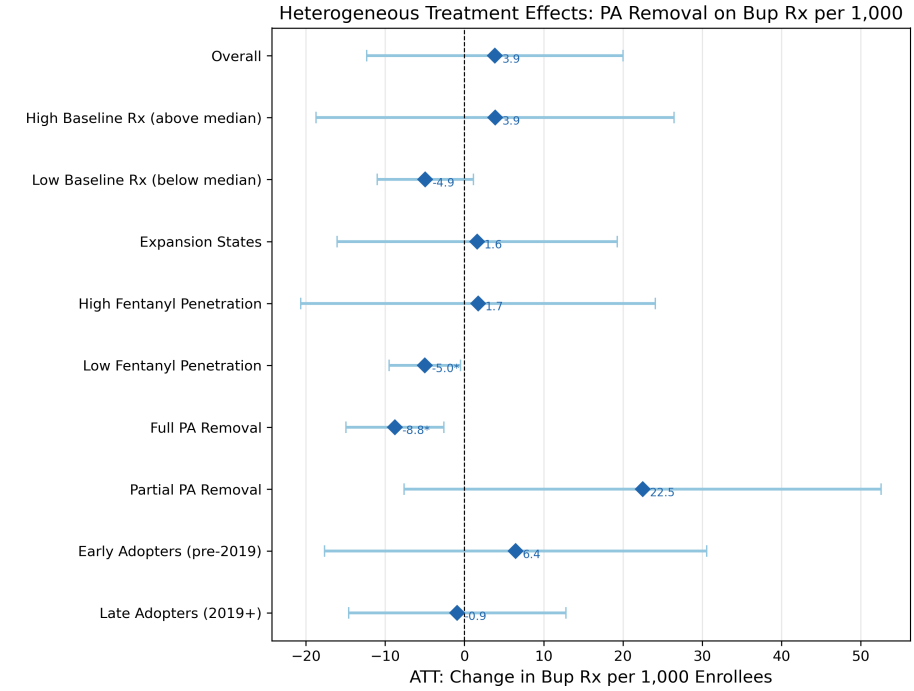
Appendix Figure A1. Saturated TWFE event study estimates of the effect of PA removal on buprenorphine prescriptions. Point estimates and 95 percent confidence intervals are plotted by quarter relative to PA removal. Pre-treatment coefficients are centered at zero at t = -1.



.

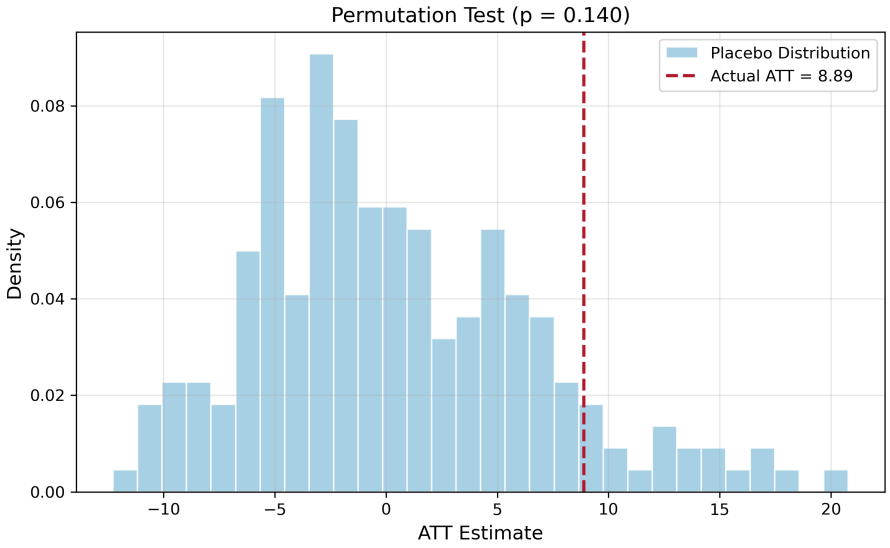
Appendix Figure A2. Forest plot of heterogeneity estimates from Appendix Exhibit A3. Point estimates and 95 percent confidence intervals for each subgroup, with the overall ATT indicated by a vertical reference line.

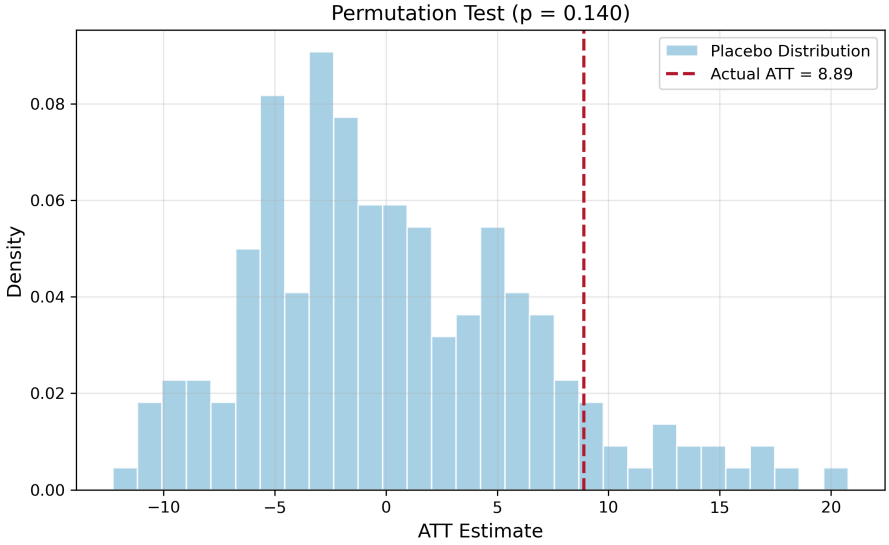

Heterogeneous Treatment Effects: PA Removal on Bup Rx per 1,000
Overall
High Baseline Rx (above median)
Low Baseline Rx (below median)
Expansion States
High Fentanyl Penetration
Low Fentanyl Penetration
Full PA Removal
Partial PA Removal
Early Adopters (pre-2019)
Late Adopters (2019+)
3.9
3.9
-4.9
1.6
1.7
-5.0*
-8.8*
22.5
6.4
-0.9
−20
−10
0
10
20
30
40
50
ATT: Change in Bup Rx per 1,000 Enrollees


Heterogeneous Treatment Effects: PA Removal on Bup Rx per 1,000
Overall
High Baseline Rx (above median)
Low Baseline Rx (below median)
Expansion States
High Fentanyl Penetration
Low Fentanyl Penetration
Full PA Removal
Partial PA Removal
Early Adopters (pre-2019)
Late Adopters (2019+)
3.9
3.9
-4.9
1.6
1.7
-5.0*
-8.8*
22.5
6.4
-0.9
−20
−10
0
10
20
30
40
50
ATT: Change in Bup Rx per 1,000 Enrollees

.

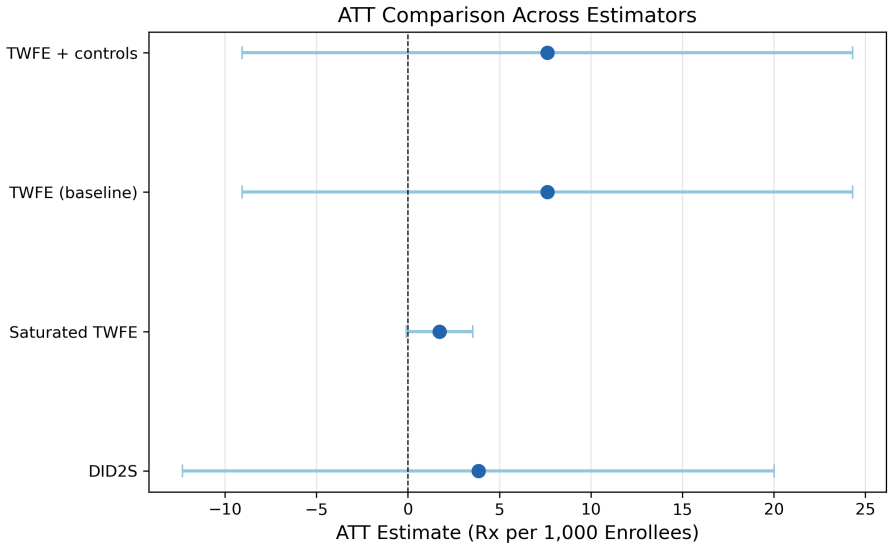
Appendix Figure A3. Distribution of placebo ATTs from 200 random permutations of treatment assignment. The observed ATT of 8.89 is indicated by a vertical line.



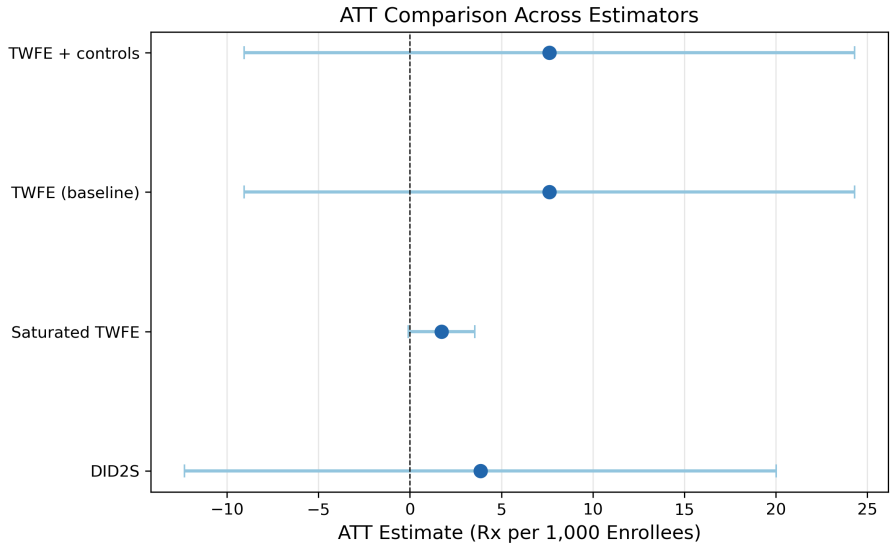


.

Appendix Figure A4. Comparison of ATT estimates across estimators (DID2S, conventional TWFE, saturated TWFE, LP-DiD) for the buprenorphine

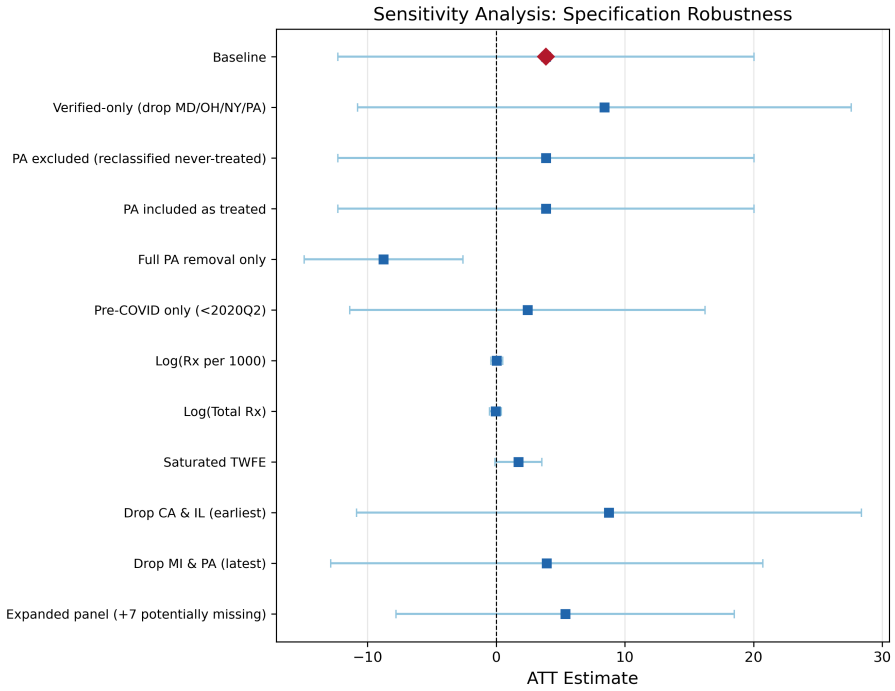


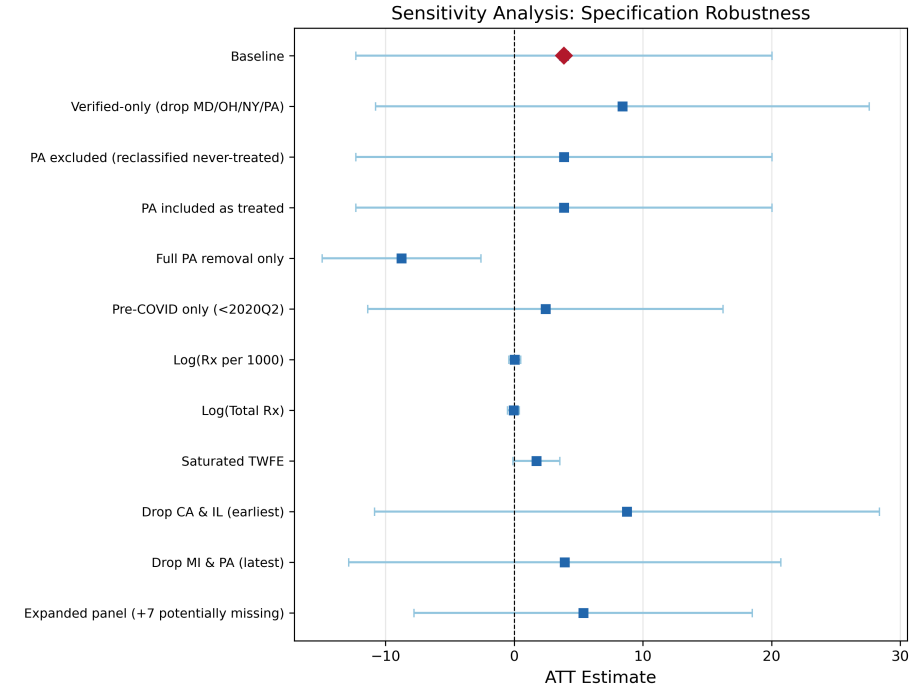
prescribing outcome.



.

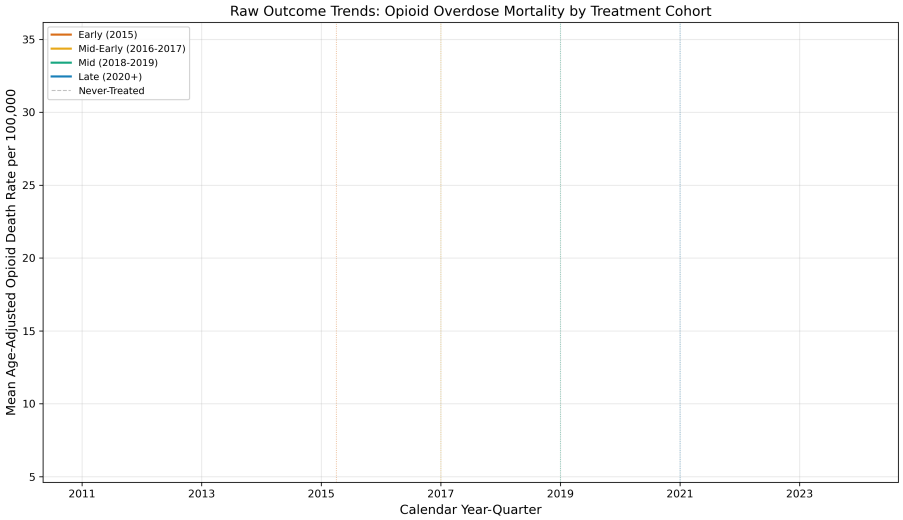
Appendix Figure A5. Sensitivity of the baseline ATT to alternative specifications, samples, and estimators from Appendix Exhibit A2.

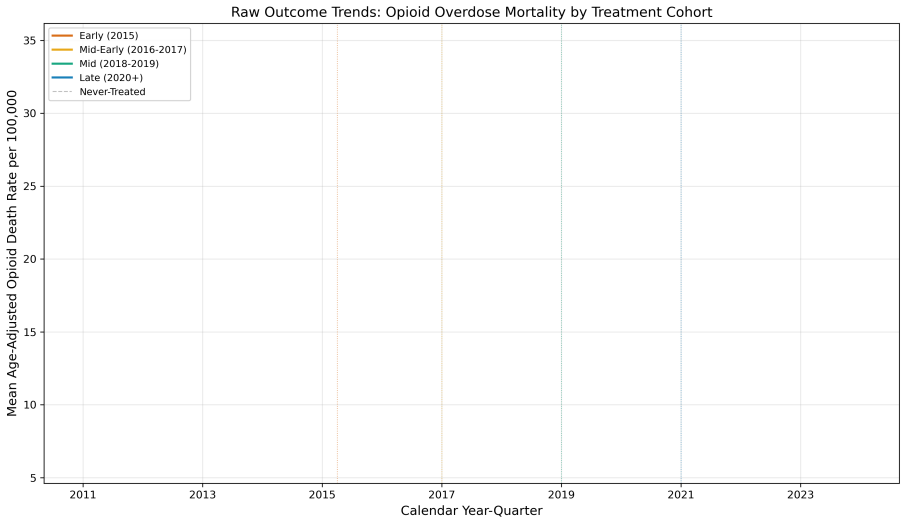

Sensitivity Analysis: Specification Robustness
Baseline
Verified-only (drop MD/OH/NY/PA)
PA excluded (reclassified never-treated)
PA included as treated
Full PA removal only
Pre-COVID only (<2020Q2)
Log(Rx per 1000)
Log(Total Rx)
Saturated TWFE
Drop CA & IL (earliest)
Drop MI & PA (latest)
Expanded panel (+7 potentially missing)
−10
0
10
20
30
ATT Estimate


Sensitivity Analysis: Specification Robustness
Baseline
Verified-only (drop MD/OH/NY/PA)
PA excluded (reclassified never-treated)
PA included as treated
Full PA removal only
Pre-COVID only (<2020Q2)
Log(Rx per 1000)
Log(Total Rx)
Saturated TWFE
Drop CA & IL (earliest)
Drop MI & PA (latest)
Expanded panel (+7 potentially missing)
−10
0
10
20
30
ATT Estimate

.

Appendix Figure A6. Raw cohort-specific opioid overdose mortality trends over 2011–2020 (limited by CDC WONDER data availability). All cohorts trend upward, consistent with the national fentanyl crisis. The Mid (2018–2019) cohort has the highest baseline mortality. No obvious visual discontinuity at treatment timing, consistent with the mixed/null mortality findings in the main analysis.

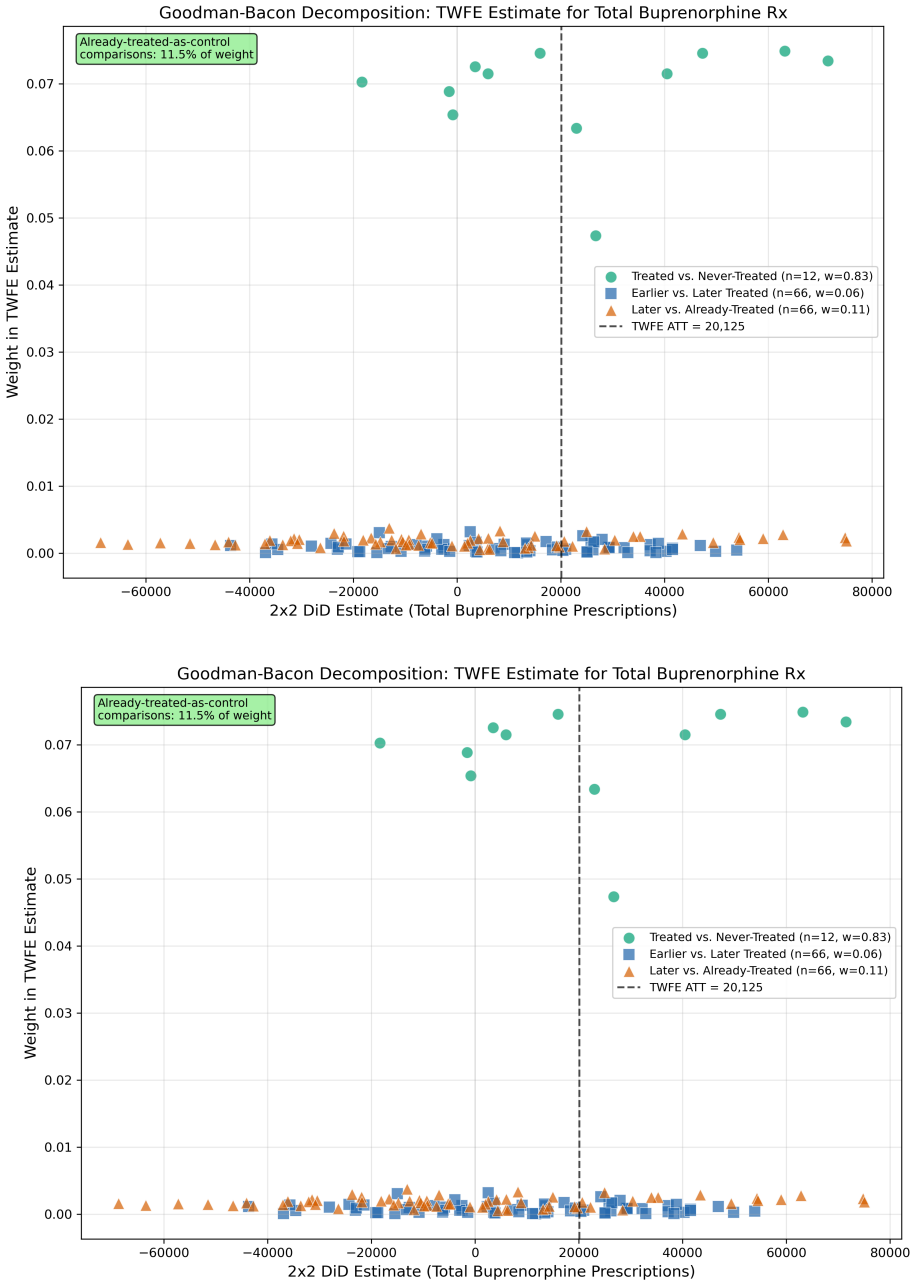




.

Appendix Figure A7. Goodman-Bacon decomposition for the standard TWFE specification of total buprenorphine prescriptions. Each point represents a 2x2 DiD estimate (x-axis) plotted against its TWFE weight (y-axis), colored by comparison type. Treated-vs-never-treated comparisons carry 82.8 percent of TWFE weight (clean comparisons); later-vs-already-treated

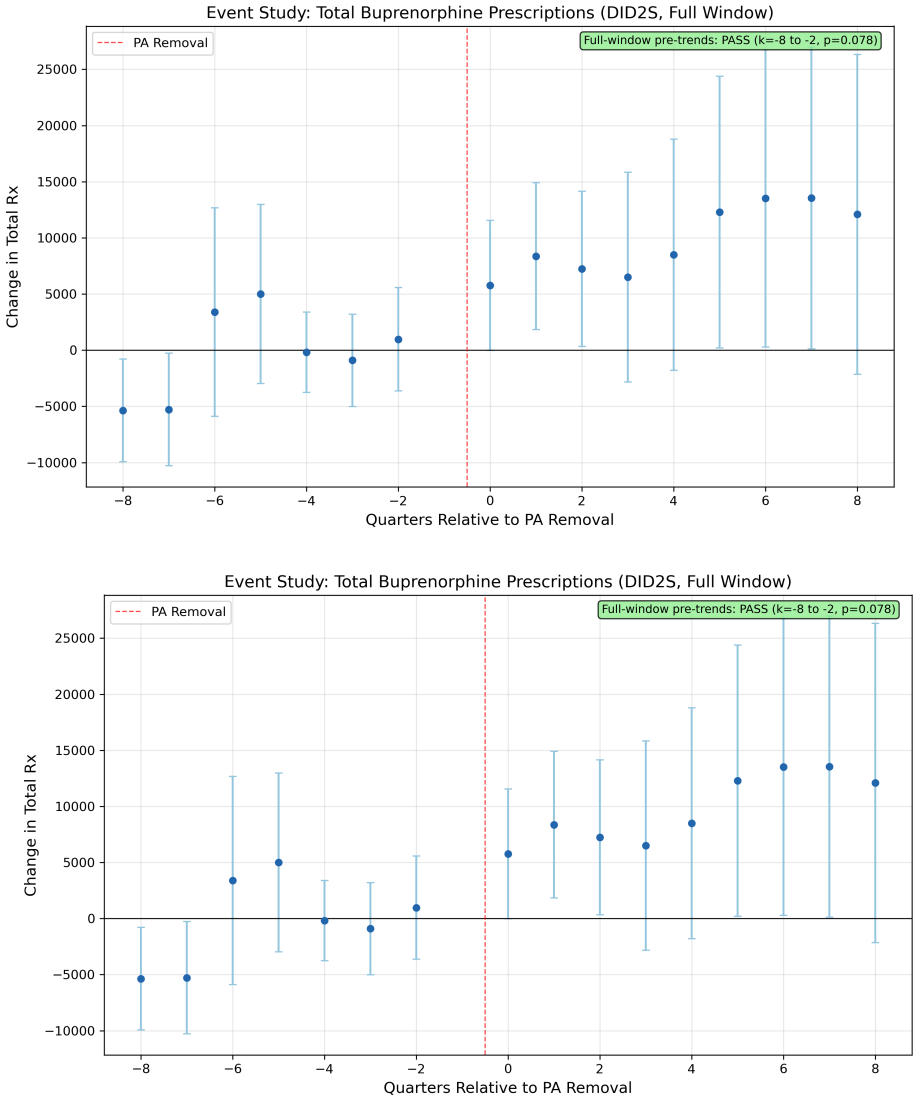
comparisons carry 11.5 percent (potentially problematic); earlier-vs-later comparisons carry 5.7 percent. The low weight on problematic comparisons is reassuring that TWFE contamination is minimal in this setting.



.

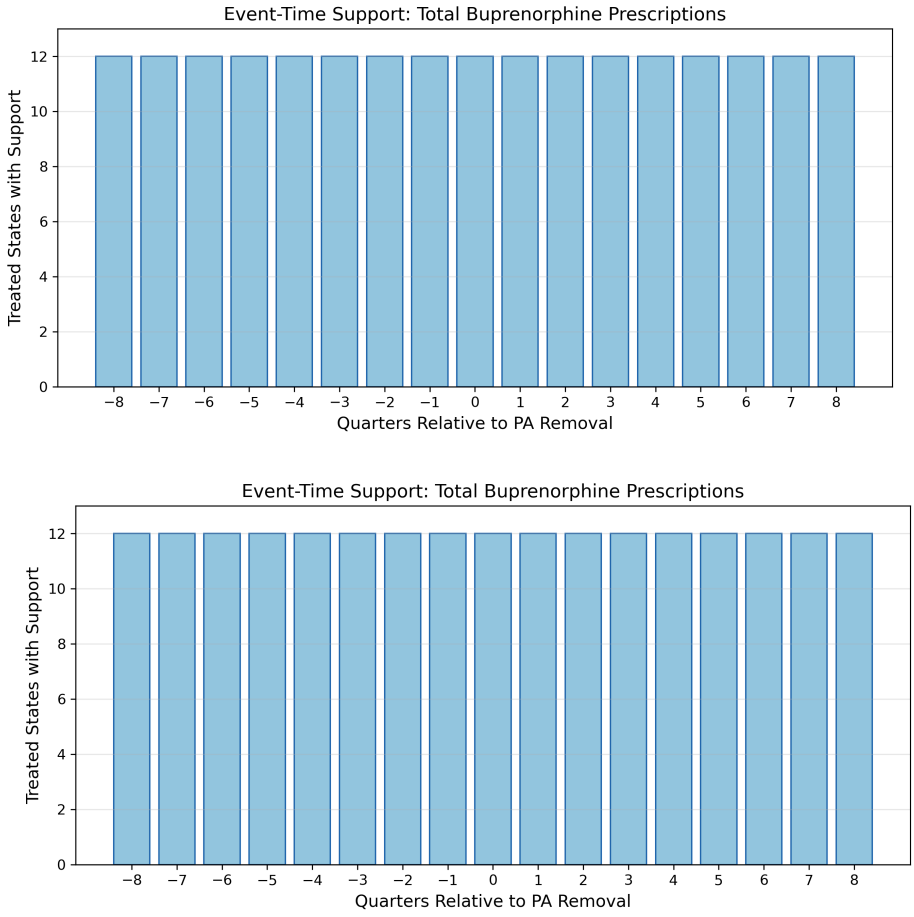
Appendix Figure A8. Full DID2S event-study path for total buprenorphine

prescriptions over the complete $k = -8$ through $k = +8$ window. The farthest leads ($k = -7$ and $k = -8$) are negative and drive the full-window clustered pre-trends rejection, while the near pre-period is comparatively flat.



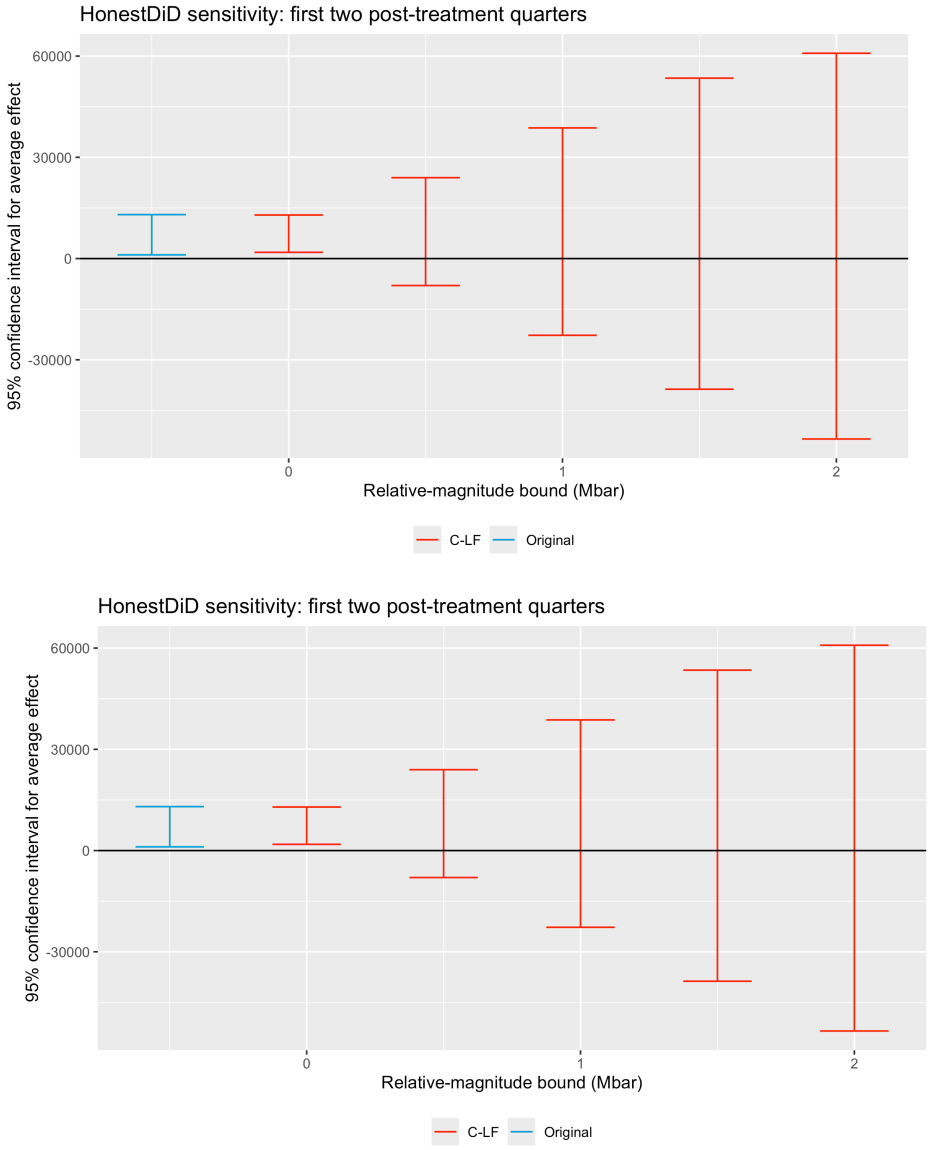
.

Appendix Figure A9. Event-time support counts for the total-prescribing event study. The figure reports the number of treated-state observations and unique treated states contributing to each lead and lag, showing that support thins in the distant pre-period and at the longest post-treatment horizons.



.

Appendix Figure A10. HonestDiD sensitivity analysis for the total-prescribing outcome, targeting the average effect over the first two post-treatment quarters. The original confidence interval is positive, but the robust interval crosses zero by $\bar{M} = 0.5$, indicating that the short-run effect is not robust to moderate departures from parallel trends.



.

References

Andraka-Christou, B., Golan, O., Totaram, R., Ohama, M., Saloner, B., Gordon, A. J., & Stein, B. D. (2023). Prior authorization restrictions on medications for opioid use disorder: Trends in state laws from 2005 to 2019. *Annals of Medicine*, *55*(1), 514–520. https://doi.org/10.1080/07853890.2023.2171107

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event study designs: Robust and efficient estimation. *Review of Economic Studies*, *91*(6), 3253–3285.

Christine, P. J., Larochelle, M. R., Lin, L. A., McBride, J., & Tipirneni, R. (2023). Removal of Medicaid prior authorization requirements and buprenorphine treatment for opioid use disorder. *JAMA Health Forum*, *4*(10), e233549. https://doi.org/10.1001/jamahealthforum.2023.3549

Gardner, J. (2022). Two-stage differences in differences. Working paper.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

Tiako, M. J. N., Dolan, A., Abrams, M., Oyekanmi, K., Meisel, Z., & Aronowitz, S. V. (2023). Thematic analysis of state Medicaid buprenorphine prior authorization requirements. *JAMA Network Open*, *6*(6), e2318487. https://doi.org/10.1001/jamanetworkopen.2023.18487