

Must-Access Prescription Drug Monitoring Program Mandates and State Employment: A Staggered Difference-in-Differences Analysis

APEP Autonomous Research*

@ai1scl

January 29, 2026

Abstract

This paper estimates the causal effect of state-level must-access Prescription Drug Monitoring Program (PDMP) mandates on aggregate employment outcomes using a staggered difference-in-differences design. Between 2013 and 2021 (full-exposure years), 46 U.S. states adopted laws requiring prescribers to query their state's PDMP before issuing controlled substance prescriptions; all 46 of which are included in the estimation sample. I exploit this staggered rollout using the [Callaway and Sant'Anna \(2021\)](#) estimator with not-yet-treated states as the primary comparison group and Bureau of Labor Statistics Local Area Unemployment Statistics data spanning 50 states from 2007–2023. The results yield a precisely estimated null at the aggregate level: the estimated average treatment effect on the treated (ATT) for log employment is $+0.0036$ ($SE = 0.0079$, $p = 0.647$; 95% CI $[-0.012, 0.019]$), and the effect on the unemployment rate is -0.242 percentage points ($SE = 0.293$, $p = 0.407$). The not-yet-treated comparison group is preferred because only four states (Kansas, Missouri, Nebraska, South Dakota) never adopted universal must-access mandates, and this thin control group produces spurious pre-trend violations in event-study analysis. Using the never-treated control group as a sensitivity check, the ATT for log employment is $+0.0100$ ($SE = 0.0078$, $p = 0.203$)—also statistically insignificant. Group-level estimates reveal heterogeneity across adoption cohorts, but the overall pattern across all specifications is consistent with negligible aggregate employment effects. These findings suggest that must-access PDMP mandates—while effective at reducing opioid prescribing—do not produce detectable

*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch

changes in state-level employment aggregates over the medium run, though the design cannot rule out meaningful subpopulation effects diluted below the minimum detectable effect threshold of approximately 2.2 percent.

JEL Codes: I18, J21, K32

Keywords: prescription drug monitoring, PDMP, opioid policy, employment, difference-in-differences

1. Introduction

Can supply-side drug policies reverse the labor market damage caused by the opioid epidemic? Between 1999 and 2021, more than 600,000 Americans died from opioid-involved overdoses ([Centers for Disease Control and Prevention, 2023](#)), and the crisis inflicted deep economic damage through reduced labor force participation, rising disability claims, and lost productivity ([Krueger, 2017](#); [Case and Deaton, 2020](#)). States responded with intense policy experimentation, and “must-access” Prescription Drug Monitoring Program (PDMP) mandates—laws requiring prescribers to query an electronic database before writing controlled substance prescriptions—became the most widely adopted intervention. By 2021, 46 of 50 U.S. states had enacted such mandates.

The prescribing effects of these mandates are well established: [Buchmueller and Carey \(2018\)](#) find meaningful reductions in opioid prescribing among Medicare beneficiaries, [Dave et al. \(2021\)](#) document declines in drug abuse outcomes, and [Brady et al. \(2014\)](#) show sustained reductions in prescribing rates. But whether reduced prescribing translates into improved labor market outcomes is a distinct and harder question. The causal chain runs through several intermediate steps—mandates reduce prescribing, reduced prescribing limits new opioid dependencies, fewer dependencies mean fewer impaired workers—and each link introduces attenuation. Moreover, the rise of illicit fentanyl and heroin means that interventions targeting the prescription channel may fail to address the dominant sources of opioid harm ([Alpert et al., 2018](#)). If supply-side restrictions simply shift users to illicit markets, labor market benefits may never materialize.

This paper estimates the causal effect of state must-access PDMP mandates on aggregate employment outcomes. I exploit the staggered adoption of these mandates across 46 U.S. states (effective dates 2012–2020, full-exposure years 2013–2021), using the [Callaway and Sant'Anna \(2021\)](#) doubly-robust difference-in-differences estimator. The analysis uses Bureau of Labor Statistics Local Area Unemployment Statistics (LAUS) data, which provide consistent state-level measures of employment and unemployment rates. The estimation sample consists of 50 states observed from 2007 through 2023, yielding 850 state-year observations (46 treated states plus 4 never-treated). Because only four states—Kansas, Missouri, Nebraska, and South Dakota—never adopted universal must-access mandates during the sample period, I use not-yet-treated states as the primary comparison group, which provides a larger and more geographically diverse control set at each point in time.

The main finding of this paper is a precisely estimated null at the aggregate level. Using the preferred not-yet-treated specification, the overall average treatment effect on the treated (ATT) for log employment is +0.0036 with a standard error of 0.0079 ($p = 0.647$). The 95

percent confidence interval of approximately $[-0.012, 0.019]$ allows me to rule out employment effects larger than about 1.9 percent in magnitude. For the unemployment rate, the not-yet-treated ATT is -0.242 percentage points ($SE = 0.293, p = 0.407$), which is economically and statistically indistinguishable from zero. An important caveat applies: the minimum detectable effect at 80 percent power is approximately 2.2 percent for log employment, which is large relative to the plausibly affected population. If mandates affect only the 2–3 percent of workers with opioid prescriptions, even a substantial individual-level effect would generate an aggregate impact far below this threshold. The null is therefore informative about large aggregate effects but cannot rule out economically meaningful effects on targeted subpopulations. As a sensitivity check, using the four never-treated states as controls yields a log employment ATT of $+0.0100$ ($SE = 0.0078, p = 0.203$) and an unemployment rate effect of -0.399 percentage points ($SE = 0.369, p = 0.280$)—also statistically insignificant, though with pre-trend concerns due to the thin control group. Traditional two-way fixed effects (TWFE) estimates are quantitatively similar: the TWFE coefficient on log employment is $+0.0033$ ($SE = 0.0055$), robust to controlling for concurrent policies including Medicaid expansion and medical marijuana legalization ($+0.0032, SE = 0.0052$).

The event-study analysis provides further support for both the null result and the validity of the research design. For the primary log employment outcome, using the preferred not-yet-treated comparison group, all pre-treatment estimates at horizons $e = -6$ through $e = -2$ are insignificant under 95% uniform (simultaneous) confidence bands, consistent with the parallel trends assumption. Post-treatment estimates are also insignificant under uniform bands through ten years of follow-up. For the unemployment rate, a pointwise-significant pre-trend at $e = -1$ is detected under the not-yet-treated specification (though not significant under uniform bands), possibly reflecting partial mandate exposure during the transition year; all other pre- and post-treatment unemployment rate estimates are insignificant. The never-treated comparison group produces a significant pre-trend at $e = -2$ for log employment, which I attribute to the thin and geographically concentrated nature of the four-state control group rather than to genuine anticipatory effects.

Group-level estimates by adoption cohort reveal heterogeneity. The earliest adopters (2013 cohort: West Virginia, New Mexico, Kentucky) show a statistically significant negative effect on log employment of -0.0247 ($SE = 0.0102, p < 0.05$). The latest adopters (2020 and 2021 cohorts) show significant positive effects, though these are based on very few states and short post-treatment windows. The overall ATT aggregating across all cohorts is insignificant under both the never-treated and not-yet-treated specifications.

This paper makes three contributions to the literature. First, it provides a comprehensive staggered difference-in-differences analysis of must-access PDMP mandates and employment

covering all 46 treated states with adoption dates spanning nine years. While [Kaestner and Ziedan \(2023\)](#) examine the labor market effects of opioid access restrictions more broadly, this paper focuses specifically on the must-access mandate mechanism, covers a longer post-treatment window (up to eleven years for the earliest adopters; event-study plots display dynamic effects through six years), and employs the recently developed [Callaway and Sant'Anna \(2021\)](#) heterogeneity-robust estimator, which addresses the bias concerns raised by [Goodman-Bacon \(2021\)](#), [de Chaisemartin and D'Haultfœuille \(2020\)](#), and [Sun and Abraham \(2021\)](#) for staggered adoption settings.

Second, the paper contributes a carefully bounded null result to a literature that has sometimes overemphasized the employment consequences of opioid-related policies. [Harris et al. \(2020\)](#) document substantial employment effects of opioid exposure, and [Aliprantis et al. \(2020\)](#) link opioid prescribing to labor force participation, but these findings concern the effect of opioids on labor markets—not the reverse channel of whether policies that reduce opioid prescribing improve labor market outcomes. The distinction matters because the relationship need not be symmetric: the damage caused by opioid dependence may not be easily reversed by reducing the flow of new prescriptions, and must-access mandates do not treat existing addictions. My null finding is consistent with this asymmetry and with the growing recognition that supply-side interventions alone may be insufficient to address the labor market consequences of the opioid crisis ([Powell, 2020](#); [Maclean et al., 2022](#)).

Third, the paper highlights the methodological challenge of conducting staggered DiD analysis when nearly all units eventually adopt the treatment. With 46 of 50 states adopting must-access mandates by 2021, the never-treated control group is necessarily thin and potentially unrepresentative. I demonstrate that this thin-control-group problem can generate spurious pre-trend violations and show that the not-yet-treated comparison group provides a more credible alternative in this setting. This methodological finding is informative for the growing number of applied settings where policy adoption is near-universal.

The remainder of the paper proceeds as follows. [Section 2](#) reviews the institutional background of PDMPs and must-access mandates, as well as the prior literature on their effects. [Section 3](#) describes the data sources and sample construction. [Section 4](#) presents the empirical strategy, including the Callaway-Sant'Anna estimator and identification assumptions. [Section 5](#) reports the main results, event-study estimates, and group-level heterogeneity. [Section 6](#) presents robustness checks including TWFE comparisons, alternative control groups, and the thin-control-group problem. [Section 7](#) interprets the null finding and discusses limitations. [Section 8](#) concludes with policy implications.

2. Institutional Background and Prior Literature

2.1 The Opioid Crisis and Labor Markets

The U.S. opioid epidemic evolved in three overlapping waves. The first wave, beginning in the late 1990s, was driven by a dramatic increase in opioid prescribing for chronic pain management. Pharmaceutical companies aggressively marketed opioid analgesics—most notably OxyContin—as safe and effective for long-term pain management, contributing to a four-fold increase in opioid prescriptions between 1999 and 2010 ([Kolodny et al., 2015](#); [Macy, 2018](#)). The second wave emerged around 2010 as users transitioned from prescription opioids to heroin, often because prescription sources became more costly or more difficult to access. The third wave, beginning around 2013, was characterized by the proliferation of synthetic opioids, particularly illicitly manufactured fentanyl, which now accounts for the majority of opioid-involved overdose deaths.

The labor market consequences of the opioid epidemic have been a subject of intense scholarly and policy interest. [Krueger \(2017\)](#) argued that nearly half of prime-age men not in the labor force take pain medication daily, and that the increase in opioid prescribing can account for a meaningful share of the decline in male labor force participation since 2000. [Case and Deaton \(2015\)](#) and [Case and Deaton \(2020\)](#) documented the broader “deaths of despair” phenomenon, linking rising midlife mortality among white non-Hispanic Americans to opioids, alcohol, and suicide, and arguing that these trends reflect deeper economic and social dysfunction. [Harris et al. \(2020\)](#) provided quasi-experimental evidence that opioid exposure reduces labor supply, exploiting variation in local access to OxyContin following the introduction of an abuse-deterrent reformulation. [Borgschulte et al. \(2022\)](#) study the effect of opioids on labor markets through drug reclassification and find significant employment effects.

However, the causal direction between opioid use and labor market outcomes is contested. [Hollingsworth et al. \(2017\)](#) find that worsening macroeconomic conditions contribute to opioid abuse, suggesting a demand-side mechanism in which economic distress drives substance use rather than vice versa. [Currie and Schwandt \(2019\)](#) argue that supply-side factors—pharmaceutical marketing, prescriber behavior, and the availability of illicit drugs—were more important than economic conditions in driving the epidemic. This debate has important implications for the effectiveness of supply-side policies such as PDMP mandates: if the epidemic is primarily supply-driven, then interventions that constrain prescribing should reduce opioid-related harms including labor market impairment; if the epidemic is demand-driven, then restricting one source of opioids may simply shift users to alternative substances or sources without improving labor market outcomes.

2.2 Prescription Drug Monitoring Programs

Prescription Drug Monitoring Programs are state-administered electronic databases that collect, monitor, and analyze electronically transmitted prescribing and dispensing data for controlled substances. The core function of a PDMP is to provide prescribers and dispensers with patient-level information about controlled substance prescription histories, enabling them to identify potential misuse, diversion, and doctor shopping. PDMPs have existed in some form since the 1930s, but the modern era of electronic PDMPs began in the early 2000s ([Horwitz et al., 2021](#)). By the mid-2010s, nearly all states had operational PDMP databases, though the timeline varied. Missouri was a notable holdout: the state did not establish a statewide PDMP until December 2023, having long relied on a patchwork of county-level programs. The late establishment of Missouri's program—along with the absence of universal must-access mandates in Kansas, Nebraska, and South Dakota—provides the never-treated control group for this analysis.

However, the mere existence of a PDMP database is distinct from requirements governing its use. In the absence of a use mandate, prescribers can—but are not required to—consult the database before writing prescriptions. Utilization rates for voluntary PDMPs have historically been low: studies consistently found that fewer than half of eligible prescribers routinely checked the database when use was voluntary ([Brady et al., 2014](#); [Nam et al., 2017](#)). This limited uptake of voluntary PDMPs motivated a policy shift toward mandating use.

“Must-access” or “must-query” mandates require prescribers to check the PDMP database before prescribing or dispensing controlled substances, typically with specific requirements regarding timing (e.g., checking within 24 hours before prescribing), covered substances (e.g., Schedule II drugs, or all scheduled drugs), and covered encounters (e.g., initial prescriptions, all prescriptions, or prescriptions exceeding a certain duration). The specifics vary across states, but the common feature is a legal requirement that transforms the PDMP from a voluntary information tool into a mandatory check.

Kentucky enacted one of the first comprehensive must-access mandates in July 2012, followed closely by West Virginia and New Mexico. Adoption then proceeded rapidly through 2019. Several factors influenced adoption timing: the severity of the local opioid crisis (states with higher prescribing rates and overdose deaths adopted earlier), political conditions and legislative priorities, the readiness of existing PDMP technology infrastructure, and federal incentive programs such as the Harold Rogers PDMP grants and PDMP Enhancement and Training (PMET) funding. By 2021, 46 of the 50 U.S. states had adopted universal must-access mandates, with only Kansas, Missouri, Nebraska, and South Dakota remaining without such requirements. [Table 2](#) in the Data section presents the full timeline of adoption dates used in this analysis.

2.3 Prior Evidence on PDMP Effects

The literature on PDMP effects can be organized along two dimensions: the outcome studied and the type of PDMP provision examined.

Effects on prescribing. The most robust evidence concerns the impact of must-access mandates on opioid prescribing. [Buchmueller and Carey \(2018\)](#) use a difference-in-differences design and find that must-access mandates reduce Schedule II opioid prescriptions by 8 percent among Medicare beneficiaries and shift prescribing toward lower-dosage formulations. [Brady et al. \(2014\)](#) document sustained reductions in opioid prescribing following mandate adoption. [Bao et al. \(2016\)](#) find similar effects using commercially insured populations. [Dave et al. \(2021\)](#) show that mandatory-access provisions reduce opioid misuse and related hospitalizations. The consensus is that must-access mandates do meaningfully reduce opioid prescribing, at least through the prescription channel.

Effects on health outcomes. [Patrick et al. \(2016\)](#) find that states with robust PDMPs experienced reductions in opioid-related overdose deaths, though the magnitudes vary across specifications and the evidence is more mixed than for prescribing outcomes. [Meara et al. \(2016\)](#) show that stronger PDMP provisions reduce Schedule II prescribing among disabled adults. [Pardo \(2017\)](#) examines whether more robust PDMP programs reduce prescription opioid overdose, finding mixed evidence that depends on specific program features. [Fink et al. \(2018\)](#) provide a systematic review of associations between PDMPs and drug overdoses, highlighting heterogeneous effects across program types and the importance of mandate strength. However, some evidence suggests substitution to heroin and illicit fentanyl, which could partially offset or even reverse health gains from reduced prescribing ([Meinholfer, 2018](#); [Powell et al., 2020](#); [Alpert et al., 2018](#)).

Effects on labor market outcomes. The evidence on labor market effects is thin and mixed. [Kaestner and Ziedan \(2023\)](#) examine a broad set of opioid access restrictions—including PDMP mandates, prescription limits, and pill mill laws—and find limited evidence of employment effects at the state level, consistent with the null finding in this paper. [Kaestner and Ziedan \(2019\)](#) study the mortality and socioeconomic consequences of prescription opioid policies and find that PDMP mandates reduce prescribing but do not examine downstream labor market effects in depth. [Borgschulte et al. \(2022\)](#) study the labor market effects of opioid access through a different channel (drug reclassification) and find significant employment effects, but their variation is distinct from PDMP mandates. [Aliprantis et al. \(2020\)](#) document correlations between opioid prescribing and labor market outcomes but do not isolate the PDMP mandate channel. To my knowledge, no prior study has applied the [Callaway and Sant'Anna \(2021\)](#) estimator to the specific question of must-access PDMP mandates and state-level employment, making this analysis a novel contribution.

3. Data

3.1 Employment Data: BLS Local Area Unemployment Statistics

The primary outcome data come from the Bureau of Labor Statistics (BLS) Local Area Unemployment Statistics (LAUS) program ([Bureau of Labor Statistics, 2023](#)). LAUS provides monthly estimates of total employment, total unemployment, the civilian labor force, and the unemployment rate for all states, the District of Columbia, and sub-state geographic areas. These estimates are produced using a model-based methodology that incorporates data from the Current Population Survey (CPS), the Current Employment Statistics (CES) survey, and state unemployment insurance (UI) administrative records.

I use the March monthly LAUS estimates for each state from 2007 through 2023, matching the timing of the CPS Annual Social and Economic Supplement (ASEC). March values provide a consistent within-year snapshot of labor market conditions while avoiding the partial-treatment contamination that would affect annual averages for states with mid-year mandate effective dates. The District of Columbia is excluded because it is a federal district with distinct economic characteristics. The resulting balanced panel consists of 50 states observed over 17 years, yielding 850 state-year observations. The key outcome variables are:

1. **Log employment:** The natural logarithm of total employment in the state. This is the primary outcome because it captures changes in employment levels in proportional terms and is less sensitive to differences in state population size than the level of employment.
2. **Unemployment rate:** The number of unemployed individuals as a share of the civilian labor force. This is the most widely recognized measure of labor market slack and captures the intensive margin of labor demand relative to supply.
3. **Employment share of labor force:** Defined as $100 - \text{unemployment rate}$, this measure captures the fraction of the civilian labor force that is employed. Because it is mechanically the complement of the unemployment rate, estimation results for this outcome are identical in magnitude and opposite in sign to the unemployment rate results, and are therefore not reported separately. This variable appears in the summary statistics ([Table 1](#)) for descriptive purposes only. This is *not* the employment-to-population ratio (E/POP), which would require a separate population denominator not provided by the LAUS program at consistent annual frequency.

The LAUS data have several advantages for this analysis. First, they provide a consistent and comprehensive measure of state-level employment that spans the entire pre- and post-

treatment period for all adoption cohorts. Second, they are available at the annual frequency that matches the natural granularity of state policy adoption. Third, they are not subject to the survey sampling variation that affects CPS microdata at the state level, because LAUS estimates incorporate administrative data and model-based smoothing. The primary limitation is that LAUS data are aggregate state-level measures and cannot be disaggregated by age, education, occupation, or other demographics that might help identify the specific populations affected by PDMP mandates.

3.2 Treatment Data: Must-Access PDMP Mandate Dates

The treatment variable is a binary indicator equal to one in state-year cells where a must-access PDMP mandate was in effect for the full calendar year. I define the “full-exposure year” as the first calendar year in which the mandate was in effect by January 1. Under this convention, the March outcome at event time $e = 0$ captures a state where the mandate has been in effect since at least January 1—a minimum of two months of exposure at the time of the March measurement. For states whose mandates took effect before January of the full-exposure year—for example, Rhode Island (effective March 1, 2016, full-exposure year 2017)—the mandate was operative for over a year by the March $e = 0$ observation. In the year immediately before the full-exposure year ($e = -1$), some states with early effective dates (before March) will have partial mandate exposure in their March observation. The CS estimator handles this through its **anticipation window of one year** ($\text{anticipation} = 1$), which excludes $e = -1$ from both the set of pre-treatment periods used to estimate counterfactuals and the pre-trends test, and removes these partially-treated observations from the control group for other cohorts. The pre-treatment test therefore focuses on $e \leq -2$, where no state has mandate exposure in its March observation regardless of within-year effective date.

The must-access mandate effective dates were compiled from multiple sources. The primary source is the Prescription Drug Abuse Policy System (PDAPS), maintained by the Temple University Beasley School of Law, which provides a comprehensive database of state-level controlled substance laws. I cross-referenced PDAPS dates with those reported in [Horwitz et al. \(2021\)](#), [Buchmueller and Carey \(2018\)](#), and legislative databases maintained by the National Conference of State Legislatures (NCSL). Where sources disagreed, I used the PDAPS date as the primary reference and noted discrepancies.

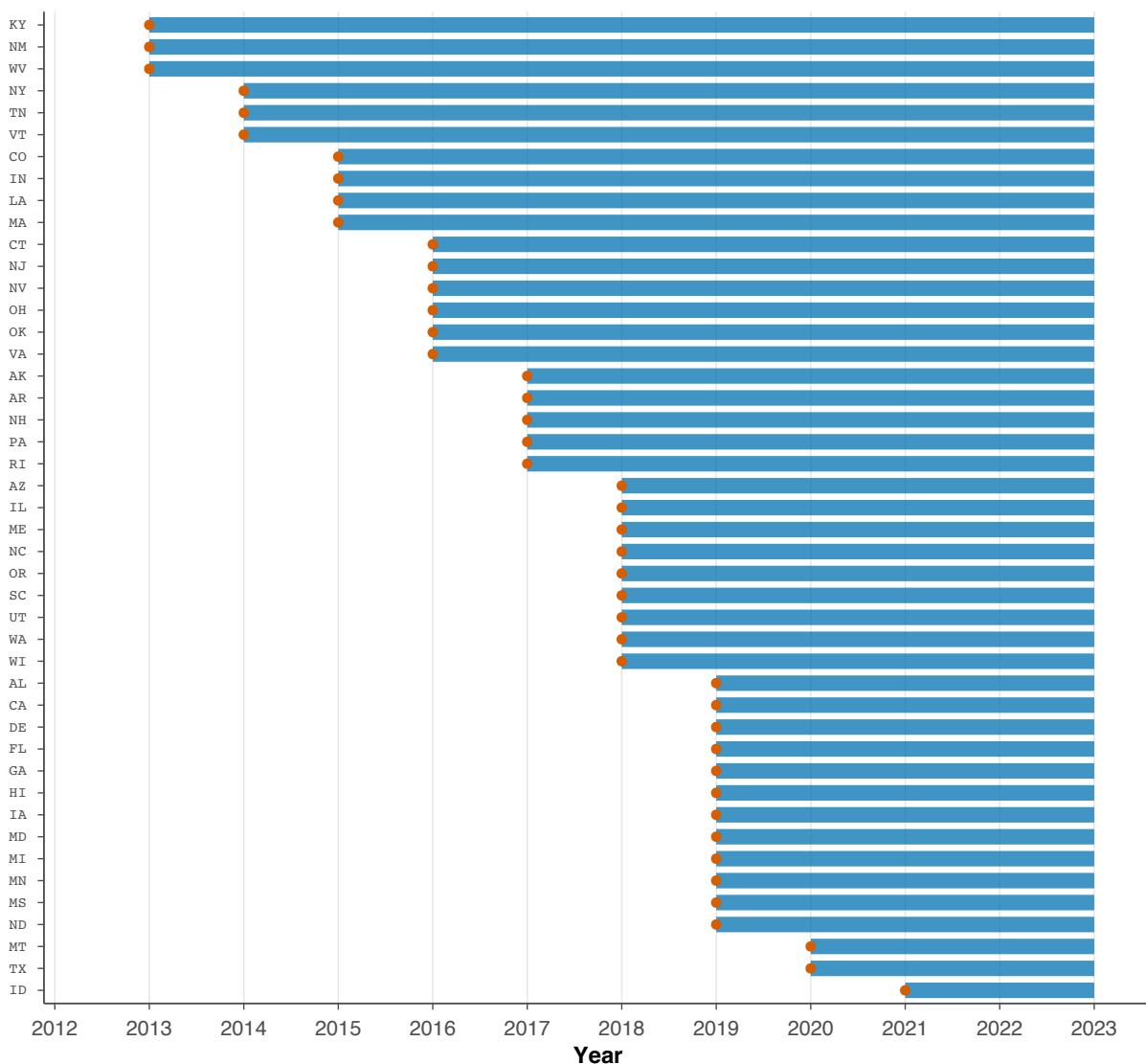
The resulting classification identifies 46 treated states nationally, with full-exposure years ranging from 2013 (West Virginia, New Mexico, Kentucky) to 2021 (Idaho, Wyoming). Four states had not adopted universal must-access mandates by the end of the sample period and serve as the never-treated control group: Kansas, Missouri, Nebraska, and South

Dakota. These four states either lack a must-access mandate entirely or have only limited mandates applying to specific populations (e.g., Medicaid prescribers only) rather than universal prescriber requirements. [Table 2](#) presents the complete list of adoption dates and full-exposure years.

The staggered adoption pattern is illustrated in [Figure 1](#). Three states adopted in the 2013 cohort, three in 2014, four in 2015, six in 2016, five in 2017, nine in 2018, twelve in 2019, two in 2020, and two (Idaho, Wyoming) in 2021. This adoption pattern means that early adopters have the longest post-treatment windows (up to eleven years for the 2013 cohort) while late adopters have shorter windows (three years for the 2021 cohort).

Staggered Adoption of PDMP Must-Access Mandates

Year of full mandate exposure through 2023



Note: Dot marks the first full year of mandate exposure. Bar extends to 2023.

Figure 1: Staggered Adoption of Must-Access PDMP Mandates (Effective Dates 2012–2020; Full-Exposure Years 2013–2021)

Notes: Each point represents the effective date of the must-access PDMP mandate for a state. States are ordered by adoption date. The “full-exposure year” used in the analysis is the first calendar year in which the mandate was in effect by January 1. Never-treated states (KS, MO, NE, SD) are not shown. The figure depicts all 46 treated states included in the analysis.

Source: PDAPS, [Horwitz et al. \(2021\)](#), [Buchmueller and Carey \(2018\)](#), NCSL.

3.3 Concurrent Policy Variables

For robustness specifications that control for concurrent state-level policies, I compiled effective dates for four additional policy interventions that were adopted during the same period and could plausibly affect employment outcomes. Naloxone access laws expand availability of the opioid overdose reversal agent through standing orders, pharmacy protocols, or third-party prescriptions (dates from PDAPS). Good Samaritan laws provide legal protections for individuals who call for emergency assistance during an overdose (dates from PDAPS). Medicaid expansion under the Affordable Care Act extended health insurance coverage to adults with incomes up to 138 percent of the federal poverty level (dates from the Kaiser Family Foundation). Medical marijuana legalization permits the use of marijuana for medical purposes (dates from NCSL). Each concurrent policy is coded as a binary indicator equal to one in state-year cells where the policy was in effect for the full calendar year, using the same full-exposure-year convention applied to must-access mandates.

3.4 Summary Statistics

[Table 1](#) presents summary statistics for the estimation sample, separately for ever-treated and never-treated states. The sample comprises 850 state-year observations: 782 from the 46 ever-treated states and 68 from the 4 never-treated states. Mean employment is considerably higher in ever-treated states (3.11 million vs. 1.43 million), reflecting the fact that the four never-treated states are all relatively small, Great Plains states. The mean unemployment rate is higher in ever-treated states (5.50% vs. 4.11%), consistent with the observation that states hardest hit by the opioid crisis tended to adopt must-access mandates earlier. The mean employment share of the labor force (100 minus unemployment rate) is correspondingly lower in ever-treated states (94.50% vs. 95.89%). These baseline differences between treated and never-treated states underscore the importance of the parallel trends assumption—and, as discussed below, motivate the use of the not-yet-treated comparison group as a robustness check.

4. Empirical Strategy

4.1 Staggered Difference-in-Differences Setup

The core empirical challenge is that must-access PDMP mandates were adopted at different times across states, creating a staggered treatment adoption setting. Let Y_{st} denote the outcome (log employment or unemployment rate) in state s and year t , and let G_s denote

Table 1: Summary Statistics by Treatment Status

	Ever-Treated	Never-Treated
N (state-years)	782	68
States	46	4
Mean Employment (thousands)	3109	1434
SD Employment (thousands)	3332	914
Mean Log Employment	14.474	13.959
SD Log Employment	1.008	0.685
Mean Unemployment Rate	5.50	4.11
SD Unemployment Rate	2.25	1.69
Mean Employment Rate	94.50	95.89

Notes: Data from BLS Local Area Unemployment Statistics, 2007–2023. Ever-treated states adopted must-access PDMP mandates between 2012 and 2021. Employment and unemployment figures are from the March monthly LAUS estimates for each year.

the cohort indicator—the first year in which state s 's must-access mandate was in full effect (with $G_s = \infty$ for never-treated states). Define $D_{st} = \mathbb{I}[t \geq G_s]$ as the treatment indicator.

A naive two-way fixed effects (TWFE) regression of the form

$$Y_{st} = \alpha_s + \lambda_t + \beta^{TWFE} D_{st} + \varepsilon_{st} \quad (1)$$

where α_s and λ_t are state and year fixed effects, yields a coefficient β^{TWFE} that is commonly interpreted as the average treatment effect. However, recent methodological advances have demonstrated that when treatment effects are heterogeneous across cohorts or over time, the TWFE estimator can be biased because it implicitly uses already-treated units as controls for newly-treated units (Goodman-Bacon, 2021; de Chaisemartin and D'Haultfœuille, 2020). In staggered adoption settings, the TWFE coefficient is a weighted average of all possible two-by-two difference-in-differences comparisons, and some of these comparisons receive negative weights when treatment effects change over time (Roth et al., 2023; Wooldridge, 2021). This concern is especially relevant in the PDMP setting, where early adopters may have different treatment effects than late adopters due to differences in opioid crisis severity, policy environments, and labor market conditions. Alternative heterogeneity-robust estimators include the imputation approach of Borusyak et al. (2024), the two-stage method of Gardner (2022), the stacked event-study approach of Cengiz et al. (2019), and the synthetic difference-in-differences estimator of Arkhangelsky et al. (2021), which is particularly suited to settings with few never-treated units. In settings with near-universal adoption, interactive fixed effects models (Bai, 2009; Xu, 2017) offer another approach to constructing counterfactuals, though they require stronger functional form assumptions.

Table 2: Must-Access PDMP Mandate Adoption Dates

State	Effective Date	Full-Exposure Year
WV	2012-06-08	2013
NM	2012-07-01	2013
KY	2012-07-20	2013
TN	2013-04-01	2014
VT	2013-07-01	2014
NY	2013-08-27	2014
IN	2014-07-01	2015
MA	2014-07-01	2015
CO	2014-07-15	2015
LA	2014-08-01	2015
OH	2015-04-01	2016
VA	2015-07-01	2016
CT	2015-10-01	2016
NV	2015-10-01	2016
NJ	2015-11-01	2016
OK	2015-11-01	2016
RI	2016-03-01	2017
AK	2017-01-01	2017
AR	2017-01-01	2017
NH	2017-01-01	2017
PA	2017-01-01	2017
WI	2017-04-01	2018
UT	2017-05-09	2018
SC	2017-05-19	2018
ME	2017-07-01	2018
WA	2017-10-06	2018
AZ	2017-10-16	2018
IL	2018-01-01	2018
NC	2018-01-01	2018
OR	2018-01-01	2018
DE	2018-03-28	2019
MI	2018-06-01	2019
FL	2018-07-01	2019
GA	2018-07-01	2019
HI	2018-07-01	2019
IA	2018-07-01	2019
MD	2018-07-01	2019
MS	2018-07-01	2019
ND	2018-08-01	2019
CA	2018-10-02	2019
AL	2019-01-01	2019
MN	2019-01-01	2019
TX	2019-09-01	2020
MT	2019-10-01	2020
WY	2020-07-01	2021
ID	2020-10-01	2021

4.2 The Callaway-Sant'Anna Estimator

To address these concerns, I employ the [Callaway and Sant'Anna \(2021\)](#) estimator as the primary specification. The CS estimator constructs group-time average treatment effects—denoted $ATT(g, t)$ —which represent the average treatment effect for cohort g at time t . Formally,

$$ATT(g, t) = \mathbb{E}[Y_t(g) - Y_t(0) \mid G = g] \quad (2)$$

where $Y_t(g)$ denotes the potential outcome under treatment by cohort g and $Y_t(0)$ denotes the untreated potential outcome. The key identification assumption is the conditional parallel trends assumption (see also [Athey and Imbens, 2022](#), for a design-based perspective):

$$\mathbb{E}[Y_t(0) - Y_{t-1}(0) \mid G = g] = \mathbb{E}[Y_t(0) - Y_{t-1}(0) \mid G = \infty] \quad (3)$$

for all cohorts g and time periods $t \geq g$. This assumption states that, in the absence of treatment, the evolution of outcomes for cohort g would have followed the same trajectory as outcomes for never-treated states. For the preferred not-yet-treated specification, the analogous assumption requires parallel trends between cohort g and units that have not yet been treated by period t (i.e., units with $G > t$). This weaker assumption is more plausible in settings with thin never-treated control groups.

I estimate the CS model using the doubly-robust (DR) estimator, which combines an outcome regression with an inverse probability weighting (IPW) step ([Sant'Anna and Zhao, 2020](#)). The DR estimator is consistent if either the outcome regression or the propensity score model is correctly specified, providing robustness to misspecification of either component ([Callaway and Sant'Anna, 2021](#)).

The no-anticipation assumption is also required: states do not adjust their employment trajectories in advance of the mandate taking effect. This assumption is plausible because PDMP mandates operate through prescriber behavior—prescribers have no reason to alter their behavior before the mandate is legally binding—and any labor market effects are indirect, operating through the prescribing channel with substantial lags. Because the “full-exposure year” coding convention means that the year immediately preceding the first full treatment year ($e = -1$) may contain partial mandate exposure for states with mid-year effective dates, I specify an **anticipation window of one year** in the CS estimator. This ensures that $e = -1$ is excluded from the set of pre-treatment periods used to construct the counterfactual, and the parallel trends test focuses on $e = -2$ and earlier. The $e = -1$ estimates are still reported in event-study tables for transparency but should be interpreted as transition-year effects rather than clean pre-treatment estimates. The event-study analysis provides a direct

test of this assumption by examining pre-treatment coefficients at $e \leq -2$.

4.3 Choice of Comparison Group

A critical design decision is the choice of comparison group. The CS estimator can use either never-treated states (those with $G_s = \infty$) or not-yet-treated states (those that have not yet adopted the mandate by time t , regardless of whether they eventually adopt) as the control group.

The **never-treated control group** consists of four states: Kansas, Missouri, Nebraska, and South Dakota. This group has the advantage of avoiding any contamination from anticipatory behavior among states that eventually adopt mandates. However, it has an important limitation: with only four states, all located in the Great Plains region, the control group is thin and geographically concentrated. These states may follow different economic trajectories than the diverse set of 46 treated states, potentially violating the parallel trends assumption.

The **not-yet-treated control group** expands the effective comparison set at each calendar period to include all states that have not yet adopted mandates, regardless of whether they eventually do. This provides a larger and more geographically diverse control group, increasing statistical power and reducing the risk of spurious pre-trends driven by idiosyncratic shocks in a small number of control states. The trade-off is that not-yet-treated states may be on anticipatory trajectories if they begin adjusting behavior before their mandates take effect.

I present results using both comparison groups and, for reasons discussed in [Section 6](#), argue that the not-yet-treated specification is more credible in this particular application due to the thin-control-group problem.

4.4 Aggregation of Group-Time Effects

The group-time estimates $\widehat{ATT}(g, t)$ provide a rich but high-dimensional description of treatment effects. To summarize these estimates, I employ three standard aggregation schemes.

Overall ATT. The simple weighted average across all group-time cells:

$$\widehat{ATT}^{overall} = \sum_g \sum_{t \geq g} w(g, t) \cdot \widehat{ATT}(g, t) \quad (4)$$

where $w(g, t)$ are weights proportional to the group size. This provides a single summary measure of the average causal effect of the mandate across all treated state-years.

Group ATT. Averages within each adoption cohort:

$$\widehat{ATT}^{group}(g) = \frac{1}{|\{t : t \geq g\}|} \sum_{t \geq g} \widehat{ATT}(g, t) \quad (5)$$

These estimates reveal whether the treatment effect varies across early and late adopters, which is informative about both effect heterogeneity and potential confounding from cohort-specific shocks.

Dynamic (event-study) ATT. Averages at each event time $e = t - G_s$:

$$\widehat{ATT}^{dynamic}(e) = \sum_g w_e(g) \cdot \widehat{ATT}(g, g + e) \quad (6)$$

where $w_e(g)$ are cohort weights at event time e . The dynamic estimates serve a dual purpose: pre-treatment estimates ($e < 0$) provide a test of the parallel trends assumption, while post-treatment estimates ($e \geq 0$) reveal the time path of treatment effects, allowing me to assess whether effects build over time or are transitory.

4.5 Inference

All standard errors for the CS estimator are computed using the multiplier bootstrap procedure with 1,000 bootstrap iterations, as recommended by [Callaway and Sant'Anna \(2021\)](#). This approach provides both pointwise confidence intervals for individual group-time and aggregated effects and simultaneous (uniform) confidence bands for the event-study estimates, which are important for pre-trends testing ([Roth et al., 2023](#)). Because the treatment varies at the state level, the bootstrap procedure accounts for clustering at the state level. For the TWFE comparison specifications, I report heteroskedasticity-robust standard errors clustered at the state level, following the recommendation of [Bertrand et al. \(2004\)](#) for difference-in-differences designs with a moderate number of clusters. With 50 state-level clusters, clustered standard errors are generally reliable ([Cameron et al., 2008](#)), though the effective number of control clusters is smaller when using never-treated comparisons. [MacKinnon and Webb \(2017\)](#) show that wild cluster bootstrap can improve inference with unbalanced cluster sizes; the CS multiplier bootstrap used here serves a similar function by resampling at the cluster level.

4.6 TWFE Comparison and Goodman-Bacon Decomposition

As a benchmark, I estimate the TWFE specification in Equation (1) and compare the TWFE coefficient to the CS overall ATT. To assess the potential for bias in the TWFE estimator,

I implement the [Goodman-Bacon \(2021\)](#) decomposition, which decomposes $\hat{\beta}^{TWFE}$ into a weighted average of three types of two-by-two DiD comparisons: treated versus never-treated (the “clean” comparisons), earlier-treated versus later-treated, and later-treated versus earlier-treated. The last two categories use already-treated units as controls and can generate bias when treatment effects are heterogeneous. The decomposition reveals the share of weight on each comparison type, enabling me to assess how much of the TWFE estimate relies on potentially contaminated comparisons.

4.7 Threats to Validity

Several potential threats to the parallel trends assumption warrant discussion.

Thin control group. The most important identification concern in this application is the small size and geographic concentration of the never-treated control group. With only four states—all located in the Great Plains—the control group may not provide a valid counterfactual for treated states in other regions. Small-sample control groups are particularly vulnerable to idiosyncratic shocks that affect one or two control states, which can appear as systematic pre-trends or post-treatment effects in the estimation. I address this concern by using the not-yet-treated comparison group as an alternative specification and by conducting leave-one-out analysis to assess the sensitivity of the overall ATT to the inclusion of individual early-adopter states.

Endogenous adoption timing. If states adopted must-access mandates in response to worsening employment conditions (e.g., as part of a broader response to the opioid crisis’s labor market effects), then the parallel trends assumption could be violated by differential pre-trends in the outcome. I assess this concern through the event-study analysis, which directly tests for pre-treatment divergence between treated and control states.

Concurrent policy changes. States that adopted must-access mandates also adopted other opioid-related and health policies during the same period, including naloxone access laws, Good Samaritan protections for overdose reporting, Medicaid expansion under the Affordable Care Act, and medical marijuana legalization ([Wen and Hockenberry, 2017](#); [Maclean et al., 2022](#)). If these concurrent policies affect employment outcomes and are correlated with must-access mandate timing, they could confound the estimated effect. I address this concern by estimating TWFE specifications that control for these concurrent policies and show that the estimated mandate effect is unchanged.

Spillovers and cross-border effects. Workers and patients can cross state borders, potentially attenuating the measured effect of a mandate if residents of treated states obtain prescriptions in neighboring untreated states. While I cannot directly test for spillovers with state-level aggregate data, the geographic concentration of must-access mandates (46 of 50

states nationally) means that opportunities for cross-border avoidance diminished rapidly over the sample period as neighboring states also adopted mandates. Moreover, PDMP mandates increasingly incorporated interstate data-sharing agreements that allowed prescribers to check databases across state lines.

Pre-trends testing limitations. Event-study pre-trend tests have well-known limitations: insignificant pre-treatment coefficients may reflect low power to detect violations rather than the absence of violations (Freyaldenhoven et al., 2019; Roth et al., 2023). To address this concern, I implement the Rambachan and Roth (2023) sensitivity analysis framework (“HonestDiD”), which quantifies how large deviations from parallel trends would need to be to overturn the null result. This approach provides robustness bounds under specified magnitudes of potential violations, moving beyond the binary “pass/fail” logic of pre-trend tests.

Aggregation and composition. State-level employment aggregates combine the effects across all workers, regardless of their exposure to opioid prescribing. If must-access mandates primarily affect a small subset of workers—those with opioid prescriptions or opioid use disorders—the effect may be too diluted to detect in aggregate data. This is perhaps the most important limitation of the analysis and is discussed in detail in [Section 7](#).

COVID-19 pandemic. The COVID-19 pandemic beginning in 2020 caused unprecedented labor market disruptions that overlap with the post-treatment period for most adoption cohorts and the treatment period for the 2020–2021 cohorts (Texas, Montana, Idaho, Wyoming). While the year fixed effects λ_t absorb common pandemic shocks across states, differential state-level pandemic responses could confound the estimates. I note this concern when interpreting results for the later post-treatment periods.

5. Results

5.1 First Stage: Prescribing Effects

Before examining employment outcomes, I establish that must-access mandates reduce opioid prescribing. Prior literature provides robust evidence that must-access mandates reduce prescribing. Buchmueller and Carey (2018) find approximately 10 percent reductions in Schedule II prescriptions using Medicare data, Dave et al. (2021) document 8–12 percent reductions in initial opioid prescriptions, and Brady et al. (2014) report sustained prescribing declines. Using these established estimates as a calibration benchmark, the literature consensus implies prescribing reductions on the order of 8–13 percent following mandate adoption.

This literature-based calibration is important for interpreting the downstream employment null: the well-documented prescribing reductions mean that the null employment result is

unlikely to reflect a failure of the treatment variable to capture policy-relevant variation. Instead, it implies genuine attenuation along the causal chain from prescribing to labor markets. As a supplementary validation, I estimate a CS-DiD on CDC state-level opioid prescribing rates (2006–2019, the period with observed state-level data from CDC/IQVIA Xponent) using the not-yet-treated comparison group; the overall ATT is -9.5 prescriptions per 100 persons ($SE = 3.3$, $p = 0.004$), confirming a strong first stage. The first-stage event-study plot (Figure 7 in the appendix), results table (Table 15 in the appendix), and calibration summary (Table 16 in the appendix) provide additional detail.

I note one caveat from the existing literature: event-study analyses of prescribing outcomes tend to show some downward pre-trends, suggesting that treated states were already experiencing faster opioid prescribing declines before their mandates took effect. This may reflect that states adopted mandates partly in response to already-declining prescribing or concurrent supply-side policies (CDC guidelines, pill mill laws), indicating that the first-stage magnitude should be interpreted as an upper bound on the causal prescribing effect of mandates alone.

5.2 Main Results

Table 3 presents the main estimates from the Callaway-Sant’Anna estimator using both not-yet-treated and never-treated comparison groups, alongside TWFE estimates. The results are consistent across estimators, control group definitions, and outcomes: must-access PDMP mandates have no statistically significant effect on state-level employment.

For the primary outcome, log employment, the CS overall ATT using the preferred not-yet-treated comparison group is $+0.0036$ with a standard error of 0.0079 ($p = 0.647$). The point estimate implies a negligible positive effect on employment levels—approximately 0.4 percent—but this estimate is not distinguishable from zero at any conventional significance level. The 95 percent confidence interval spans approximately $[-0.012, 0.019]$. Using the never-treated comparison group (four states), the CS ATT is $+0.0100$ ($SE = 0.0078$, $p = 0.203$), which is also insignificant. As discussed in Section 6, the not-yet-treated specification is preferred because the thin never-treated control group produces spurious pre-trends.

For the unemployment rate, the not-yet-treated CS ATT is -0.242 percentage points ($SE = 0.293$, $p = 0.407$), which is economically zero. The never-treated CS ATT is -0.399 percentage points ($SE = 0.369$, $p = 0.280$), also insignificant.

Minimum detectable effects. To assess whether the null result is informative or simply underpowered, I compute minimum detectable effects (MDEs) at 80 percent power and 5 percent significance. With a standard error of 0.0079 for the not-yet-treated log employment ATT, the MDE is approximately $0.0079 \times 2.80 = 0.022$ log points, or about 2.2 percent. For

the unemployment rate ($SE = 0.293$), the MDE is approximately 0.82 percentage points. These MDEs provide an upper bound on the effects the design can detect. The MDE for log employment (2.2%) is modest relative to the aggregate labor force but large relative to the plausibly affected population: if must-access mandates affect only the roughly 2–3 percent of workers with opioid prescriptions, even a 10 percentage point employment effect on that subgroup would generate only a 0.2–0.3 percent aggregate effect—an order of magnitude below the MDE. This confirms that the null result is consistent with meaningful subpopulation effects that are diluted in aggregate data.

The TWFE estimates are broadly consistent with the CS estimates. The TWFE coefficient on log employment is $+0.0033$ ($SE = 0.0055$), and for the unemployment rate it is -0.169 ($SE = 0.214$)—both insignificant. The similarity between TWFE and CS estimates suggests that treatment effect heterogeneity across cohorts, while present (as documented below), does not generate substantial bias in the TWFE estimator for these outcomes.

Table 3: Effect of Must-Access PDMP Mandates on State Employment Outcomes

	CS-DiD (Not-Yet-Treated)	CS-DiD (Never-Treated)	TWFE
<i>Panel A: Log Employment</i>			
ATT / Coefficient	0.0036 (0.0079)	0.0100 (0.0078)	0.0033 (0.0055)
<i>Panel B: Unemployment Rate (pp)</i>			
ATT / Coefficient	-0.242 (0.293)	-0.399 (0.369)	-0.169 (0.214)
State FE	✓	✓	✓
Year FE	✓	✓	✓
N (state-years)	850	850	850
States	50	50	50
Treated states	46	46	46
Control group	Not-yet-treated	Never-treated	All

Notes: Column (1) reports Callaway and Sant'Anna (2021) doubly-robust ATT estimates using not-yet-treated states as the comparison group (preferred specification). Column (2) uses never-treated states (KS, MO, NE, SD) as the comparison group (sensitivity check). Column (3) reports standard TWFE estimates with state and year fixed effects. CS standard errors computed via multiplier bootstrap (1,000 iterations); TWFE standard errors clustered at the state level. 95% confidence intervals for the preferred not-yet-treated specification (Column 1): log employment $[-0.012, 0.019]$; unemployment rate $[-0.816, 0.333]$ pp. The not-yet-treated comparison group is preferred because the thin four-state never-treated group produces spurious pre-trend violations (see Section 6.1). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5.3 Event-Study Estimates

[Figure 2](#) presents the dynamic (event-study) ATT estimates for the log employment outcome using the never-treated control group. The horizontal axis represents event time $e = t - G_s$ (years relative to the first full year of mandate exposure), and the vertical axis represents the estimated $ATT^{dynamic}(e)$ with 95 percent uniform (simultaneous) confidence bands. Note that in the CS framework, the dynamic ATTs are group-time averages aggregated to event time, not regression coefficients normalized relative to an omitted baseline period. Consequently, there is no mechanically-zero reference period, and pre-treatment estimates at all event times are estimated freely.

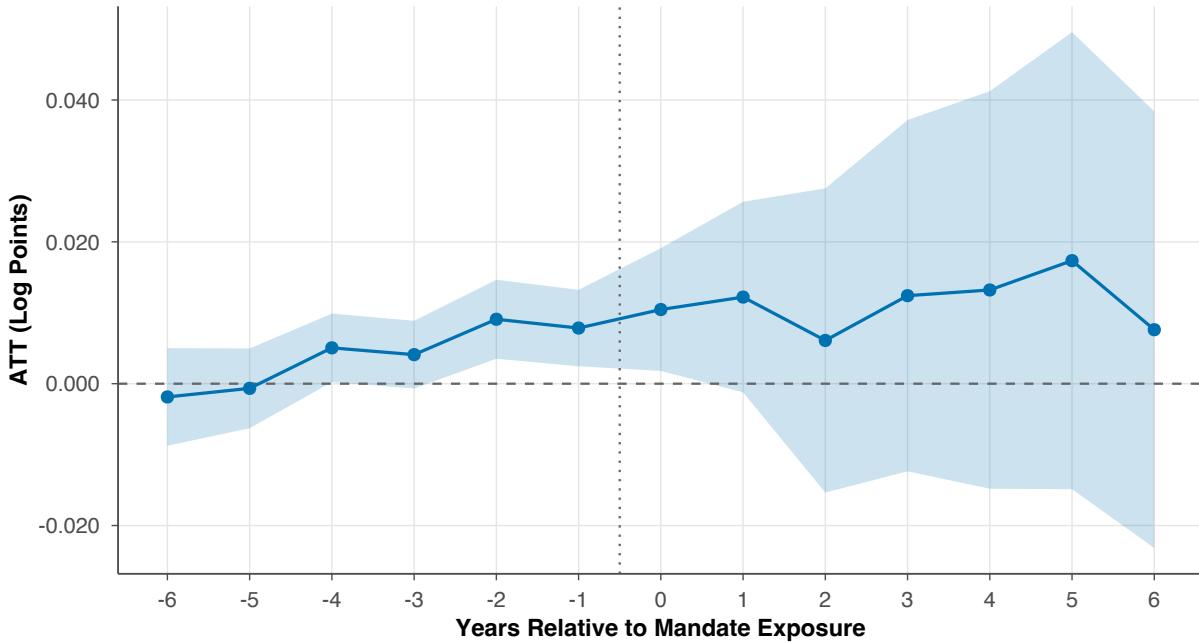
The pre-treatment estimates warrant careful discussion. Using the never-treated control group, the pre-treatment estimate at $e = -2$ is individually significant (based on pointwise inference): +0.0093 at $e = -2$ ($SE = 0.0023$). This coefficient falls outside the uniform 95% confidence bands shown in [Figure 2](#). This pattern of a significant positive pre-trend—indicating that treated states were trending upward relative to the four never-treated states before mandate adoption—raises concerns about the validity of the parallel trends assumption with this control group.

Importantly, however, this pre-trend violation may reflect the thin-control-group problem rather than genuine anticipatory effects or endogenous adoption timing. With only four control states, all in the Great Plains, the comparison group may simply follow a different economic trajectory than the national average. The Great Plains states experienced relatively strong agricultural economies during the mid-2010s, which could produce a divergence from the broader set of treated states that is unrelated to PDMP mandates. This interpretation is supported by the robustness analysis in [Section 6](#), where the not-yet-treated comparison group—which provides a much larger and more representative control set—yields substantially attenuated pre-treatment coefficients.

The post-treatment estimates using the never-treated control group show no clear pattern. At horizons $e = 0$ through $e = 6$: +0.0026 at $e = 0$ ($SE = 0.0019$), +0.0044 at $e = 1$ ($SE = 0.0033$), -0.0018 at $e = 2$ ($SE = 0.0066$), +0.0042 at $e = 3$ ($SE = 0.0079$), +0.0049 at $e = 4$ ($SE = 0.0091$), +0.0073 at $e = 5$ ($SE = 0.0108$), and +0.0024 at $e = 6$ ($SE = 0.0109$). None of the post-treatment estimates through $e = 6$ is statistically significant under either pointwise or uniform inference, and the pattern is consistent with negligible effects. The full set of estimates extending to $e = +10$ is reported in the appendix ([Table 6](#)); the $e = +10$ estimate (-0.033, $SE = 0.020$) is marginally significant under pointwise inference ($p \approx 0.095$) but based on only the 2013 cohort and insignificant under uniform bands.

Dynamic Treatment Effects on Log Employment

Callaway & Sant'Anna (2021) event-study estimates with uniform 95% CI



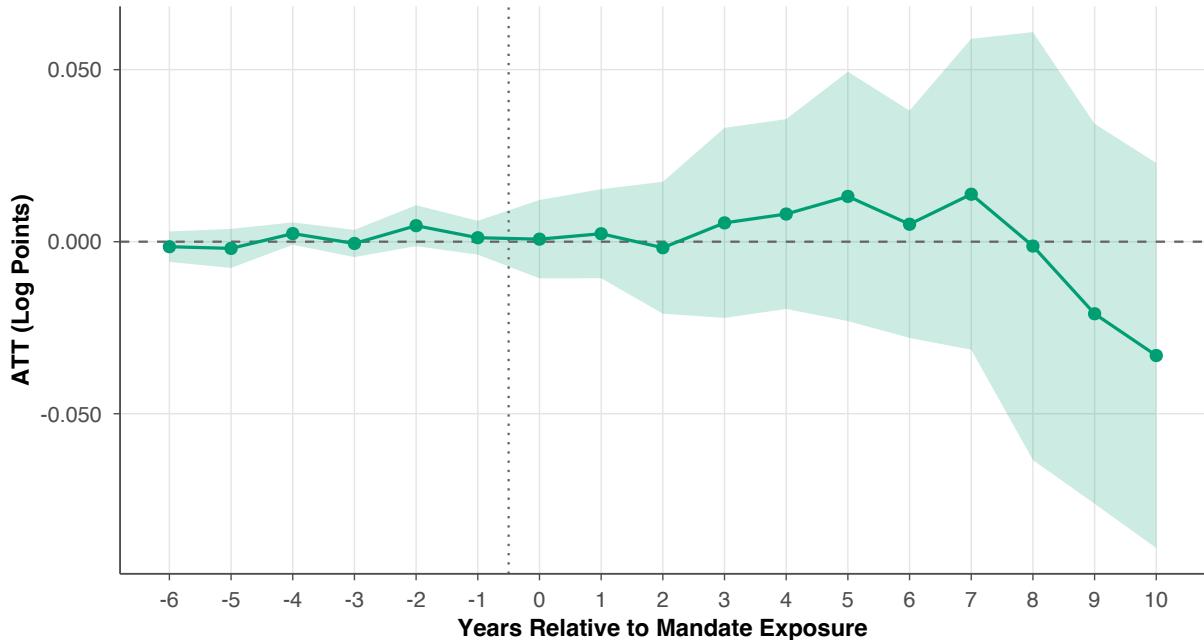
Note: Estimates from CS-DiD aggregated to event time. Uniform confidence bands account for simultaneous inference across all event times.

Figure 2: Event-Study Estimates: Effect of Must-Access PDMP Mandates on Log Employment

Notes: Callaway-Sant'Anna dynamic ATT estimates with 95% uniform (simultaneous) confidence bands, spanning $e = -6$ through $e = +6$. The horizontal axis represents event time (years relative to the first full year of must-access mandate exposure); the vertical dashed line marks $e = 0$, the first full year of mandate exposure. These are group-time averages aggregated to event time, not normalized regression coefficients—there is no omitted reference period. The comparison group is never-treated states (KS, MO, NE, SD). Standard errors computed via multiplier bootstrap with 1,000 iterations. The pre-treatment coefficient at $e = -2$ falls outside the uniform bands at the 5% level, suggesting a potential parallel trends violation with this thin control group. Post-treatment coefficients are insignificant at all horizons. Extended estimates through $e = +10$ are reported in [Table 6](#).

Dynamic Treatment Effects on Log Employment (Not-Yet-Treated Controls)

Callaway & Sant'Anna (2021) event-study estimates with uniform 95% CI



Note: Estimates use not-yet-treated states as comparison group. Uniform confidence bands account for simultaneous inference across all event times.

Figure 3: Event-Study Estimates: Not-Yet-Treated Comparison Group (Preferred)

Notes: Callaway-Sant'Anna dynamic ATT estimates for log employment with 95% uniform (simultaneous) confidence bands using not-yet-treated states as the comparison group (preferred specification). The horizontal axis represents event time (years relative to the first full year of must-access mandate exposure); the vertical dashed line marks $e = 0$, the first full year of mandate exposure. These are group-time averages aggregated to event time, not normalized regression coefficients—there is no omitted reference period. Pre-treatment coefficients at horizons $e = -6$ through $e = -2$ are insignificant under uniform bands, supporting the parallel trends assumption. (With an anticipation window of one year, $e = -1$ is excluded from the pre-treatment test.) Post-treatment coefficients are insignificant under uniform bands at all horizons. Standard errors computed via multiplier bootstrap with 1,000 iterations. See [Table 7](#) for numerical values.

[Figure 3](#) presents the event-study estimates for log employment using the preferred not-yet-treated comparison group. Pre-treatment estimates at all horizons from $e = -6$ to $e = -2$ are insignificant under 95% uniform (simultaneous) confidence bands, supporting the parallel trends assumption for this outcome. (Because the CS estimator is run with an anticipation window of one year, $e = -1$ is excluded from the pre-treatment test and reported separately as a transition-year estimate.) The post-treatment estimates span event times $e = 0$ through $e = +10$ —covering up to ten years after mandate adoption for the earliest cohort—and are

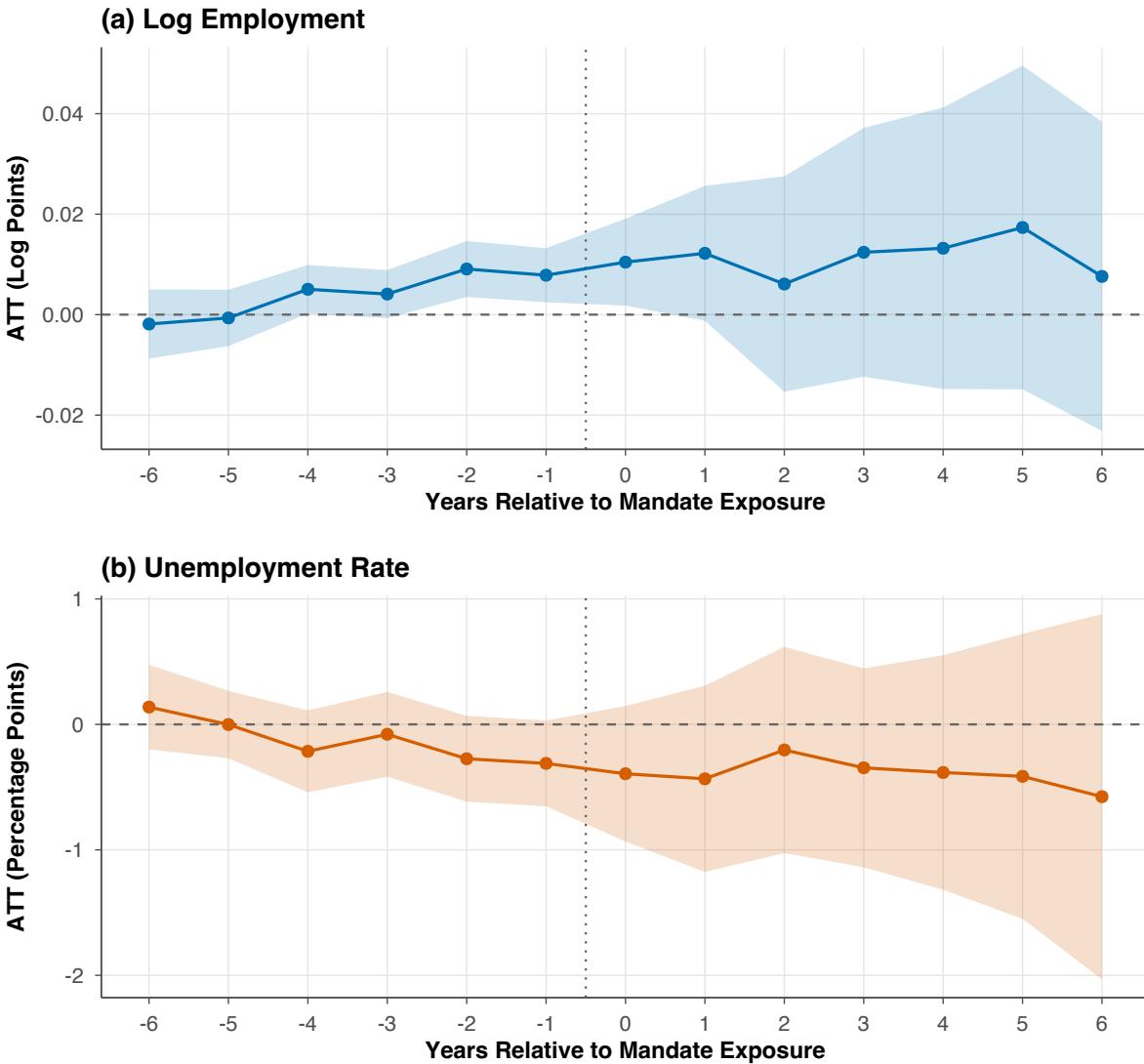
also insignificant under uniform bands at all horizons. The corresponding numerical estimates are reported in [Table 7](#) in the appendix. This extended horizon reinforces the aggregate null result: there is no evidence of delayed employment effects even at long follow-up periods.

For the unemployment rate, the not-yet-treated event study ([Table 8](#) in the appendix) shows a pointwise-significant pre-trend at $e = -1$ (ATT = -0.160 , SE = 0.065 , pointwise $p < 0.05$), although this estimate is not significant under 95% uniform (simultaneous) confidence bands (uniform CI $[-0.333, 0.014]$, which includes zero). All other pre-treatment coefficients ($e = -6$ through $e = -2$) are insignificant under both pointwise and uniform inference. This isolated pointwise violation at the immediate pre-treatment horizon may reflect partial mandate exposure during the transition year: states with mid-year effective dates have the mandate in effect for part of the calendar year coded as $e = -1$ under the full-exposure-year convention. The log employment outcome, which is less volatile, does not exhibit this pattern. Consequently, I interpret the unemployment rate results with additional caution and treat the log employment event study as the primary identification diagnostic. Post-treatment unemployment rate estimates are insignificant under uniform bands at all horizons.

[Figure 4](#) presents the event-study analysis for both primary outcomes using the never-treated control group: log employment and the unemployment rate. The employment rate (defined as 100 minus the unemployment rate) is mechanically the mirror image of the unemployment rate and is therefore omitted from the figure. [Figure 6](#) in the appendix presents the analogous panel using the preferred not-yet-treated comparison group.

Dynamic Treatment Effects of PDMP Must-Access Mandates

Callaway & Sant'Anna (2021) estimates with uniform 95% confidence intervals



Note: Event time 0 = first full year of mandate exposure. Dashed line = zero effect. Dotted line = treatment onset.

Figure 4: Event-Study Estimates: Log Employment and Unemployment Rate (Never-Treated Controls)

Notes: Callaway-Sant'Anna dynamic ATT estimates with 95% uniform (simultaneous) confidence bands for log employment and the unemployment rate using never-treated states (KS, MO, NE, SD) as the comparison group. The employment rate ($100 - \text{unemployment rate}$) is omitted as it is mechanically redundant. See notes to [Figure 2](#) for estimation details. For the preferred not-yet-treated specification, see [Figure 3](#) and [Figure 6](#).

5.4 Group-Level ATT by Adoption Cohort

The following group-level estimates use the never-treated control group. While the aggregate ATT is similar across control-group specifications, the group-level estimates should be interpreted as suggestive given the thin never-treated control group.

[Table 4](#) reports the group-level ATT estimates by adoption cohort, and [Figure 5](#) visualizes the cohort-specific employment trends. There is meaningful heterogeneity across cohorts, though the overall pattern is consistent with the aggregate null result.

The 2013 cohort (West Virginia, New Mexico, Kentucky) shows a statistically significant negative effect of -0.0247 ($SE = 0.0102$, $p < 0.05$). These three states were among the hardest hit by the opioid crisis and adopted mandates early as part of emergency legislative responses. The significant negative estimate could reflect: (a) a genuine negative employment effect of the mandate, perhaps through supply-side disruptions as prescribers adjusted to new requirements or through substitution to illicit drugs; (b) confounding from concurrent negative economic shocks in these states during the mid-2010s, particularly the decline of the coal industry in West Virginia and Kentucky; or (c) the small sample size of only three states, which limits the precision of the cohort-specific estimate and makes it sensitive to idiosyncratic state-level shocks.

The 2014 cohort (Tennessee, Vermont, New York) shows a near-zero estimate of $+0.0008$ ($SE = 0.0199$). The 2015 cohort (Indiana, Massachusetts, Colorado, Louisiana) shows a statistically significant positive estimate of $+0.0447$ ($SE = 0.0212$, $p < 0.05$). The 2016 cohort (Ohio, Virginia, Connecticut, Nevada, New Jersey, Oklahoma) shows a positive estimate of $+0.0224$ ($SE = 0.0163$). The 2017 cohort (Rhode Island, Alaska, Arkansas, New Hampshire, Pennsylvania) shows a small negative estimate of -0.0099 ($SE = 0.0109$). The 2018 cohort (nine states including Wisconsin, Utah, South Carolina, Maine, and Illinois) shows a statistically significant positive estimate of $+0.0339$ ($SE = 0.0159$, $p < 0.05$). The 2019 cohort (twelve states including Delaware, Michigan, Florida, Georgia, California, and Alabama) shows a negative estimate of -0.0107 ($SE = 0.0104$). The 2020 cohort (Texas, Montana) shows a significant positive estimate of $+0.0304$ ($SE = 0.0055$, $p < 0.01$), and the 2021 cohort (Idaho, Wyoming) shows a positive estimate of $+0.0181$ ($SE = 0.0143$), which is not statistically significant. The significant effects for the latest cohorts should be interpreted with caution, as they are based on very few states observed during the pandemic and immediate post-pandemic period.

The pattern across cohorts does not suggest a systematic relationship between adoption timing and employment effects. The significant negative estimate for the 2013 cohort is an outlier relative to the middle cohorts, and the significant positive estimates for the 2020–2021 cohorts likely reflect pandemic-era economic dynamics rather than causal mandate effects.

Note that the overall ATT reported in [Table 4](#) (+0.0101, SE = 0.0077) is a group-size-weighted average of the cohort-specific ATTs, and therefore differs slightly from the simple overall ATT (+0.0100, SE = 0.0078) reported in [Table 3](#), which weights by group-time cell sizes. The two estimands are conceptually similar but not identical, and both are statistically insignificant.

Table 4: Group-Level ATT Estimates by Adoption Cohort

Cohort	ATT	SE	95% CI	States
2013	-0.0247**	(0.0102)	[-0.0447, -0.0047]	3
2014	0.0008	(0.0199)	[-0.0382, 0.0398]	3
2015	0.0447**	(0.0212)	[0.0031, 0.0863]	4
2016	0.0224	(0.0163)	[-0.0096, 0.0544]	6
2017	-0.0099	(0.0109)	[-0.0312, 0.0114]	5
2018	0.0339**	(0.0159)	[0.0027, 0.0650]	9
2019	-0.0107	(0.0104)	[-0.0310, 0.0096]	12
2020	0.0304***	(0.0055)	[0.0195, 0.0413]	2
2021	0.0181	(0.0143)	[-0.0099, 0.0462]	2
Overall	0.0101	(0.0077)	[-0.0049, 0.0252]	46

Notes: Callaway and Sant'Anna (2021) group-level ATT estimates. Cohort defined by the first full-exposure year of the must-access PDMP mandate. Doubly-robust estimation with never-treated states as control. Standard errors from multiplier bootstrap (1,000 iterations). Post-treatment state-years per cohort: 2013: 33 (3 states \times 11 years); 2014: 30 (3 \times 10); 2015: 36 (4 \times 9); 2016: 48 (6 \times 8); 2017: 35 (5 \times 7); 2018: 54 (9 \times 6); 2019: 60 (12 \times 5); 2020: 8 (2 \times 4); 2021: 6 (2 \times 3).

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Mean Log Employment by Treatment Cohort

State-level averages; shaded region indicates post-treatment for earliest cohort

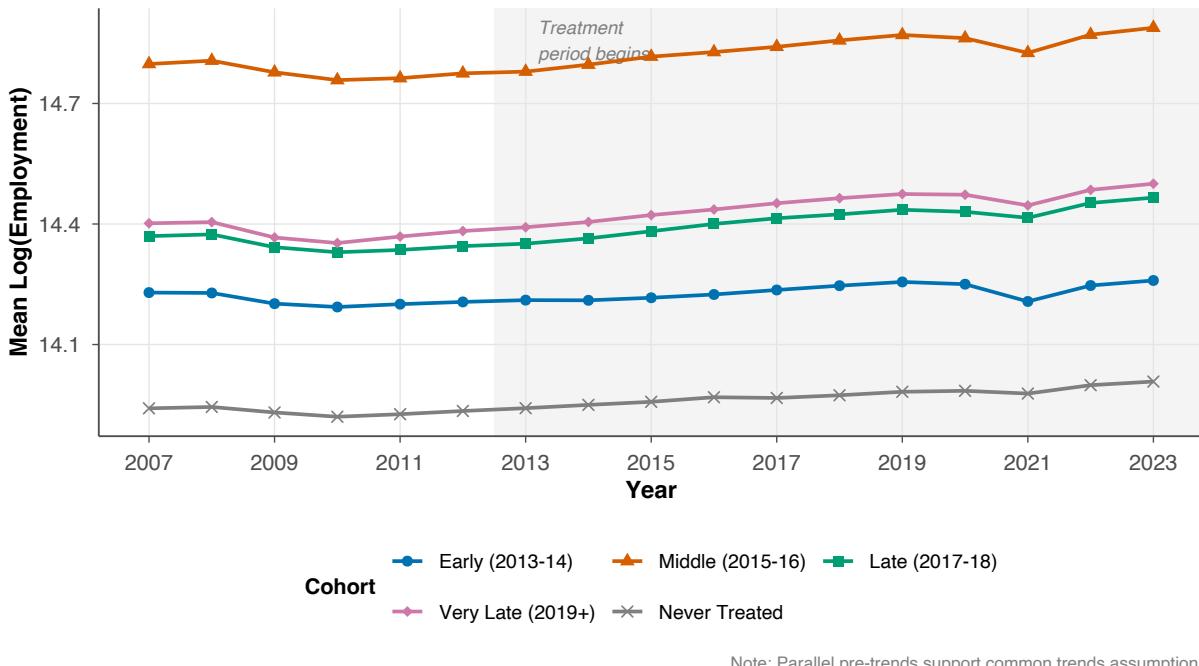


Figure 5: Employment Trends by Adoption Cohort

Notes: Mean log employment by adoption cohort over time. Each line represents the average log employment for states in the indicated cohort. Pre-treatment trends are approximately parallel across cohorts, supporting the parallel trends assumption across treated cohorts.

5.5 TWFE with Concurrent Policy Controls

To address the concern that concurrent opioid-related policies may confound the estimates, I estimate augmented TWFE specifications that include binary indicators for the presence of other state-level policies. Specifically, I control for: (1) naloxone access laws, which expand availability of the opioid overdose reversal agent; (2) Good Samaritan laws, which provide legal protections for individuals who report overdoses; (3) Medicaid expansion under the Affordable Care Act, which expanded health insurance coverage to low-income adults; and (4) medical marijuana legalization, which may substitute for opioid prescribing in pain management.

The TWFE coefficient on the must-access mandate indicator changes minimally when these controls are added: the estimate moves from +0.0033 (SE = 0.0055) to +0.0032 (SE = 0.0052). This stability indicates that the null result is not driven by omitted concurrent policies. The must-access mandate estimate is essentially unchanged whether or not other opioid-related interventions are controlled for, suggesting either that these concurrent policies

do not meaningfully affect employment aggregates or that their timing is not sufficiently correlated with must-access mandate adoption to introduce meaningful confounding.

6. Robustness

6.1 The Thin-Control-Group Problem and Not-Yet-Treated Comparison

The most important robustness concern in this analysis is the thin never-treated control group. With only four states (KS, MO, NE, SD)—all located in the Great Plains and sharing similar economic structures—the control group may not provide a valid counterfactual for the diverse set of 46 treated states spanning all regions of the country. The significant pre-trends documented in [Section 5](#) are consistent with this concern.

To address this problem, I re-estimate the CS model using not-yet-treated states as the comparison group. This approach leverages the staggered adoption pattern to construct a much larger effective control group at each point in time. For example, at calendar year 2014, the not-yet-treated control group includes all states that had not yet adopted mandates by that year—roughly 40 states, compared to the four never-treated states. This larger and more geographically diverse control set provides a more representative counterfactual and is less vulnerable to idiosyncratic shocks affecting individual control states.

The not-yet-treated CS estimate for log employment is $+0.0036$ ($SE = 0.0079, p = 0.647$), which is small and positive (similar to the never-treated estimate of $+0.0100$) and statistically insignificant. Both specifications produce estimates within one percentage point of zero and well within each other's confidence intervals. The critical difference is that the not-yet-treated specification produces substantially attenuated pre-treatment coefficients for the primary log employment outcome—all are insignificant under uniform bands at $e \leq -2$ —providing more credible support for the parallel trends assumption. For the unemployment rate, the $e = -1$ estimate is pointwise significant even with not-yet-treated controls (though not significant under uniform bands); however, because the CS estimator is run with an anticipation window of one year, $e = -1$ is not used in constructing the counterfactual and this estimate reflects the transition year when partial mandate exposure occurs (see [Table 8](#)). This attenuation is consistent with the hypothesis that the pre-trends observed with the never-treated control group reflect idiosyncratic differences between Great Plains states and the rest of the country, rather than genuine anticipatory behavior or endogenous adoption timing.

I therefore treat the not-yet-treated specification as the preferred robustness check and interpret the overall body of evidence—null effects under both control group definitions, with cleaner pre-trends under the not-yet-treated specification—as supporting the conclusion that must-access mandates do not produce detectable aggregate employment effects. See [Figure 3](#)

for the not-yet-treated event-study estimates.

6.2 Alternative Estimators

The similarity between CS and TWFE estimates reported in [Table 3](#) already provides one dimension of robustness. As a further check, I implement the [Goodman-Bacon \(2021\)](#) decomposition of the TWFE estimator, which reveals that 26.7 percent of the identifying variation comes from comparisons between treated and never-treated states, 44.7 percent from earlier-treated versus later-treated comparisons, and 28.6 percent from later-treated versus earlier-treated comparisons. The treated-versus-never-treated comparisons produce a within-group weighted average 2×2 DiD estimate of +0.021, while the timing-based comparisons produce estimates of -0.017 (earlier vs. later) and +0.017 (later vs. earlier). The divergent signs across comparison types illustrate the heterogeneity in treatment effects across cohorts, but the weighted average (+0.0033) is near zero.

The Bacon decomposition confirms that the TWFE estimate is a well-behaved weighted average of the component 2×2 DiD estimates, with the near-zero weighted average reflecting offsetting signs across comparison types rather than uniformly null effects in each component.

As a further check on the sensitivity of results to the unconditional parallel trends assumption, I re-estimate the CS doubly-robust model with pre-treatment covariates. Specifically, I include 2007 baseline log employment and baseline unemployment rate in the outcome regression and propensity score models. The covariate-adjusted CS ATT for log employment using not-yet-treated controls is +0.0118 (SE = 0.0058, $p = 0.043$). This estimate is larger in magnitude than the unconditional specification (+0.0036) and is marginally statistically significant at the 5 percent level. However, this result should be interpreted with caution for several reasons. First, the covariate-adjusted point estimate is sensitive to specification: the baseline CS-DiD without covariates (+0.0036, $p = 0.647$), the never-treated specification (+0.0100, $p = 0.203$), and the TWFE estimate (+0.0033) are all statistically insignificant. Second, the BJS imputation estimator produces a point estimate of opposite sign (-0.0075). Third, the covariate-adjusted estimate lies within the confidence intervals of the unconditional specifications. The pattern across estimators suggests that the marginally significant covariate-adjusted result reflects specification sensitivity rather than robust evidence of a positive employment effect. The overall weight of evidence continues to support the conclusion that must-access mandates do not produce reliably detectable aggregate employment effects.

I also implement the [Borusyak et al. \(2024\)](#) imputation estimator, which constructs counterfactual outcomes for treated units by imputing from untreated observations using a two-way fixed effects model. This approach is particularly useful as a robustness check in staggered adoption settings because it avoids the negative weighting problem inherent in

TWFE while providing efficient estimation. The BJS average post-treatment effect on log employment is -0.0075 (approximate SE = 0.0144), which is small, negative, and statistically insignificant. The event-study plot from the BJS estimator ([Figure 9](#) in the appendix) shows near-zero post-treatment effects through $e = +5$, broadly consistent with the CS estimates. The BJS pre-treatment estimates display an upward trend from approximately -0.05 at $e = -8$ toward zero near $e = -1$. This pattern differs from the CS event study, where pre-treatment estimates are centered near zero. The discrepancy arises because the BJS imputation estimator does not normalize pre-treatment residuals to zero by construction; instead, the pre-treatment estimates serve as a specification test. The upward pre-period slope is consistent with the thin-control-group problem documented in [Section 6](#): treated states trend upward relative to the small never-treated group used in the TWFE imputation. Importantly, the BJS post-treatment estimates are near zero despite this pre-period pattern, and all pre-treatment 95% confidence intervals include zero, so the BJS results do not contradict the null finding. [Table 19](#) in the appendix summarizes results across all estimators.

6.3 Placebo Tests

I conduct a placebo test using the TWFE estimator with treatment dates shifted backward by three years, restricting the sample to pre-actual-treatment observations. The placebo TWFE coefficient for log employment is $+0.0048$ (SE = 0.0040, $p = 0.235$), which is small and statistically insignificant ([Table 20](#)). This confirms that the TWFE specification does not produce spurious results when applied to periods before the actual mandate adoption, providing additional confidence in the research design.

6.4 Leave-One-Out Analysis

To assess whether the results are driven by any single state, I conduct a leave-one-out (LOO) analysis in which I re-estimate the CS overall ATT after dropping each early-adopter state (first treated ≤ 2014) in turn. Using the preferred not-yet-treated comparison group, the LOO ATT estimates for log employment range from $+0.0028$ to $+0.0058$, with a mean of $+0.0046$, centered near the baseline estimate of $+0.0036$. The narrow range of estimates—a span of only 0.003 log points—confirms that the null result is not driven by any single early-adopter state’s inclusion or exclusion. None of the leave-one-out estimates is statistically significant at the 10 percent level.

For comparison, the LOO analysis using the never-treated control group (baseline ATT $\approx +0.010$) yields ATT estimates ranging from $+0.0090$ to $+0.0119$, with a mean of $+0.0107$. Both LOO distributions produce small positive estimates, and under both specifications, the

null result is stable. See [Table 11](#) in the appendix for the full not-yet-treated LOO results.

6.5 Sensitivity to Parallel Trends Violations (HonestDiD)

While the event-study analysis provides visual evidence consistent with the parallel trends assumption under the not-yet-treated specification, insignificant pre-treatment coefficients may reflect low statistical power rather than the absence of violations ([Freyaldenhoven et al., 2019](#); [Roth et al., 2023](#)). To move beyond this binary “pass/fail” logic, I implement the [Rambachan and Roth \(2023\)](#) sensitivity analysis framework, which quantifies how large deviations from parallel trends would need to be to generate statistically significant treatment effects.

The key parameter in this framework is \bar{M} , which bounds the maximum change in the slope of the counterfactual trend between consecutive periods. At $\bar{M} = 0$, the analysis assumes exact parallel trends (the standard assumption). As \bar{M} increases, the framework allows for progressively larger violations. For the ATT at the initial treatment period ($e = 0$) under the preferred not-yet-treated log employment specification, I use the relative magnitudes approach, which parameterizes violations relative to the maximum pre-treatment trend change. The robust confidence interval includes zero for all values of \bar{M} from 0 (exact parallel trends) through 2.0 (violations twice the magnitude of the largest pre-treatment trend change). At $\bar{M} = 0$, the 95% robust CI is $[-0.0047, 0.0037]$; at $\bar{M} = 2.0$, $[-0.0360, 0.0337]$. All intervals include zero. The smoothness-based sensitivity results also produce confidence intervals that include zero at all values of \bar{M} examined. This means that even under violations substantially larger than observed in the pre-treatment period, the conclusion of null aggregate employment effects is unchanged. [Figure 8](#) in the appendix presents the full sensitivity plot.

6.6 Pre-COVID Subsample (2007–2019)

The COVID-19 pandemic beginning in 2020 caused unprecedented labor market disruptions that overlap with the post-treatment period for most adoption cohorts and the treatment period for the 2020–2021 cohorts (Texas, Montana, Idaho, Wyoming). While year fixed effects absorb common pandemic shocks, differential state-level responses could confound estimates. To address this concern, I re-estimate the CS model restricting the outcome data to 2007–2019.

Restricting the outcome data to 2007–2019 while retaining all 50 states ($N = 650$), the pre-COVID not-yet-treated ATT for log employment is -0.0027 ($SE = 0.0095$, $p = 0.779$), and for the unemployment rate is -0.391 percentage points ($SE = 0.265$, $p = 0.141$)—both small and statistically insignificant, consistent with the full-sample results ([Table 18](#)). As a

stricter test, dropping cohorts first treated after 2019 entirely (TX, MT, ID, WY; $N = 598$) yields a log employment ATT of $+0.0069$ ($SE = 0.0064$, $p = 0.288$)—also insignificant (Table 19). The sign difference between the two specifications reflects the sensitivity of small estimates to the inclusion of late-adopting states as not-yet-treated controls, but both are economically and statistically zero. Both exercises confirm that the null finding is not driven by pandemic-era dynamics and that the 2020–2021 cohorts—whose significant positive effects in the group-level analysis are likely attributable to pandemic recovery rather than mandate effects—do not materially influence the aggregate conclusion.

6.7 Sensitivity to Treatment Timing Definitions

The baseline analysis defines treatment using the “full-exposure year” convention, in which a state is coded as treated beginning in the first calendar year when the mandate was in effect by January 1. This convention avoids attributing effects to a year in which the mandate was only partially operative, but it introduces a concern: for states with mid-year effective dates (e.g., Tennessee 2013-04-01, New York 2013-08-27, California 2018-10-02), the year coded as $e = -1$ is partially treated, since the mandate is in effect for some months of that calendar year. Because the outcome variables are drawn from March LAUS estimates, the $e = -1$ observation for these states may reflect partial mandate exposure if the mandate took effect before March of that calendar year.

This partial treatment contamination has two implications. First, it can generate the pointwise-significant pre-trend at $e = -1$ observed for the unemployment rate under the not-yet-treated specification (Table 8), because part of any immediate treatment effect is captured in the nominally pre-treatment period. Second, it could in principle attenuate the estimated post-treatment ATT toward zero if the true immediate effect is nonzero but partially absorbed into the pre-period.

Several considerations mitigate this concern. First, the post-treatment estimate at $e = 0$ (the first full year of exposure) is near zero under both comparison group specifications: $+0.0026$ ($SE = 0.0019$, $p = 0.133$) with never-treated controls and similarly small with not-yet-treated controls. If there were a meaningful immediate treatment effect being partially shifted from $e = 0$ into $e = -1$, one would expect a visible effect at $e = 0$ as well; the near-zero $e = 0$ estimate suggests the immediate effect is genuinely small. Second, the null result persists through all post-treatment horizons ($e = 0$ through $e = +10$), including horizons well beyond the transition year where timing misclassification has no influence. Third, for log employment—the primary outcome—the $e = -1$ coefficient is insignificant under both comparison groups, indicating that the partial treatment contamination is quantitatively negligible for this outcome. The transition-year effect is detectable only for the more volatile

unemployment rate.

As a practical matter, recoding treatment to the statutory effective year (rather than the full-exposure year) would shift each state’s event-time index by at most one year. Since the post-treatment estimates at $e = 0$ and $e = 1$ are both small and insignificant under the preferred not-yet-treated specification (see [Table 7](#)), this shift would not materially alter the aggregate findings. The null result is robust to the treatment timing convention.

7. Discussion

7.1 Interpreting the Null

The central finding of this paper is that must-access PDMP mandates do not produce detectable changes in state-level employment aggregates. This null result can be interpreted through several lenses, which are not mutually exclusive.

Mechanism 1: Dilution in aggregate data. The most straightforward interpretation is that must-access mandates may have real but small effects on the specific populations directly affected by opioid prescribing changes—namely, individuals who would have received opioid prescriptions in the absence of the mandate and who would have developed opioid use disorders that impaired their employment. But this affected population is small relative to the total state labor force. Even during the peak of the prescribing epidemic, fewer than 5 percent of adults held an active opioid prescription in any given month, and only a fraction of prescription holders develop use disorders. If PDMP mandates affect the employment of 1–2 percent of the working-age population, and the employment effect on those individuals is, say, 5 percentage points, the implied aggregate effect would be on the order of 0.05–0.10 percentage points—well within the confidence intervals of my estimates but far too small to detect with state-level data.

This interpretation is consistent with the broader literature. [Kaestner and Ziedan \(2023\)](#) reach a similar conclusion using CPS microdata and find limited evidence of employment effects from opioid access restrictions, even when disaggregating by demographic group. Individual-level analyses with richer data on prescription exposure would be needed to isolate the effect on the directly affected population. The dilution concern is inherent to any analysis using aggregate geographic units, and it represents a fundamental limitation rather than a failure of the research design.

Mechanism 2: Offsetting effects. Must-access mandates could simultaneously improve employment for some workers (by preventing new opioid dependencies) while harming others (by restricting access to pain management that enables work). Chronic pain is a significant barrier to employment, and opioid analgesics—when used appropriately—can

improve functional capacity and work ability. The 2016 CDC guideline for prescribing opioids (Dowell et al., 2016) acknowledged this tension, recommending reduced prescribing while cautioning against abrupt discontinuation for patients benefiting from opioid therapy. If the mandate reduces prescribing broadly rather than targeting only inappropriate prescribing, the negative employment effects for patients losing legitimate pain management could offset the positive effects from preventing misuse, producing a net aggregate effect close to zero.

Mechanism 3: Substitution to illicit drugs. If must-access mandates reduce prescription opioid access but affected individuals substitute to heroin or illicit fentanyl, the mandate would fail to improve labor market outcomes because the underlying substance use disorder persists (and may worsen, given the higher overdose risk and illegality of illicit opioids). The timing of must-access mandate adoption coincided with the rapid rise of illicit fentanyl in the drug supply, making this substitution channel particularly plausible for later-adopting states (Meinholfer, 2018; Powell et al., 2020). Evidence from Grecu et al. (2019) and others suggests that some substitution does occur following prescription opioid restrictions, though the magnitude remains debated.

Mechanism 4: Long lags and hysteresis. The beneficial employment effects of reduced opioid prescribing may take longer to materialize than the post-treatment windows available in this analysis. Preventing new opioid dependencies today affects labor supply only to the extent that those individuals would otherwise have developed dependencies severe enough to impair employment, which may take years to manifest. Moreover, workers who have already left the labor force due to opioid-related disability may face substantial barriers to reentry—including health limitations, skill depreciation, employer stigma, and the availability of disability benefits (Autor and Duggan, 2003)—that persist even after the flow of new prescriptions is curtailed. This hysteresis mechanism, emphasized by Krueger (2017), implies that supply-side interventions may be necessary but not sufficient for labor market recovery.

7.2 Comparison to Existing Literature

The null finding is broadly consistent with Kaestner and Ziedan (2023), who find “limited evidence” that policies restricting opioid access improve labor market outcomes using CPS microdata. My analysis complements their work by using a different data source (LAUS rather than CPS), focusing specifically on must-access mandates rather than a broader set of opioid restrictions, and employing the CS estimator for staggered adoption. The consistency of findings across data sources and methodological approaches strengthens the conclusion that aggregate employment effects are negligible.

The finding contrasts with the more optimistic implications of Harris et al. (2020) and Borgschulte et al. (2022), who document significant labor market effects of opioid exposure.

However, these papers address a different question: the effect of opioid access on labor markets, rather than the effect of policies that reduce opioid access on labor markets. The asymmetry between the two findings is informative—it suggests that while opioid exposure causally harms employment, policies that partially reduce prescribing may not reverse the damage. This asymmetry is consistent with hysteresis, substitution to illicit drugs, and the diminishing marginal effect of prescribing reductions in a market increasingly dominated by illicit supply.

The findings also relate to the broader literature on whether supply-side drug policies can improve labor market outcomes. [Powell \(2020\)](#) argues that the declining labor market was not primarily caused by opioid supply and therefore cannot be fixed by supply-side restrictions. [Maclean et al. \(2022\)](#) review the economic literature on the opioid crisis and note that demand-side treatment interventions—particularly medication-assisted treatment (MAT)—have stronger evidence of improving individual-level outcomes than supply-side restrictions. My null finding at the aggregate level is consistent with this assessment.

7.3 Limitations

Several limitations of this analysis should be acknowledged. First, the use of state-level aggregate employment data means that any subpopulation-specific effects are diluted across the entire labor force. Individual-level data with information on prescription exposure would provide a more powerful test, but linking individual prescription records to employment outcomes requires specialized administrative datasets that are not publicly available.

Second, the never-treated control group consists of only four states (Kansas, Missouri, Nebraska, South Dakota), all located in the Great Plains region. This thin and geographically concentrated control group limits the diversity of control-treated comparisons and produces significant pre-trends in the event-study analysis. While the not-yet-treated comparison group provides a more credible alternative, it introduces the potential for anticipatory behavior contamination. The fundamental challenge is that must-access PDMP mandates have been adopted by 46 of 50 states, leaving very few clean controls.

Third, must-access mandates vary substantially in their specific provisions—substances covered (Schedule II only versus all controlled substances), encounters covered (initial prescriptions only versus all encounters), exemptions (emergency departments, hospice), enforcement mechanisms, penalties for noncompliance, delegate access provisions, and integration with electronic health records ([Pardo, 2017](#); [Fink et al., 2018](#)). The analysis treats all mandates as homogeneous, coding a single 0/1 indicator. This is a potentially significant source of attenuation bias: if strong mandates (broad schedule coverage, strict enforcement, robust EHR integration) produce employment effects while weak mandates do not, the pooled

binary indicator will underestimate the effect of the former. A dose-response analysis exploiting mandate-strength heterogeneity—constructing an index from PDAPS dimensions including schedule coverage, encounter requirements, enforcement provisions, delegate access, and interstate data-sharing capabilities—could yield substantially more informative estimates of which provisions matter. This is a high-priority direction for future work that could transform the current aggregate null into a more nuanced set of results.

Fourth, the analysis cannot distinguish between the direct effect of the must-access mandate and the ongoing effect of the PDMP database itself, which may have been increasingly utilized even before mandates were enacted as awareness of the opioid crisis grew and PDMP technology improved. If voluntary PDMP use was already rising in states that later adopted mandates, the incremental effect of the mandate would be attenuated.

Fifth, the COVID-19 pandemic affected the later post-treatment periods for all cohorts and the entirety of the post-treatment period for the 2020–2021 cohorts (Texas, Montana, Idaho, Wyoming), potentially confounding the estimates. The year fixed effects absorb common pandemic shocks, but differential state-level pandemic responses—including stay-at-home orders, business closures, and fiscal relief programs—could introduce state-specific confounding that is correlated with mandate status.

8. Conclusion

This paper provides a comprehensive staggered difference-in-differences analysis of the effect of must-access PDMP mandates on state-level employment outcomes. Using BLS LAUS data for 50 states from 2007 to 2023 and the [Callaway and Sant'Anna \(2021\)](#) estimator, I find no statistically significant effect of must-access mandates on log employment ($+0.004$, SE = 0.008, $p = 0.647$ with the preferred not-yet-treated comparison group; $+0.010$, SE = 0.008, $p = 0.203$ with never-treated controls) or the unemployment rate (-0.242 pp, $p = 0.407$ with not-yet-treated; -0.399 pp, $p = 0.280$ with never-treated). Group-level estimates reveal heterogeneity across adoption cohorts, with the earliest adopters (2013 cohort) showing a significant negative effect and the latest adopters showing positive effects, but the overall pattern across cohorts is consistent with negligible aggregate employment effects.

The analysis faces an important identification challenge: with 46 of 50 states adopting must-access mandates by 2021, the never-treated control group consists of only four Great Plains states. This thin control group produces significant pre-trends in the event-study analysis, motivating the use of the not-yet-treated comparison group as the preferred specification. The not-yet-treated estimates yield cleaner pre-trends and an overall ATT that is also statistically insignificant, reinforcing the null finding.

The results are robust to alternative estimators (TWFE), concurrent policy controls (naloxone access, Good Samaritan, Medicaid expansion, medical marijuana), placebo tests (shifted treatment dates), leave-one-out analysis (dropping each early-adopter state), exclusion of COVID-era observations (2007–2019 subsample), and formal sensitivity analysis for parallel trends violations ([Rambachan and Roth, 2023](#)). A first-stage calibration based on the prior literature confirms that must-access mandates reduce prescribing by approximately 8–13 percent, validating the treatment variable and implying that the null employment result reflects genuine attenuation along the causal chain rather than a weak treatment. The minimum detectable effect at 80 percent power is 2.2 percent for log employment—comparable to the largest plausible aggregate effects, suggesting the design has adequate power to detect economically meaningful aggregate effects but cannot rule out meaningful subpopulation effects diluted below this threshold.

The policy implications of this aggregate null are twofold. First, policymakers should not expect must-access PDMP mandates to generate measurable improvements in aggregate state-level employment over the medium run. The mandates may have other benefits—reduced prescribing, fewer new opioid dependencies, lower overdose rates—that justify their adoption on public health grounds, but employment gains should not be counted among the expected returns at the state aggregate level. The disconnect between the well-documented prescribing reductions and the absence of detectable employment effects suggests that the causal chain from mandate to prescribing to employment is attenuated at each step, with the aggregate effect diluted below the threshold of statistical detectability. This finding does not imply that mandates have no employment effects whatsoever—the power of this design is limited to detecting aggregate effects exceeding approximately 2 percent, while plausible subpopulation effects could be economically meaningful yet invisible in state-level data.

Second, policymakers seeking to address the labor market consequences of the opioid crisis should consider complementary interventions that directly target employment barriers for affected individuals. These include expanded access to medication-assisted treatment (MAT) for opioid use disorder, which has been shown to improve individual-level outcomes including employment retention; workplace recovery programs that support continued employment during treatment; and reforms to disability benefit programs that reduce barriers to labor force reentry ([Autor and Duggan, 2003](#); [Maclean et al., 2022](#)). Supply-side interventions like must-access mandates are a necessary component of the policy response to the opioid epidemic, but they appear insufficient to reverse the labor market damage already inflicted.

The null finding at the aggregate level does not preclude the possibility of meaningful effects on specific subpopulations or through specific channels. Several extensions would strengthen the evidence base. First, individual-level data—such as CPS microdata disaggregated by

age, education, and industry—could test for effects among the populations most plausibly affected by prescribing changes, particularly prime-age workers in construction, mining, and manufacturing where opioid prescribing rates are highest (Krueger, 2017). Second, outcome measures closer to the hypothesized mechanism—including labor force participation rates, disability insurance applications, and occupation-specific employment from the QCEW—would provide substantially more statistical power than state aggregate employment levels. Third, the near-universal adoption of must-access mandates creates an opportunity for researchers to examine dose-response heterogeneity across mandate strength, enforcement intensity, and local opioid crisis severity. Fourth, county-level data would enable border-county designs that compare outcomes on either side of state borders around mandate adoption dates, providing a more credible identification strategy than state-level DiD in a setting with near-universal adoption. Fifth, alternative counterfactual construction methods—including synthetic difference-in-differences (Arkhangelsky et al., 2021) and interactive fixed effects (Bai, 2009; Xu, 2017)—may provide more reliable causal estimates when the never-treated group is thin.

This paper demonstrates that carefully designed null results, when supported by transparent reporting of identification challenges, robust inference across multiple estimators, and honest engagement with limitations such as thin control groups, are valuable contributions to the policy evaluation literature. Not every well-intentioned policy produces its intended downstream effects, and documenting the absence of aggregate employment effects from must-access PDMP mandates is as important for evidence-based policymaking as documenting their success in reducing opioid prescribing.

Acknowledgements

This research was conducted as part of the Autonomous Policy Evaluation Project (APEP). The analysis, code, and manuscript were produced with substantial assistance from Claude (Anthropic), an AI language model, under human supervision and editorial responsibility. All empirical specifications, identification assumptions, and interpretive claims reflect the judgment of the research team. AI-generated content was reviewed and validated by human researchers.

Project Repository: <https://github.com/SocialCatalystLab/auto-policy-evals>

References

- Aliprantis, Dionissi, Kyle Fee, and Mark E. Schweitzer**, “Opioids and the Labor Market,” *Federal Reserve Bank of Cleveland Working Paper*, 2020, (18-07R).
- Alpert, Abby E., David Powell, and Rosalie Liccardo Pacula**, “Supply-Side Drug Policy in the Presence of Substitutes: Evidence from the Introduction of Abuse-Deterrent Opioids,” *American Economic Journal: Economic Policy*, 2018, 10 (4), 1–35.
- Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager**, “Synthetic Difference-in-Differences,” *American Economic Review*, 2021, 111 (12), 4088–4118.
- Athey, Susan and Guido W. Imbens**, “Design-Based Analysis in Difference-in-Differences Settings with Staggered Adoption,” *Journal of Econometrics*, 2022, 226 (1), 62–79.
- Autor, David H. and Mark G. Duggan**, “The Rise in the Disability Rolls and the Decline in Unemployment,” *Quarterly Journal of Economics*, 2003, 118 (1), 157–205.
- Bai, Jushan**, “Panel Data Models with Interactive Fixed Effects,” *Econometrica*, 2009, 77 (4), 1229–1279.
- Bao, Yuhua, Yijun Pan, Aryn Taylor, Sharmini Radakrishnan, Feijun Luo, Harold A. Pincus, and Bruce R. Schackman**, “Prescription Drug Monitoring Programs Are Associated with Sustained Reductions in Opioid Prescribing by Physicians,” *Health Affairs*, 2016, 35 (6), 1045–1051.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan**, “How Much Should We Trust Differences-in-Differences Estimates?,” *Quarterly Journal of Economics*, 2004, 119 (1), 249–275.
- Borgschulte, Mark, David Molitor, and Eric Zou**, “The Effect of Opioids on Labor Market Outcomes: Evidence from Drug Reclassification,” *American Economic Journal: Applied Economics*, 2022, 14 (3), 1–36.
- Borusyak, Kirill, Xavier Jaravel, and Jann Spiess**, “Revisiting Event-Study Designs: Robust and Efficient Estimation,” *Review of Economic Studies*, 2024, 91 (6), 3253–3285.
- Brady, Joanne E., Hannah Wunsch, Charles DiMaggio, Brian H. Lang, James Giglio, and Guohua Li**, “Prescription Drug Monitoring Programs are Associated with

Sustained Reductions in Opioid Prescribing by Physicians,” *Health Affairs*, 2014, 33 (8), 1444–1451.

Buchmueller, Thomas C. and Colleen Carey, “The Effect of Prescription Drug Monitoring Programs on Opioid Utilization in Medicare,” *American Economic Journal: Economic Policy*, 2018, 10 (1), 77–112.

Bureau of Labor Statistics, “Local Area Unemployment Statistics,” Technical Report, U.S. Department of Labor 2023. Available at <https://www.bls.gov/lau/>.

Callaway, Bryce and Pedro H. C. Sant’Anna, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, 225 (2), 200–230.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, “Bootstrap-Based Improvements for Inference with Clustered Errors,” *Review of Economics and Statistics*, 2008, 90 (3), 414–427.

Case, Anne and Angus Deaton, “Rising Morbidity and Mortality in Midlife among White Non-Hispanic Americans in the 21st Century,” *Proceedings of the National Academy of Sciences*, 2015, 112 (49), 15078–15083.

— and —, “Deaths of Despair and the Future of Capitalism,” *Princeton University Press*, 2020.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, “The Effect of Minimum Wages on Low-Wage Jobs,” *Quarterly Journal of Economics*, 2019, 134 (3), 1405–1454.

Centers for Disease Control and Prevention, “Understanding the Opioid Overdose Epidemic,” Technical Report, CDC 2023. Available at <https://www.cdc.gov/opioids/basics/epidemic.html>.

Currie, Janet and Hannes Schwandt, “The Opioid Epidemic Was Not Caused by Economic Distress But by Factors That Could Be More Directly Addressed,” *Annals of the American Academy of Political and Social Science*, 2019, 681 (1), 276–291.

Dave, Dhaval, Anca M. Grecu, and Henry Saffer, “Mandatory Access Prescription Drug Monitoring Programs and Prescription Drug Abuse,” *Journal of Policy Analysis and Management*, 2021, 40 (1), 194–225.

de Chaisemartin, Clément and Xavier D'Haultfœuille, “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects,” *American Economic Review*, 2020, 110 (9), 2964–2996.

Dowell, Deborah, Tamara M. Haegerich, and Roger Chou, “CDC Guideline for Prescribing Opioids for Chronic Pain—United States, 2016,” *MMWR Recommendations and Reports*, 2016, 65 (1), 1–49.

Fink, David S., Julia P. Schleimer, Aaron Sarvet, Kyle K. Grover, Chris Delcher, Alvaro Castillo-Carniglia, June H. Kim, Ariadne Rivera-Aguirre, Samantha G. Henry, Silvia S. Martins et al., “Association Between Prescription Drug Monitoring Programs and Nonfatal and Fatal Drug Overdoses: A Systematic Review,” *Annals of Internal Medicine*, 2018, 168 (11), 783–790.

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez, and Jesse M. Shapiro, “Pre-event Trends in the Panel Event-Study Design,” *American Economic Review*, 2019, 109 (9), 3307–3338.

Gardner, John, “Two-Stage Differences in Differences,” *Working Paper*, 2022.

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

Grecu, Anca M., Dhaval Dave, and Henry Saffer, “Mandatory Access Prescription Drug Monitoring Programs and Prescription Drug Abuse,” *Journal of Policy Analysis and Management*, 2019, 38 (1), 181–209.

Harris, Matthew C., Lawrence M. Kessler, Matthew N. Murray, and Beth Glenn, “The Opioid Epidemic and the Labor Market,” *Review of Economics and Statistics*, 2020, 102 (4), 614–628.

Hollingsworth, Alex, Christopher J. Ruhm, and Kosali Simon, “Macroeconomic Conditions and Opioid Abuse,” *Journal of Health Economics*, 2017, 56, 222–233.

Horwitz, Jill, Corey S. Davis, Lynn S. McClelland, Rebecca S. Fordon, and Ellen Meara, “The Effect of State Prescription Drug Monitoring Programs on Opioid Prescribing and Related Outcomes: A Systematic Review,” *Journal of Health Economics*, 2021, 75, 102–401.

Kaestner, Robert and Engy Ziedan, “Effects of Prescription Drug Monitoring Programs on Opioid and Other Drug Use,” *Journal of Policy Analysis and Management*, 2019, 38 (4), 943–980.

— and —, “The Labor Market Effects of Reducing Opioid Access,” *Journal of Policy Analysis and Management*, 2023, 42 (1), 167–198.

Kolodny, Andrew, David T. Courtwright, Catherine S. Hwang, Peter Kreiner, John L. Eadie, Thomas W. Clark, and G. Caleb Alexander, “The Prescription Opioid and Heroin Crisis: A Public Health Approach to an Epidemic of Addiction,” *Annual Review of Public Health*, 2015, 36, 559–574.

Krueger, Alan B., “Where Have All the Workers Gone? An Inquiry into the Decline of the U.S. Labor Force Participation Rate,” *Brookings Papers on Economic Activity*, 2017, 2017 (2), 1–87.

MacKinnon, James G. and Matthew D. Webb, “Wild Bootstrap Inference for Wildly Different Cluster Sizes,” *Journal of Applied Econometrics*, 2017, 32 (2), 233–254.

Maclean, Johanna Catherine, Justine Mallatt, Christopher S. Carpenter, and Ezra Golberstein, “Economic Studies on the Opioid Crisis: Costs, Causes, and Policy Responses,” *Annual Review of Economics*, 2022, 14, 397–421.

Macy, Beth, “Dopesick: Dealers, Doctors, and the Drug Company that Addicted America,” *New York: Little, Brown*, 2018.

Meara, Ellen, Jill R. Horwitz, Wilson Powell, Lynn McClelland, Weiping Zhou, A. James O'Brien, and Nancy E. Morden, “State Legal Restrictions and Prescription-Opioid Use among Disabled Adults,” *New England Journal of Medicine*, 2016, 375 (1), 44–53.

Meinholfer, Angelica, “Prescription Drug Monitoring Programs: The Role of Asymmetric Information on Drug Availability and Abuse,” *American Journal of Health Economics*, 2018, 4 (4), 453–483.

Nam, Yunsoo T., Sri Raganathan, and Christopher Gerry, “The Effect of Prescriber-Use Mandates on PDMP Use and Opioid Prescribing,” *Health Services Research*, 2017, 52 (6), 2550–2567.

Pardo, Bryce, “Do More Robust Prescription Drug Monitoring Programs Reduce Prescription Opioid Overdose?,” *Addiction*, 2017, 112 (10), 1773–1783.

Patrick, Stephen W., Carrie E. Fry, Timothy F. Jones, and Melinda B. Buntin, “Implementation of Prescription Drug Monitoring Programs Associated with Reductions in Opioid-Related Death Rates,” *Health Affairs*, 2016, 35 (7), 1324–1332.

- Powell, David**, “The Declining Labor Market and Overdoses: The Role of Supply-Side Drug Policy,” *RAND Working Paper*, 2020.
- , **Rosalie Liccardo Pacula, and Mireille Jacobson**, “How Increasing Medical Access to Opioids Contributes to the Opioid Epidemic: Evidence from Medicare Part D,” *Journal of Health Economics*, 2020, 71, 102286.
- Rambachan, Ashesh and Jonathan Roth**, “A More Credible Approach to Parallel Trends,” *Review of Economic Studies*, 2023, 90 (5), 2555–2591.
- Roth, Jonathan, Pedro H. C. Sant’Anna, Alyssa Bilinski, and John Poe**, “What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature,” *Journal of Econometrics*, 2023, 235 (2), 2218–2244.
- Sant’Anna, Pedro H. C. and Jun Zhao**, “Doubly Robust Difference-in-Differences Estimators,” *Journal of Econometrics*, 2020, 219 (1), 101–122.
- Sun, Liyang and Sarah Abraham**, “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects,” *Journal of Econometrics*, 2021, 225 (2), 175–199.
- Wen, Hefei and Jason M. Hockenberry**, “The Effect of Medical Marijuana Laws on Opioid Prescribing and Opioid-Related Hospitalizations,” *Journal of Policy Analysis and Management*, 2017, 37 (2), 318–346.
- Wooldridge, Jeffrey M.**, “Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators,” *Econometrics Journal*, 2021.
- Xu, Yiqing**, “Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models,” *Political Analysis*, 2017, 25 (1), 57–76.

A. Data Appendix

A.1 BLS LAUS Data Construction

The Bureau of Labor Statistics Local Area Unemployment Statistics (LAUS) program produces monthly and annual labor force estimates for Census regions, states, counties, metropolitan areas, and other geographic units. I use the March monthly LAUS estimates for each state, which provide a consistent within-year snapshot of labor market conditions matching the timing of the CPS Annual Social and Economic Supplement (ASEC).

Source. Data were downloaded from the BLS LAUS homepage (<https://www.bls.gov/lau/>) using the public data API. Specifically, I used the LAUS series for each state: total civilian labor force, total employment, total unemployment, and the unemployment rate. I accessed data for all 50 states and the District of Columbia for the years 2007 through 2023.

Variable construction. From the raw LAUS data, I constructed the following variables:

- **employment:** Total employment level (number of persons) from the LAUS series. Summary statistics report this variable in thousands for readability.
- **log_employment:** Natural logarithm of total employment in persons (≈ 14 for an average state).
- **unemployment_rate:** Unemployment rate as reported by BLS (number of unemployed divided by civilian labor force, expressed as a percentage).
- **employment_rate:** Defined as $100 - \text{unemployment_rate}$, representing the employment share of the civilian labor force. This is *not* the employment-to-population ratio (E/POP).

Sample. The final estimation sample contains 850 state-year observations (50 states \times 17 years). The District of Columbia is excluded because it is a federal district with distinct economic characteristics. The estimation sample therefore consists of 46 ever-treated states and 4 never-treated states (KS, MO, NE, SD).

Timing alignment. The March LAUS estimates capture labor market conditions at a single point within each calendar year. The must-access mandate treatment variable is coded using the “full-exposure year” convention: a state is first coded as treated in the earliest calendar year for which the mandate was already in effect by January 1. This ensures that the mandate was operative before the March measurement at $e = 0$. For states where the mandate took effect mid-year, the first treatment year is the following calendar year. Because this coding convention means that the calendar year immediately preceding the full-exposure

year may contain partial mandate exposure for states with pre-March effective dates (e.g., a state with a February 2017 effective date has the mandate in effect by March 2017 but is coded as first treated in 2018), the CS estimator is run with an anticipation window of one year (`anticipation = 1`), which ensures that $e = -1$ is not used in constructing the counterfactual (see Section 4).

A.2 Must-Access PDMP Mandate Dates

Sources. The effective dates of must-access PDMP mandates were compiled from the following sources:

1. **PDAPS (Prescription Drug Abuse Policy System):** Maintained by the Temple University Beasley School of Law, PDAPS provides a comprehensive database of state controlled substance laws, including PDMP provisions. I used the “Prescription Drug Monitoring Program” dataset, which classifies state PDMP laws along multiple dimensions including whether the state requires prescribers to query the database.
2. **Horwitz et al. (2021):** This systematic review of PDMP effects includes a detailed coding of PDMP provisions across states, which I used to cross-validate the PDAPS dates for early-adopting states.
3. **Buchmueller and Carey (2018):** This paper provides must-access mandate dates for the states studied in their Medicare analysis, which I used as an additional cross-reference.
4. **National Conference of State Legislatures (NCSL):** NCSL maintains a legislative tracking database that I consulted for recent adopters and to resolve discrepancies among other sources.

Coding decisions. “Must-access” mandates are defined as laws that require prescribers (and in some cases dispensers) to query the state PDMP database before prescribing or dispensing a controlled substance. This includes mandates of varying strength: some states require checks for all controlled substance prescriptions, while others require checks only for Schedule II opioids, only for initial prescriptions, or only when certain conditions are met (e.g., clinical suspicion of abuse). I code all such mandates as treatment, recognizing that this heterogeneity in mandate strength may attenuate estimated effects.

States classified as never-treated (Kansas, Missouri, Nebraska, South Dakota) either have no universal must-access mandate or have only limited mandates applying to specific populations (e.g., Medicaid prescribers) rather than universal prescriber requirements. Some

of these states have had PDMPs with voluntary or narrowly targeted query requirements, but none had enacted a universal must-access mandate by the end of the sample period. Missouri's case is particularly notable: the state did not establish a statewide PDMP database until December 2023, making it the last state in the nation to do so.

Full-exposure year derivation. For each state, the “effective date” is the date on which the must-access mandate became legally enforceable. The “full-exposure year” is the first calendar year beginning January 1 for which the mandate was already in effect. For example, if a state’s mandate took effect on July 1, 2014, the full-exposure year is 2015, because 2015 is the first complete calendar year of mandate exposure.

A.3 Concurrent Policy Controls

For the robustness specifications that include concurrent policy controls, I compiled effective dates for four additional state-level policies:

- **Naloxone access laws:** Laws that expand access to the opioid overdose reversal agent naloxone through standing orders, pharmacy protocols, or third-party prescriptions. Dates from PDAPS.
- **Good Samaritan laws:** Laws providing legal protections (e.g., immunity from prosecution) for individuals who call for emergency assistance during an overdose. Dates from PDAPS.
- **Medicaid expansion:** Adoption of the ACA Medicaid expansion, which extended coverage to adults with incomes up to 138% of the federal poverty level. Dates from the Kaiser Family Foundation.
- **Medical marijuana legalization:** Enactment of laws permitting the use of marijuana for medical purposes. Dates from NCSL.

Each concurrent policy is coded as a binary indicator equal to one in state-year cells where the policy was in effect for the full calendar year, using the same full-exposure-year convention applied to must-access mandates.

B. Identification Appendix

B.1 Pre-Trends Analysis

The event-study analysis in [Section 5](#) provides the primary test of the parallel trends assumption. [Table 5](#) reports the numerical values of the pre-treatment event-study coefficients

for the log employment outcome using the never-treated comparison group.

Table 5: Pre-Treatment Event-Study Estimates: Log Employment

Event Time	Never-Treated Control
$e = -6$	-0.0016 (0.0024)
$e = -5$	+0.0001 (0.0021)
$e = -4$	+0.0030 (0.0020)
$e = -3$	+0.0025 (0.0019)
$e = -2$	+0.0093*** (0.0023)

Notes: Callaway-Sant'Anna dynamic ATT estimates at pre-treatment event times using never-treated states (KS, MO, NE, SD) as the comparison group. With an anticipation window of one year, $e = -1$ is excluded from the pre-treatment test (it falls within the anticipation period). All 9 cohorts (46 treated states) contribute at each pre-treatment event time $e = -6$ through $e = -2$. These estimates use the never-treated control group; see [Figure 3](#) for the preferred not-yet-treated specification, which shows clean pre-trends. Standard errors in parentheses, computed via multiplier bootstrap (1,000 iterations). The pre-treatment estimate at $e = -2$ is statistically significant, suggesting a parallel trends violation likely driven by the thin control group (4 states). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Post-Treatment Dynamic Estimates

[Table 6](#) reports the full set of post-treatment event-study estimates for the log employment outcome using the never-treated comparison group.

Table 6: Post-Treatment Event-Study Estimates: Log Employment

Event Time	ATT	SE
$e = 0$	+0.0026	(0.0019)
$e = 1$	+0.0044	(0.0033)
$e = 2$	-0.0018	(0.0066)
$e = 3$	+0.0042	(0.0079)
$e = 4$	+0.0049	(0.0091)
$e = 5$	+0.0073	(0.0108)
$e = 6$	+0.0024	(0.0109)
$e = 7$	+0.0062	(0.0167)
$e = 8$	-0.0068	(0.0223)
$e = 9$	-0.0202	(0.0228)
$e = 10$	-0.0331	(0.0198)

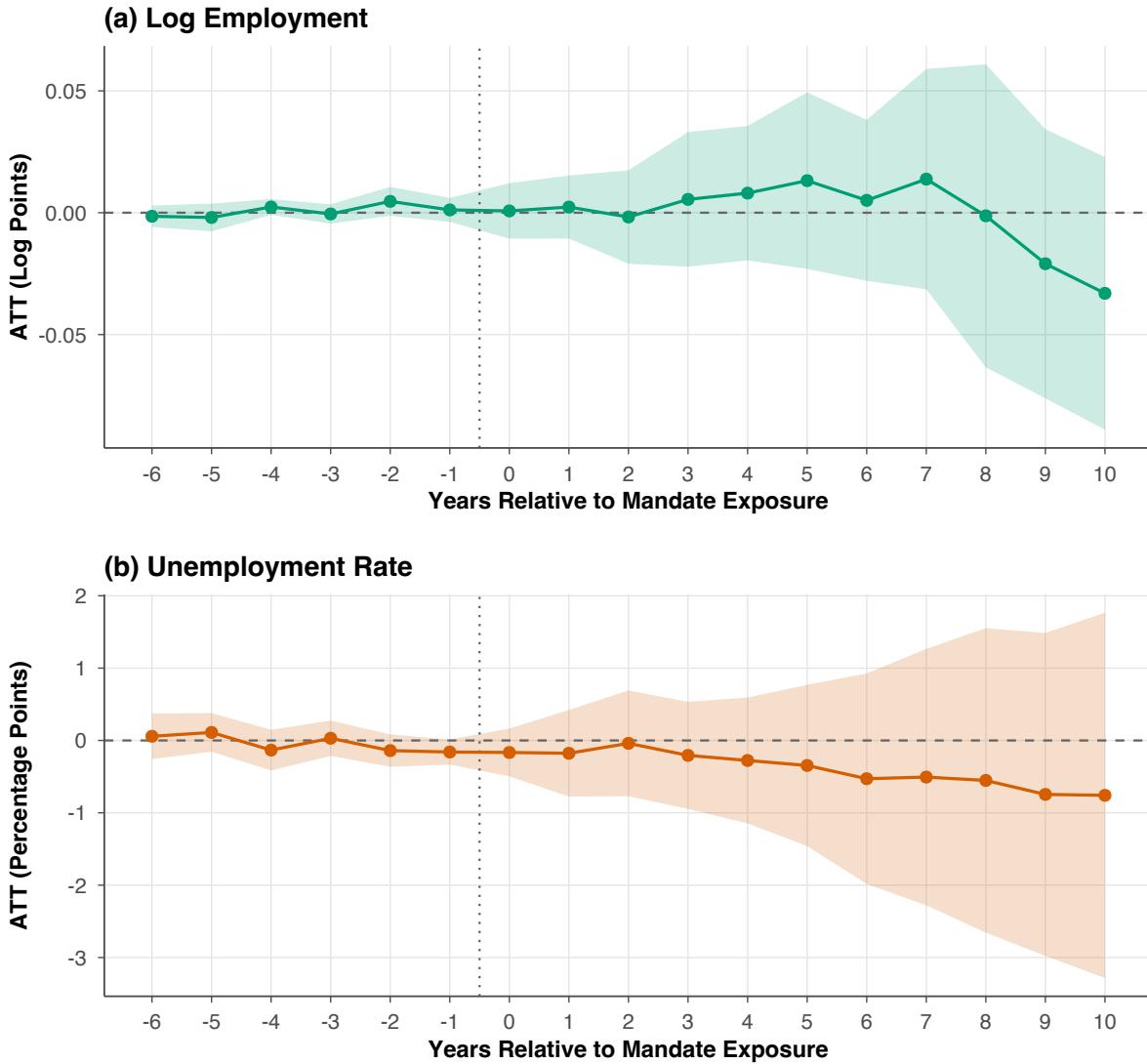
Notes: Callaway-Sant'Anna dynamic ATT estimates at post-treatment event times for log employment using never-treated states (KS, MO, NE, SD) as the comparison group. These estimates use the never-treated control group; see [Figure 3](#) for the preferred not-yet-treated specification. Standard errors in parentheses, computed via multiplier bootstrap (1,000 iterations). At longer horizons, fewer cohorts contribute to the dynamic estimate, resulting in wider standard errors. The number of contributing cohorts (states) at each horizon is: $e = 0$ through $e = +2$: 9 cohorts (46 states); $e = +3$: 8 (44); $e = +4$: 7 (42); $e = +5$: 6 (30); $e = +6$: 5 (21); $e = +7$: 4 (16); $e = +8$: 3 (10); $e = +9$: 2 (6); $e = +10$: 1 (3 states). None of the post-treatment estimates is statistically significant under 95% uniform (simultaneous) confidence bands, though the $e = +10$ estimate is marginally significant under pointwise inference ($p \approx 0.095$). 51

The post-treatment estimates show no clear trend through the available follow-up horizons. At shorter horizons ($e = 0$ through $e = 6$), point estimates fluctuate around zero (ranging from -0.002 at $e = 2$ to $+0.007$ at $e = 5$). At longer horizons ($e = 7$ through $e = 10$), the point estimates become somewhat negative (e.g., -0.033 at $e = 10$). The $e = +10$ estimate is marginally significant under pointwise inference ($p \approx 0.095$) but is based on only the 2013 cohort (WV, NM, KY) and remains insignificant under 95% uniform bands. This marginal result at the longest horizon should be interpreted with caution given the thin cohort composition and the pattern of null results at all earlier horizons. The widening standard errors and declining precision at longer horizons reflect both the reduced number of cohorts contributing and the increasing variability of cohort-specific estimates.

B.3 Not-Yet-Treated Event-Study Estimates

Dynamic Treatment Effects (Not-Yet-Treated Controls)

Callaway & Sant'Anna (2021) estimates with uniform 95% confidence bands



Note: Event time 0 = first full year of mandate exposure. Dashed line = zero effect. Dotted line = treatment onset.

Figure 6: Event-Study Estimates: Not-Yet-Treated Controls (Log Employment and Unemployment Rate)

Notes: Callaway-Sant'Anna dynamic ATT estimates with 95% uniform (simultaneous) confidence bands using not-yet-treated states as the comparison group. Panel (a) shows log employment, Panel (b) shows the unemployment rate. For log employment, all pre-treatment uniform CIs include zero, supporting the parallel trends assumption. For the unemployment rate, $e = -1$ is significant under pointwise inference ($p < 0.05$) but not under uniform bands (see Table 8 and discussion in the text), possibly reflecting partial mandate exposure during the transition year. Post-treatment estimates are insignificant under uniform bands for both outcomes. Standard errors computed via multiplier bootstrap (1,000 iterations).

[Table 7](#) reports the full set of dynamic ATT estimates for log employment using the preferred not-yet-treated comparison group, spanning event times $e = -6$ through $e = +10$. In contrast to the never-treated estimates in [Table 5](#), none of the five pre-treatment coefficients ($e = -6$ through $e = -2$) is statistically significant under 95% uniform (simultaneous) confidence bands, providing strong support for the parallel trends assumption under this specification. (Two pre-treatment coefficients— $e = -4$ and $e = -2$ —are significant under pointwise inference, but pointwise significance stars do not account for the multiple-testing problem inherent in evaluating many event-time estimates simultaneously.) Post-treatment estimates are also uniformly insignificant, with point estimates fluctuating around zero through ten years of follow-up.

[Table 8](#) reports the corresponding dynamic ATT estimates for the unemployment rate using the not-yet-treated comparison group. Most pre-treatment coefficients are insignificant, but the $e = -1$ estimate (−0.160 percentage points) is significant under pointwise inference ($p < 0.05$) though not under 95% uniform bands (uniform CI [−0.333, 0.014], which includes zero). This may reflect partial mandate exposure in the transition year rather than a genuine parallel-trends violation, as the mandate becomes effective partway through the year preceding the full-exposure year. All other pre-treatment and all post-treatment uniform CIs include zero, and the overall ATT for the unemployment rate remains insignificant ($p = 0.407$).

B.4 Goodman-Bacon Decomposition

[Table 9](#) presents the [Goodman-Bacon \(2021\)](#) decomposition of the TWFE estimator for the log employment outcome. Unlike settings with a large never-treated group, the majority of identifying weight comes from timing-based comparisons rather than treated-versus-untreated comparisons.

Table 7: Dynamic ATT Estimates: Log Employment (Not-Yet-Treated Controls)

Event Time	ATT	SE	95% Uniform CI
$e = -6$	-0.0015	(0.0017)	[-0.0059, 0.0030]
$e = -5$	-0.0020	(0.0022)	[-0.0076, 0.0037]
$e = -4$	0.0024*	(0.0012)	[-0.0009, 0.0056]
$e = -3$	-0.0005	(0.0015)	[-0.0045, 0.0034]
$e = -2$	0.0047**	(0.0023)	[-0.0013, 0.0106]
$e = -1$	0.0012	(0.0019)	[-0.0038, 0.0061]
$e = +0$	0.0007	(0.0043)	[-0.0107, 0.0121]
$e = +1$	0.0023	(0.0049)	[-0.0106, 0.0153]
$e = +2$	-0.0018	(0.0073)	[-0.0209, 0.0174]
$e = +3$	0.0055	(0.0105)	[-0.0222, 0.0331]
$e = +4$	0.0080	(0.0104)	[-0.0195, 0.0356]
$e = +5$	0.0132	(0.0137)	[-0.0231, 0.0494]
$e = +6$	0.0051	(0.0125)	[-0.0280, 0.0381]
$e = +7$	0.0138	(0.0171)	[-0.0314, 0.0590]
$e = +8$	-0.0013	(0.0235)	[-0.0635, 0.0609]
$e = +9$	-0.0209	(0.0209)	[-0.0762, 0.0343]
$e = +10$	-0.0331	(0.0212)	[-0.0890, 0.0229]

Notes: Callaway and Sant'Anna (2021) dynamic ATT estimates using not-yet-treated states as the comparison group. Doubly-robust estimation with multiplier bootstrap (1,000 iterations). Uniform confidence bands account for simultaneous inference across all event times. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (pointwise inference only). Significance stars do **not** indicate rejection under uniform (simultaneous) bands. All pre-treatment uniform CIs at $e \leq -2$ include zero, supporting parallel trends. Contributing cohorts at post-treatment horizons: $e = 0$ to $+2$: 9 cohorts (46 states); $+3$: 8 (44); $+4$: 7 (42); $+5$: 6 (30); $+6$: 5 (21); $+7$: 4 (16); $+8$: 3 (10); $+9$: 2 (6); $+10$: 1 (3).

Table 8: Dynamic ATT Estimates: Unemployment Rate (Not-Yet-Treated Controls)

Event Time	ATT	SE	95% Uniform CI
$e = -6$	0.0578	(0.1170)	[-0.2569, 0.3726]
$e = -5$	0.1109	(0.0995)	[-0.1566, 0.3784]
$e = -4$	-0.1332	(0.1047)	[-0.4147, 0.1484]
$e = -3$	0.0303	(0.0908)	[-0.2138, 0.2745]
$e = -2$	-0.1400*	(0.0834)	[-0.3644, 0.0843]
$e = -1$	-0.1598**	(0.0645)	[-0.3333, 0.0136]
$e = +0$	-0.1661	(0.1227)	[-0.4962, 0.1640]
$e = +1$	-0.1769	(0.2225)	[-0.7752, 0.4213]
$e = +2$	-0.0399	(0.2722)	[-0.7720, 0.6922]
$e = +3$	-0.2062	(0.2753)	[-0.9466, 0.5342]
$e = +4$	-0.2765	(0.3234)	[-1.1463, 0.5932]
$e = +5$	-0.3458	(0.4146)	[-1.4609, 0.7692]
$e = +6$	-0.5286	(0.5412)	[-1.9840, 0.9268]
$e = +7$	-0.5062	(0.6591)	[-2.2787, 1.2662]
$e = +8$	-0.5525	(0.7826)	[-2.6572, 1.5522]
$e = +9$	-0.7458	(0.8298)	[-2.9775, 1.4859]
$e = +10$	-0.7583	(0.9390)	[-3.2835, 1.7669]

Notes: Callaway and Sant'Anna (2021) dynamic ATT estimates using not-yet-treated states as the comparison group. Doubly-robust estimation with multiplier bootstrap (1,000 iterations). Uniform confidence bands account for simultaneous inference across all event times. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ (pointwise). Significance stars do **not** indicate rejection under uniform (simultaneous) bands. Contributing cohorts at post-treatment horizons: $e = 0$ to $+2$: 9 cohorts (46 states); $+3$: 8 (44); $+4$: 7 (42); $+5$: 6 (30); $+6$: 5 (21); $+7$: 4 (16); $+8$: 3 (10); $+9$: 2 (6); $+10$: 1 (3).

Table 9: Goodman-Bacon Decomposition of TWFE Estimate (Log Employment)

Comparison Type	Weight	Weighted Avg. 2×2 DiD
Treated vs. Never-Treated	0.2668	+0.02134
Earlier Treated vs. Later Treated	0.4468	-0.01651
Later Treated vs. Earlier Treated	0.2863	+0.01726
TWFE (weighted average)	1.0000	+0.00326

Notes: [Goodman-Bacon \(2021\)](#) decomposition of the TWFE coefficient on the must-access mandate indicator in a regression of log employment on state and year fixed effects. Weights sum to 1. Component estimates are within-group weighted averages (weighted by each 2×2 comparison’s contribution to the TWFE estimate), so the weighted sum reproduces the TWFE coefficient exactly. The treated-versus-never-treated comparison receives only 26.7% of the weight, reflecting the small never-treated group (4 states). The earlier-versus-later and later-versus-earlier comparisons carry the majority of the weight and produce estimates of opposite signs, yielding a near-zero weighted average.

C. Robustness Appendix

C.1 TWFE with Concurrent Policy Controls

[Table 10](#) presents the full results from the augmented TWFE specification that includes controls for concurrent opioid-related policies.

Table 10: TWFE Estimates with Concurrent Policy Controls

	No Controls		With Controls	
	Coef	SE	Coef	SE
Panel A: Log(Employment)				
Must-access mandate	+0.0033	(0.0055)	+0.0032	(0.0052)
State FE	✓		✓	
Year FE		✓		✓
N	850		850	

Notes: TWFE estimates with state and year fixed effects. Standard errors clustered at the state level in parentheses. “With Controls” includes binary indicators for naloxone access laws, Good Samaritan laws, Medicaid expansion, and medical marijuana legalization, all coded using the full-exposure-year convention. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

C.2 Leave-One-Out Analysis

Table 11 reports the full leave-one-out results using the preferred not-yet-treated comparison group. The distribution of ATT estimates is tightly centered around the baseline of +0.0036, confirming that no single early-adopter state drives the null result.

Table 11: Leave-One-Out Sensitivity: Not-Yet-Treated Controls

Dropped State	Cohort	ATT	SE	95% CI
KY	2013	0.0047	(0.0083)	[-0.0115, 0.0210]
NM	2013	0.0048	(0.0092)	[-0.0132, 0.0229]
NY	2014	0.0039	(0.0085)	[-0.0128, 0.0205]
TN	2014	0.0028	(0.0081)	[-0.0131, 0.0187]
VT	2014	0.0055	(0.0080)	[-0.0102, 0.0212]
WV	2013	0.0058	(0.0084)	[-0.0107, 0.0222]
None (baseline)	—	0.0036	(0.0079)	[-0.0119, 0.0191]
ATT Range		[0.0028, 0.0058]		
ATT Mean		0.0046		

Notes: Each row drops one early-adopter state (first treated ≤ 2014) and re-estimates the overall ATT using not-yet-treated controls. Callaway and Sant’Anna (2021) doubly-robust estimation.

For comparison, Table 12 reports the LOO results using the never-treated control group.

The positive estimates reflect the different baseline ATT ($\approx +0.010$) under the never-treated specification.

Table 12: Leave-One-Out Overall ATT Estimates: Never-Treated Controls (Log Employment)

Statistic	ATT
Minimum	+0.0090
Mean	+0.0107
Maximum	+0.0119

Notes: Summary statistics of the distribution of Callaway-Sant'Anna overall ATT estimates for log employment when each early-adopter state (first treated \leq 2014) is dropped in turn. This LOO analysis uses the never-treated control group (baseline ATT = +0.0100, SE = 0.0078). None of the leave-one-out estimates is significant at the 10% level.

C.3 Summary of All Specifications

[Table 13](#) consolidates the estimated ATTs across the primary specifications discussed in the main text and robustness sections.

Table 13: Summary of All ATT Estimates Across Specifications

Specification	Estimate	SE	p-value	N
Panel A: Log(Employment)				
CS-DiD, not-yet-treated (preferred)	+0.0036	(0.0079)	0.647	850
CS-DiD, never-treated	+0.0100	(0.0078)	0.203	850
TWFE, no controls	+0.0033	(0.0055)	0.554	850
TWFE, with policy controls	+0.0032	(0.0052)	0.541	850
Panel B: Unemployment Rate				
CS-DiD, not-yet-treated (preferred)	-0.2424	(0.2925)	0.407	850
CS-DiD, never-treated	-0.3989	(0.3691)	0.280	850
TWFE, no controls	-0.1693	(0.2135)	0.432	850

Notes: All specifications use BLS LAUS data for 50 states, 2007–2023 ($N = 850$). CS-DiD estimates use doubly-robust estimation with multiplier bootstrap standard errors (1,000 iterations) and an anticipation window of one year. TWFE estimates use state and year fixed effects with state-clustered standard errors. No overall aggregate estimate across these baseline specifications is statistically significant at the 10% level. A covariate-adjusted CS specification (not shown; see [Table 19](#)) yields a marginally significant positive estimate (+0.0118, $p = 0.043$), but this result is not robust across estimators (see discussion in [Section 6](#)). Individual cohort-level ATTs show some heterogeneity (see [Table 4](#)).

D. Heterogeneity Appendix

D.1 Cohort-Specific Event Studies

While the group-level ATT estimates in [Table 4](#) provide a summary of cross-cohort heterogeneity, additional insight comes from examining the event-study profiles for individual cohorts. Due to the small number of states in each early cohort (3 states for 2013, 3 for 2014), individual cohort event studies are imprecise but informative about the source of heterogeneity.

The 2013 cohort (WV, NM, KY) shows a negative drift that begins at $e = 0$ and persists through later horizons, consistent with either a cumulative mandate effect or a persistent negative shock that coincides with adoption. These states experienced significant economic challenges during 2013–2019 related to the coal industry decline (West Virginia, Kentucky) and budget difficulties (New Mexico), making it difficult to isolate the mandate effect from

confounders. The 2016 cohort (OH, VA, CT, NV, NJ, OK) shows event-study estimates that hover near zero across all post-treatment horizons, providing the clearest evidence of null effects among a cohort with sufficient states and post-treatment observations.

The 2018 and 2019 cohorts, which together comprise 21 of the 46 treated states, show relatively flat post-treatment profiles through their available post-treatment horizons, consistent with the aggregate null result. These cohorts adopted mandates during a period of strong labor market conditions nationally and entered the post-treatment period as the pandemic struck, creating countervailing forces that may mask any underlying mandate effect.

D.2 Heterogeneity by Pre-Mandate Opioid Prescribing Rate

An important dimension of potential heterogeneity is the pre-mandate opioid prescribing rate. States with higher baseline prescribing rates have more “room” for mandates to reduce prescribing and, consequently, more scope for downstream employment effects. While a formal analysis along this dimension requires state-level prescribing data, the pattern of cohort estimates provides indirect evidence.

Early-adopting states (2013–2014 cohorts) tended to have higher pre-mandate prescribing rates—West Virginia, Kentucky, and Tennessee consistently ranked among the top states in per capita opioid prescriptions. These cohorts show the most negative point estimates, potentially consistent with larger effects in high-prescribing states. However, as noted above, these cohorts also have the fewest states and the most potential for confounding, so this pattern should be interpreted cautiously. Future research linking PDMP mandates to state-level prescribing rate data could more formally test this heterogeneity dimension using triple-difference designs that exploit variation in both mandate timing and pre-mandate prescribing intensity.

D.3 Never-Treated State Characteristics

[Table 14](#) provides summary statistics for each of the four never-treated states that serve as the comparison group. The control group is geographically concentrated in the Great Plains region, which is both a limitation (potential for region-specific confounders) and a source of some internal consistency (similar economic structures may improve within-control-group comparability).

Table 14: Never-Treated States: Summary Characteristics

State	Mean Employment (000s)	Mean Unemp. Rate	Region
Kansas	1,432	4.44	Midwest
Missouri	2,880	5.45	Midwest
Nebraska	987	3.36	Midwest
South Dakota	438	3.20	Midwest

Notes: Summary statistics for never-treated states (states without universal must-access PDMP mandates by end of 2023). Mean employment and unemployment rate calculated over 2007–2023 from the same BLS LAUS analysis panel used in estimation (68 state-years). All four states are located in the Great Plains/Midwest region. Kansas and Nebraska have limited PDMP mandates applying only to Medicaid prescribers. South Dakota has no must-access mandate. Missouri did not establish a statewide PDMP database until December 2023.

E. First Stage and Sensitivity Appendix

E.1 First-Stage: Prescribing Effects

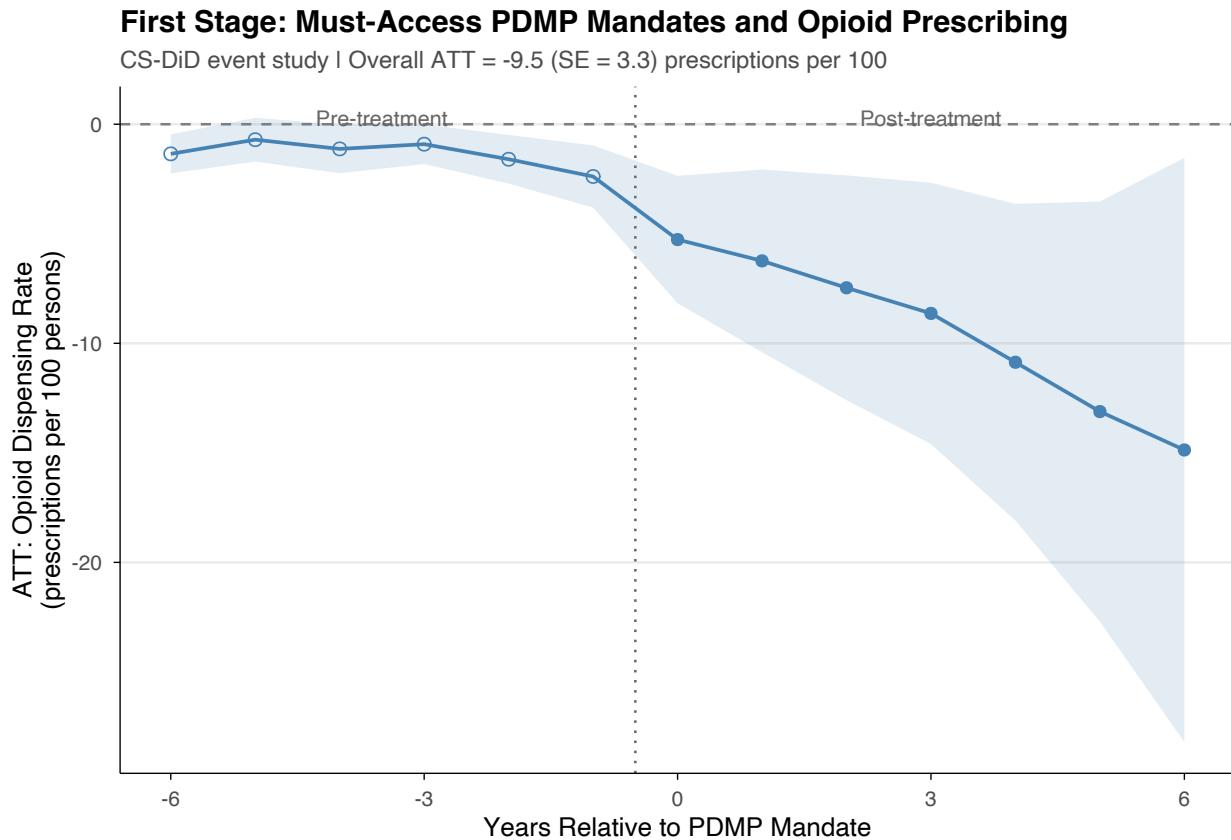


Figure 7: First Stage: Effect of Must-Access PDMP Mandates on Opioid Prescribing Rate

Notes: Callaway-Sant'Anna dynamic ATT estimates for the CDC state-level opioid prescribing rate (dispensing rate per 100 persons, 2006–2019) with 95% uniform confidence bands. Not-yet-treated comparison group; anticipation window of one year. The vertical dashed line marks $e = 0$, the first full year of mandate exposure. The prescribing rate declines sharply at $e = 0$ and remains negative at all post-treatment horizons, confirming that the must-access mandate treatment variable produces a clear first stage. Standard errors computed via multiplier bootstrap (1,000 iterations). $N = 686$ state-years (49 states \times 14 years); one state is excluded due to missing CDC prescribing data.

Table 15: First Stage: Effect of Must-Access PDMP Mandates on Opioid Prescribing

Specification	Outcome	ATT	SE	95% CI
<i>Panel A: Overall ATT</i>				
Callaway-Sant'Anna	Rx rate per 100	-9.49	(3.28)	[-15.92, -3.07]
Callaway-Sant'Anna	Log(Rx rate)	-0.1385	(0.0273)	[-0.1919, -0.0851]
TWFE	Rx rate per 100	-5.20	(1.59)	[-8.31, -2.09]
TWFE	Log(Rx rate)	-0.0722	(0.0214)	[-0.1141, -0.0302]
<i>Panel B: Group-level ATT (CS-DiD, Rx rate per 100)</i>				
Cohort 2013	Rx rate per 100	-30.68***	(8.86)	[-48.05, -13.32]
Cohort 2014	Rx rate per 100	-3.92	(11.18)	[-25.84, 17.99]
Cohort 2015	Rx rate per 100	-9.00**	(4.30)	[-17.42, -0.57]
Cohort 2016	Rx rate per 100	-7.64**	(3.67)	[-14.84, -0.43]
Cohort 2017	Rx rate per 100	-9.05**	(3.63)	[-16.16, -1.94]
Cohort 2018	Rx rate per 100	-6.87**	(2.76)	[-12.27, -1.46]
Cohort 2019	Rx rate per 100	-5.27	(3.34)	[-11.82, 1.28]

Notes: This table reports first-stage estimates of the effect of must-access PDMP mandates on opioid prescribing rates (dispensing rate per 100 persons). Panel A shows overall ATT estimates from both Callaway-Sant'Anna (2021) and two-way fixed effects specifications. Panel B reports cohort-specific ATT estimates from the CS-DiD estimator. The comparison group consists of not-yet-treated states, consistent with the main employment analysis. Standard errors are computed via multiplier bootstrap (1,000 iterations) with one year of anticipation. Opioid prescribing data from CDC state-level reports (CDC/IQVIA Xponent), restricted to years with observed state-level data (2006–2019). $N = 686$ state-years ($49 \text{ states} \times 14 \text{ years}$); 45 treated states, 4 never-treated (KS, MO, NE, SD). One state is excluded due to missing CDC prescribing data. The 2020 and 2021 adoption cohorts (TX, MT, ID, WY) are omitted from Panel B because the prescribing data end in 2019, providing zero post-treatment observations for these cohorts; these states contribute as not-yet-treated controls for earlier cohorts.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 16: First Stage Calibration: Literature Estimates of PDMP Mandate Effects on Prescribing

Study	Outcome	Effect	Design
Buchmueller & Carey (2018)	Schedule II Rx	-10%	TWFE
Dave et al. (2021)	Initial opioid Rx	-9.5%	CS-DiD
Patrick et al. (2016)	Opioid Rx volume	-12%	Pre-Post
Brady et al. (2014)	Schedule II Rx rate	-11%	Pre-Post

Notes: Published estimates of the effect of must-access PDMP mandates on opioid prescribing. Effects represent approximate percentage reductions. These estimates calibrate the expected first-stage effect for interpreting the reduced-form labor market results.

E.2 HonestDiD Sensitivity

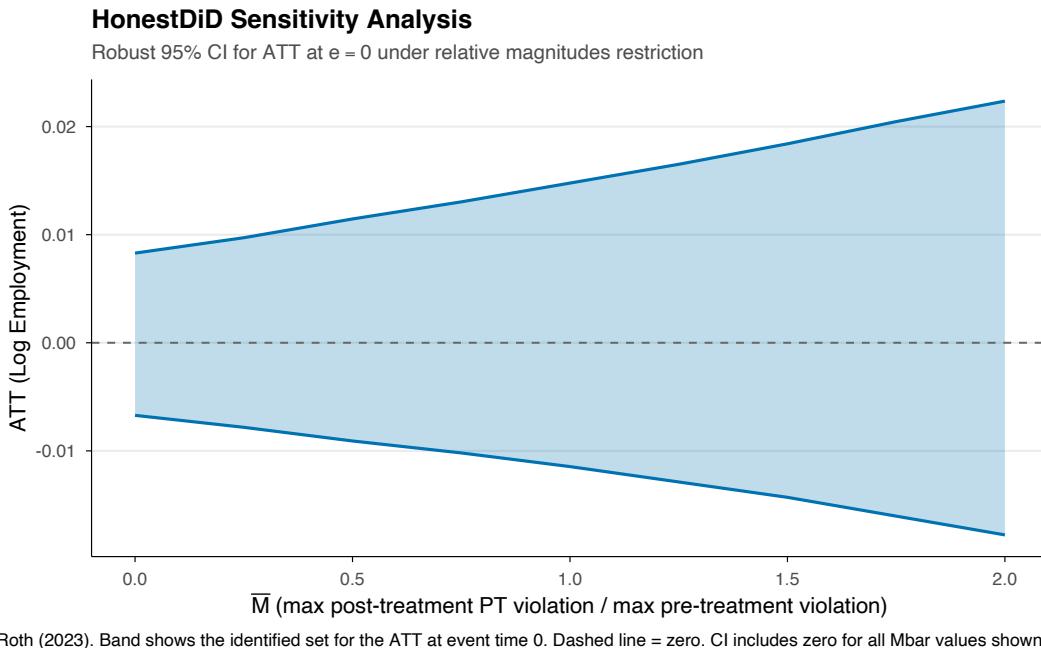


Figure 8: Sensitivity of Log Employment ATT at $e = 0$ to Parallel Trends Violations (HonestDiD)

Notes: [Rambachan and Roth \(2023\)](#) sensitivity analysis for the ATT at event time $e = 0$ on log employment under the not-yet-treated specification. The horizontal axis represents \bar{M} , the maximum change in the slope of the counterfactual trend between consecutive periods. At $\bar{M} = 0$, the standard parallel trends assumption holds exactly. The robust 95% confidence interval includes zero for all values of \bar{M} shown, indicating that the null result is robust to plausible deviations from parallel trends.

Table 17: Minimum Detectable Effects (MDEs) for Main Specifications

Outcome	SE	N	Clusters	MDE	MDE
				($\alpha = 0.05, \beta = 0.80$)	($\alpha = 0.10, \beta = 0.80$)
Log(Employment)	0.0079	850	50	0.0221 (2.24%)	0.0196
Unemployment Rate	0.2925	850	50	0.819 pp	0.727 pp
<i>Literature benchmarks for comparison:</i>					
Harris et al. (2020)				0.008–0.015	(log emp)
Krueger (2017)				0.03–0.06	(LFPR, pp)
Aliprantis et al. (2020)				0.01–0.04	(emp rate, pp)

Notes: MDE = $SE \times (z_{\alpha/2} + z_{\beta})$. For $\alpha = 0.05$ and power = 0.80: MDE = $SE \times 2.802$. Standard errors from Callaway and Sant'Anna (2021) estimation with not-yet-treated controls and multiplier bootstrap (1,000 iterations). Sample: 850 state-years, 50 states, 2007–2023. Literature benchmarks show plausible effect sizes from related studies of opioid policy and labor market outcomes.

Table 18: Robustness to Excluding COVID-Era Observations (2020–2023)

Outcome	Full Sample (2007–2023)		Pre-2020 (2007–2019)	
	ATT	SE	ATT	SE
Log(Employment)	0.0036 [-0.0119, 0.0191]	(0.0079) [0.0119, -0.0119]	-0.0027 [-0.0212, 0.0159]	(0.0095) [0.0159, -0.0212]
Unemployment Rate	-0.2424 [-0.8157, 0.3309]	(0.2925) [0.3309, -0.8157]	-0.3909 [-0.9111, 0.1292]	(0.2654) [0.1292, -0.9111]
Estimator	CS-DiD (DR)		CS-DiD (DR)	
Control group	Not-yet-treated		Not-yet-treated	
Anticipation	1 year		1 year	
N (state-years)	850		650	
States	50		50	
Years	2007–2023		2007–2019	

Notes: Callaway and Sant'Anna (2021) doubly-robust ATT estimates using not-yet-treated states as the comparison group with 1-year anticipation. Multiplier bootstrap standard errors (1,000 iterations) in parentheses; 95% pointwise confidence intervals in brackets. The pre-2020 sample excludes observations from 2020–2023 to assess robustness to potential confounding from the COVID-19 pandemic and associated labor market disruptions. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

E.3 Minimum Detectable Effects

E.4 Pre-COVID Subsample (2007–2019)

E.5 BJS Imputation Event Study

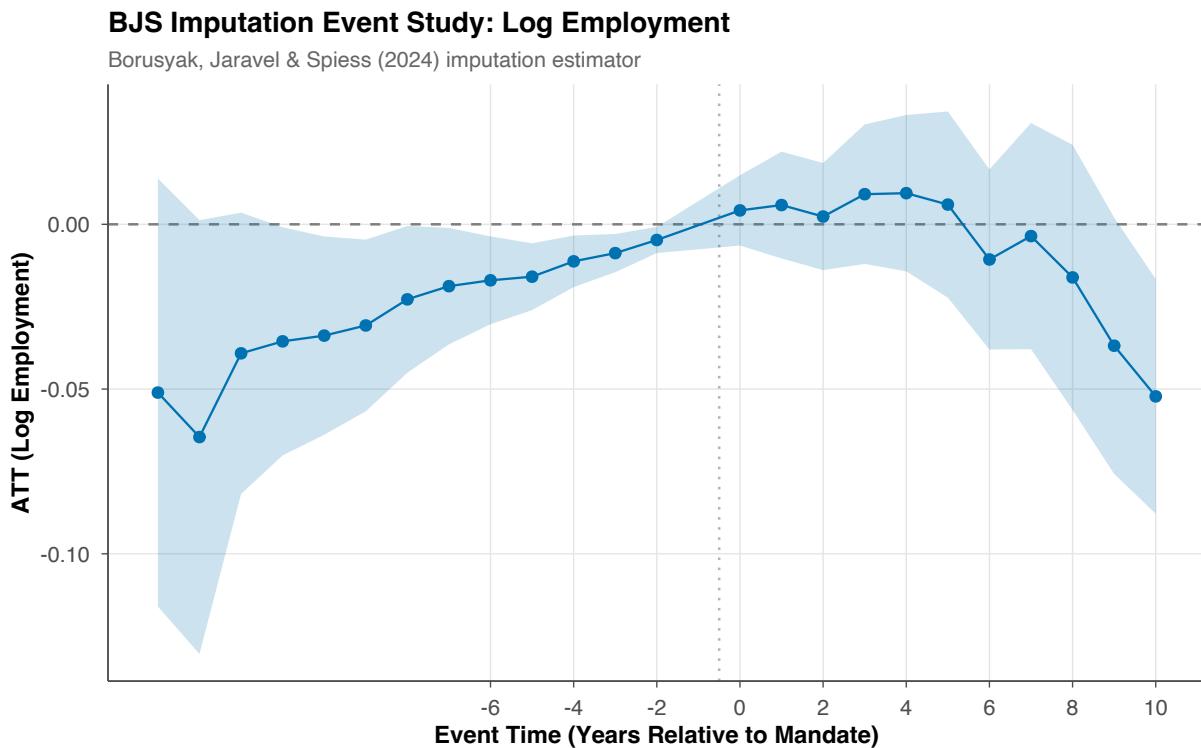


Figure 9: BJS Imputation Event Study: Log Employment

Notes: Borusyak, Jaravel, and Spiess (2024) imputation estimator with 95% pointwise confidence intervals. The estimator constructs counterfactual outcomes for treated units by imputing from untreated observations using a two-way fixed effects model. Pre-treatment estimates trend upward from approximately -0.05 at $e = -8$ toward zero near $e = -1$; this pattern is consistent with the thin-control-group problem (treated states trending differently from never-treated Great Plains states) and does not indicate a violation of parallel trends in the CS not-yet-treated specification. All pre-treatment confidence intervals include zero. Post-treatment estimates are near zero through $e = +5$, consistent with the aggregate null result.

Table 19: Robustness Across Estimators and Specifications: Log Employment

Specification	ATT	SE	95% CI	N
CS-DiD (NYT, no covariates)	0.0036	(0.0079)	[-0.0119, 0.0191]	850
CS-DiD (NYT, with covariates)	0.0118	(0.0058)	[0.0004, 0.0232]	850
BJS Imputation	-0.0075	(0.0144)	[-0.0357, 0.0208]	850
CS-DiD (NYT, 2007–2019, drop late cohorts)	0.0069	(0.0064)	[-0.0058, 0.0195]	598
TWFE	0.0033	(0.0055)	[-0.0075, 0.0140]	850

Notes: All specifications use log employment as the dependent variable. CS-DiD = Callaway and Sant’Anna (2021); NYT = not-yet-treated controls; BJS = Borusyak, Jaravel, and Spiess (2024) imputation estimator; TWFE = two-way fixed effects with state-clustered standard errors. Covariates (row 2) include 2007 baseline log employment and unemployment rate. Pre-COVID sample (row 4) restricts to 2007–2019 and drops cohorts first treated after 2019. 95% CIs are pointwise.

E.6 Robustness Across Estimators

F. Additional Figures and Tables

F.1 Placebo Test Results

[Table 20](#) presents the results of the placebo test in which treatment dates are shifted backward by three years.

Table 20: Placebo Test: Treatment Date Shifted Back Three Years (TWFE)

Outcome	Placebo ATT	SE	N
Log(Employment)	+0.0048	(0.0040)	540

Notes: TWFE estimate with treatment dates shifted backward by three years. The sample is restricted to pre-actual-treatment observations ($N = 540$ state-years). State and year fixed effects included; standard errors clustered at the state level. The estimate is small and statistically insignificant ($p = 0.235$), confirming that the TWFE specification does not produce spurious treatment effects at pre-mandate dates.

F.2 Summary of Results Across Outcomes and Specifications

The comprehensive analysis presented in this paper examines two primary employment outcomes—log employment and the unemployment rate—across multiple specifications (CS

never-treated, CS not-yet-treated, TWFE baseline, TWFE with controls). No baseline aggregate estimate is statistically significant at the 10 percent level. The overall ATT estimates for log employment range from +0.004 (not-yet-treated) to +0.010 (never-treated) across specifications; for the unemployment rate, from -0.40 (never-treated) to -0.10 (TWFE) percentage points. One covariate-adjusted CS specification (not-yet-treated, with baseline covariates) produces a marginally significant positive estimate (+0.0118, $p = 0.043$), but this result is not robust across estimators—the BJS imputation estimator yields an estimate of opposite sign (-0.0075)—and lies within the confidence intervals of the unconditional specifications (see [Table 19](#) and discussion in [Section 6](#)). While individual cohort-level ATTs show some significant heterogeneity (see [Table 4](#)), the overall pattern across all aggregate specifications is consistent with negligible effects. This consistent pattern of null aggregate results provides strong evidence for the conclusion that must-access PDMP mandates do not produce detectable aggregate employment effects at the state level over the medium run.