NBER WORKING PAPER SERIES

DID DRUG DECRIMINALIZATION INCREASE OVERDOSES

David Hall
Benjamin Hansen
Kyutaro Matsuzawa

Working Paper 35427
http://www.nber.org/papers/w35427

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2026

Supported by Grant 24-12072 from Arnold Ventures. Thank you to Adam Soliman, Jonathan Zhang, Emily Leslie, Jonathan Davis, Edward Rubin, Michael Kuhn, Lars Lefgren, and Jeffrey Denning for thoughtful advice and comments on earlier versions of this paper. We are grateful for the feedback from seminar participants at SDSU Brown Bag, Penn State, and UNLV. We also appreciate comments and thoughts from conference participants at APPAM, ASHE, SOLE, WEAI, and the 2025 IESR Labor Conference. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

Did Drug Decriminalization Increase Overdoses
David Hall, Benjamin Hansen, and Kyutaro Matsuzawa
NBER Working Paper No. 35427
July 2026
JEL No. I12, I18, K40, K42, K49

ABSTRACT

We study the effect of drug decriminalization on overdose mortality using Oregon's Measure 110 and Washington's Blake decision. Synthetic-control estimates show that both decriminalization regimes sharply reduced drug arrests and were followed by sustained increases in overdose mortality relative to matched counterfactuals. The estimates imply approximately 1,186 excess deaths in Oregon and 1,895 in Washington from 2021 through 2023. We also revisit the role of fentanyl supply shocks. A fentanyl-share control may partly capture enforcement-driven declines in non-fentanyl NFLIS reports, while alternative fentanyl measures leave positive overdose effects.

David Hall
Univeristy of Oregon
dhall7@uoregon.edu

Kyutaro Matsuzawa
San Diego State University
kmatsuzawa2@sdsu.edu

Benjamin Hansen
University of Oregon
Department of Economics
and NBER
bchansen@uoregon.edu

1 Introduction

The overdose epidemic has been one of the most devastating public health crises in recent history. The United States saw more than 100,000 overdose deaths in each year from 2021-2024. Unintentional overdose mortality is the leading cause of external mortality (Center for Disease Control and Prevention, National Center for Health Statistics, 2022) surpassing transportation accidents ($\sim 47,000$) and firearm-related deaths ($\sim 46,000$).

While the social costs of drug addiction are vast, the criminalization of drug use imposes burdens, including reduced employment opportunities for individuals with criminal records, lower lifetime earnings, and deterrence from seeking emergency care during overdose events (Agan and Starr, 2018, Holzer et al., 2006, Byles et al., 2024). Decriminalization may mitigate these harms, but the net benefit depends on whether decriminalization can reduce (or at least not increase) drug abuse. If decriminalization increases access and use of treatment and safer use of illicit substances, it could have benefits through reducing overdoses and the consequences of contact with the criminal justice system. However, it also could increase use and abuse by reducing stigma and perhaps the costs associated with illicit drug consumption. In this paper, we estimate the impact of decriminalizing drug possession on a crucial outcome in the ongoing debate over decriminalization: drug-related overdose mortality.

We study the impact of decriminalization on the fatal overdose rate in the context of Oregon's Measure 110 and Washington's Blake Fix. Enacted in February 2021, Measure 110 was a voter ballot initiative that became a harm reduction policy primarily intended to reduce racial inequities within the incarceration system via decriminalizing the possession of small amounts of drugs (Anthony Johnson et al., 2019). In particular, Measure 110 reduced the penalty for small amounts of drug possession from a Class A misdemeanor to a Class E violation carrying a fine of $100. Similarly, Washington also experimented with a similar drug decriminalization law beginning in March 2021, motivated by a court decision.

Our main empirical specification uses the synthetic control method. Synthetic control is a reasonable approach in this context because of its ability to create a counterfactual in case studies with a single treated unit, and also because of the design's transparency (Abadie et al., 2010). Joshi et al. (2023) and Spencer (2023) both provide early estimates of the causal impact via synthetic control models. However, despite

using seemingly identical approaches and asking the same question, they vary in their conclusions, with one suggesting a statistically insignificant effect and the other finding a significant increase. Zoorob et al. (2024) use a matrix-completion approach with controls for fentanyl's arrival to estimate the causal impact of Measure 110 on longer-term overdose-related mortality through 2022 and find little evidence of long-run effects. In our paper, we expand on these studies to estimate the longer-run effects of Measure 110 by extending the analysis through 2023 and revisiting the central question of whether we should directly control for fentanyl's geographic spread in our modeling choices.

We begin by extending the methodologies of Joshi et al. (2023) and document that the implementation of Measure 110 led to an increase of approximately 0.806 per 100,000 overdose-related deaths per month. This corresponds to approximately 1,186 additional deaths from 2021 to 2023—an increase of 38 percent relative to the baseline mean. We also consider whether the primary driver of the observed mortality increase was a rise in fentanyl supply. We interpret these estimates as reduced-form effects of the decriminalization policy environment, including the legal change, enforcement responses, and implementation choices, rather than as direct evidence on a single behavioral mechanism. The analysis therefore speaks to the net mortality consequences of the policy regime, while mechanism-specific channels require additional evidence.

We conduct several robustness tests to ensure that the estimated effects are not driven by either unobservable factors or design choices in synthetic control estimation. We show results from various extensions of synthetic control, including augmented synthetic control (Ben-Michael et al., 2021), generalized synthetic control (Xu, 2017), and synthetic difference-in-differences (Arkhangelsky et al., 2021). Results from these alternative specifications are in line with our main results. We also show results from an exercise where we switch the weights of Spencer (2023) and Joshi et al. (2023) while using their respective definitions. While this results in poor pre-treatment matches, the conclusion of an increase in overdose-related mortality remains. To test whether particular donor units are driving our results, we perform a remove-one robustness check.

Our study also builds on Zoorob et al. (2024), whose methodology differs from Spencer (2023) and Joshi et al. (2023) principally by explicitly attempting to control for fentanyl's market saturation through a proxy: the fraction of NFLIS drug reports involving fentanyl. They find there is no increase in overdoses after

controlling for fentanyl's arrival. We replicate and explore the sensitivity of this approach to a variety of measures of fentanyl saturation. While we can replicate their key findings, we also find their measure is driven by the *drop* in non-fentanyl NFLIS reports, rather than the increase in fentanyl reports. We also explore other potential proxies, finding evidence that other reasonable measures of fentanyl supply suggest region-specific supply shocks do not mitigate the estimated increases in overdoses due to drug decriminalization.

We also consider additional controls and strategies to deal with region-specific trends related to fentanyl supply, which may be a confounder with drug decriminalization. For instance, during COVID-19, there may have been region-specific differences in enforcement creating potential structural breaks in the donor pool: states that served as good counterfactuals for Oregon before 2020 may no longer be valid comparisons post-pandemic. To address this, we implement a sensitivity check by restricting the synthetic control match to data through February 2020, allowing us to isolate pre-pandemic trends and test whether the estimated effects are confounded by state-specific COVID-19 impacts on substance use. To assess the possibility of regional fentanyl shocks, we construct placebo synthetic control estimates for California, Nevada, and Idaho. We evaluate the robustness of our findings across various definitions of overdose mortality. This includes opioid-related, non-opioid, and non-synthetic deaths to assess whether Measure 110 disproportionately influenced synthetic opioid outcomes or affected drug use more broadly.

This paper contributes to two areas of research. First, it adds to the literature on opioid abatement strategies. Most states have focused on doctor shopping restrictions, targeting pill mills (Soliman, 2025), and prescription drug monitoring programs (Maclean et al., 2020). In contrast, decriminalization is intended to be a part of a broader harm reduction framework, alongside interventions like syringe exchange programs (SEPs) and expanded access to naloxone. Harm reduction programs can have benefits but also unintended consequences. For example, Packham (2022) finds that while SEPs effectively reduce the spread of HIV and HCV, they may unintentionally increase overdose mortality. Compared to these strategies, decriminalization has received relatively little causal analysis, particularly outside the context of Oregon's Measure 110, due to limited quasi-experimental opportunities. The primary evidence to date comes from Portugal's 2001 decriminalization, where Félix et al. (2017) estimate a significant reduction in drug-related deaths following the policy change. Beyond its policy relevance, empirical analysis of decriminalization also tests classical

economic theory. Our results speak to the Economic Theory of Illegal Goods proposed by Becker et al. (2004). Becker et al. (2004) argue that public expenditure on apprehension of suppliers of illegal goods, such as narcotics traffickers, depends on the social value of consuming that illegal good. Under plausible estimates of demand elasticity, they show that criminal penalties for drug use may be welfare-reducing. We contribute to this theoretical framework by examining one critical welfare-relevant outcome of decriminalization: overdose mortality.

We also contribute to ongoing efforts to replicate and verify published results in scientific research (Hillary and Medaglia, 2020, Camerer et al., 2016, Chang and Li, 2015). Importantly, we show that while Spencer (2023) and Joshi et al. (2023) arrived at different conclusions based on one year of follow-up, extending the post-period through 2023 yields consistently positive overdose gaps across their main outcome and timing choices. In Appendix A we fully document the researcher choices that led to different conclusions, replicate prior work estimating the overdose mortality effects of Measure 110, synthesize the methodological differences, and provide our estimates transparently. In using the synthetic control method, we ensure reproducibility by making the donor weights used to construct our counterfactual explicit and matching on outcome only, following Ferman et al. (2020). We document this approach in both the paper and our replication package. In doing so, we underscore how synthetic control models can help align empirical research with the broader goals of transparency and replicability in applied economics, but also highlight the fragility of studies with limited time horizons and results that are only marginally statistically significant.

The remainder of the paper is structured as follows. Section 2 provides context on Measure 110 and Washington's "Blake Fix" decision. Section 2.3 reviews prior literature estimating the impact of Measure 110 on overdose mortality, as well as other evaluations of the policy's broader effects. Section 3 introduces our data and methods. Section 4 contains our main results. Section 5 concludes.[1]

[1] Appendix A contains a full replication and reconciliation exercise of Spencer (2023) and Joshi et al. (2023), and additional results.

2 Background

2.1 Oregon's Measure 110

Oregon voters passed Ballot Measure 110 on November 3, 2020.[2] The measure received 58.46 percent of the vote, exceeding the majority threshold by 8.46 percentage points, and was enacted on February 1, 2021. The primary component of the policy was the decriminalization of possession of specified controlled substances at personal-use amounts. The bill adjusts drug possession violations depending on the categorized amount of drugs an individual is found in possession of. Individuals found with "user amounts" of drugs are subject to a Class E violation as opposed to a Class A misdemeanor (Anthony Johnson et al., 2019).[3] This particular Class E violation entails a $100 fine which can be waived if the individual calls a helpline that provides information on treatment services. Individuals possessing more than the "user" threshold are subject to a Class A misdemeanor rather than a Class B or C felony.[4]

The bill's goal was to shift legal attitudes surrounding addiction (Oregon Health Justice Recovery Alliance, 2021). Supporters of the bill, including the Drug Policy Alliance and Chan Zuckerberg Initiative advocated that substance use disorders (SUDs) should be treated as a disease rather than as a crime, thus an individual with a SUD should receive treatment rather than incarceration. This would alleviate concerns about capacity constraints in jails. Black people who use drugs (PWUD) are also more likely to be incarcerated for drug possession, and Measure 110 would reduce racial inequities within the Oregon incarceration system.

Critics of the bill included law enforcement and legislative opposition (Vote No on Measure 110, 2020). Oregon was already progressive regarding drug decriminalization; those found in possession of drugs were already given opportunities to seek treatment through the legal system via drug courts, for instance. Another concern was that the increase in treatment funding was overstated. Treatment providers were already at capacity and providing more beds required immediate funding that was not provisioned in the bill.[5]

Due to the limited time frame in which the full decriminalization policy was active, little published

[2] Measure 110 is also known as the Drug Addiction and Recovery Act.

[3] A Class A misdemeanor is punishable by up to 364 days of imprisonment and up to $6,250 in fines. `https://oregon.public.law/statutes/ors_161.635`

[4] Consequences of a Class B (C) felony are up to 10 (5) years imprisonment and a $250,000 ($125,000) fine.

[5] The bill does include provisions for funding treatment services, but the process of funding health providers would not begin until more than a year after Measure 110's enactment.

research has been conducted.[6] Davis et al. (2023) find a statistically significant decrease in drug possession arrests following the enactment of Measure 110. They also find a significant increase in "displaced" arrests, providing evidence of a substitution pattern among law enforcement agents. Smiley-McDonald et al. (2023) find evidence of asymmetric information among law enforcement agents and the public regarding the entirety of Measure 110 provisions. Law enforcement officers (LEOs) were unaware of treatment funding provided by Measure 110. LEOs were also hesitant to interact with PWUD after M110's enactment due to perceptions that the policy lacks teeth (Henderson et al., 2023). The decreased interaction with law enforcement agents may be the reason for the limited use of the statewide referral hotline (Russoniello et al., 2023). Those found in possession of user amounts of drugs were given a $100 fine or the option to call the statewide telephone hotline to have the fine removed and be referred to treatment services (Russoniello et al., 2023). By the end of 2023, only 577 people had called into the hotline (Gaitán, 2023).

After just three years since the enactment of the bill, Governor Tina Kotek signed House Bill 4002, repealing much of the progress made in the way of decriminalization (Representative Jason Kropf and Senator Kate Lieber, 2024). Advocates of recriminalization cite increasing overdose rates in the state of Oregon; the state itself declared an emergency in June of 2024 to address the fentanyl crisis in Portland (A. B. C. News, 2024). On September 1, 2024, the repeal became effective, making the possession of drugs a misdemeanor offense.

2.2 Washington's "Blake Fix"

On February 25, 2021, 24 days after Oregon's decriminalization went into effect, the Washington Supreme Court found the state's possession of a controlled substance statute unconstitutional.[7] In the case of *State v. Blake*, the court ruled that "convictions for Possession of Controlled Substance aka Violation of Uniform Controlled Substances Act (VUCSA), were not lawful," and made it possible for those convicted under the invalid statute to vacate those convictions (State of Washington, 2021). This ruling decriminalized the possession of controlled substances until SB5476 ("Blake Fix") took effect. Washington lawmakers enacted

[6] We acknowledge there is active and ongoing as-yet-unpublished research, particularly that of researchers who attended Comagine Health's Measure 110 Research Symposium in January of 2024.

[7] Washington's controlled substance statute did not require prosecution to prove that the defendant was aware they possessed such a substance.

SB5476 on May 13, 2021, with an effective date of July 25, 2021; the law recriminalized possession but lowered sentencing guidelines for simple possession (Manka, 2021).[8]

2.3 Literature Review

As mentioned before, Spencer (2023) and Joshi et al. (2023) study the effect of drug decriminalization on overdoses. These two prior studies reached different conclusions about the one-year impacts of decriminalization. We show in Appendix A and Appendix Tables 1 and 2 that these differences are primarily explained by small differences in outcome definitions and sample window. However, as shown in Appendix Table 3, the estimated ATTs are not statistically different from one another, suggesting that limited post-period power may also explain these differences. This motivates our central question: what are the longer-term impacts of drug decriminalization on overdoses in Oregon and Washington?

Zoorob et al. (2024) extend the time period analyzed in Spencer (2023) and Joshi et al. (2023), focusing on the second-year update of Measure 110, although their methodology differs slightly from these earlier studies. They aggregate their panel data into half-year periods rather than monthly, and they use a longer pre-treatment window starting in January 2008. In addition, they adopt a matrix completion approach to synthetic control. Initially, they find a statistically significant half-year increase in drug mortality of 1.83 per 100,000 (equivalent to a monthly increase of 0.3 per 100,000). However, in their preferred specification—which controls for fentanyl supply as measured by the percentage of all NFLIS drug reports involving fentanyl—they find no association between Measure 110 and drug mortality. Using a structural break test, Zoorob et al. (2024) find evidence that a fentanyl regime shift arrived in Oregon at the same time as Measure 110, which they interpret as evidence for why Measure 110 is not associated with an increase in drug mortality.

Our interpretation of fentanyl measures depends on their timing. Pre-treatment fentanyl exposure can improve the comparability of Oregon and Washington to donor states, and we therefore view pre-treatment fentanyl measures as candidate predictors or donor-pool restrictions. By contrast, post-treatment fentanyl report-based measures are potentially affected by the policy through changes in enforcement, submissions, and testing. We therefore do not treat post-treatment report shares as clean confounder controls in our

[8]Simple possession is similar to personal use as defined by the Oregon law. SB5476 was effective until mid-2023. In May of 2023, WA Governor Jay Inslee signed SB5536, a bill that increased sentencing on simple possession.

preferred specification. Instead, we report specifications using alternative fentanyl measures as sensitivity analyses that ask whether the estimated overdose gap is robust to different ways of proxying for regional fentanyl exposure.

We raise two methodological concerns with the approach taken by Zoorob et al. (2024). In their preferred specification, the authors explicitly control for changes in the supply of illicitly manufactured fentanyl. If the saturation of fentanyl in the drug market coincided with the passage of Measure 110, synthetic control estimates could be biased upward unless the donor pool units used to construct the synthetic Oregon also experienced similar supply-side shocks. To address this, the authors use the ratio of fentanyl-related reports to the total number of drug reports in Oregon to control for fentanyl availability.

One concern relates to the assumption underlying the inclusion of this control; the other pertains to the specific measurement strategy used. First, the authors implicitly assume that Measure 110 had no effect on the fentanyl supply chain itself. We find this assumption limiting. Consider a rational agent engaged in illicit fentanyl trafficking, as modeled in Becker's (1968) economic theory of crime (Becker, 1968). If drug traffickers and retailer sellers weigh the costs of apprehension and conviction, decriminalization may reduce their perceived risk and lead to increased offending. In other words, if drug-related arrests fell post-Measure 110 then the assumption that fentanyl supply remained unaffected by the policy change is questionable. This raises concerns about controlling for fentanyl supply if decriminalization also changes fentanyl supply.

Second, their preferred control, the ratio of fentanyl reports to all drug reports, may be a poor proxy for fentanyl supply. Aside from documented concerns about the accuracy of NFLIS data compared to actual illicit supply (Pitts et al., 2023), this ratio can vary either because of changes in fentanyl prevalence or because of reporting patterns for drugs. If this ratio changes due to changes in the reporting patterns of other drugs, which is directly induced by decriminalization, then this ratio may be a bad control (Caetano et al., 2024). If the goal is to capture a structural shift in fentanyl availability, then more direct measures, less affected by the broader seizure landscape, would be preferable.

A separate empirical concern involves the decision to use half-year data and to extend the pre-treatment window back to 2008. The authors aim to address potential temporal confounding between Measure 110 and the transition from the heroin wave to the fentanyl wave of the opioid epidemic. However, by extending

the pre-period to 2008, they allow the synthetic control to span across multiple regime shifts, most notably the period when overdose deaths shifted from being driven by prescription opioids to heroin following the reformulation of OxyContin (Evans et al., 2018, Ciccarone, 2019). Prior research suggests that these transitions occurred idiosyncratically across states (Ciccarone, 2019, Shover et al., 2020, West et al., 2021), raising concerns about the plausibility of achieving valid synthetic matches, not just for the fentanyl era but for each of the preceding waves as well.

3 Data and Methods

3.1 Data

For our primary analyses, we use restricted-use records from the National Center for Health Statistics Multiple Cause of Death files. Our primary outcome is monthly state-level drug overdose mortality, defined using ICD-10 underlying cause-of-death codes X40–X44, X60–X64, X85, and Y10–Y14.[9] This all-overdose definition captures fatal drug poisonings regardless of intent and reduces sensitivity to differences in how medical examiners and coroners classify intent. Spencer (2023) instead focuses on unintentional drug overdose mortality, which corresponds to codes X40–X44. Appendix Figures 1 and 2 compare the all-overdose and unintentional-overdose definitions nationally and for Oregon and Washington, respectively.

We use drug report data from the National Forensic Laboratory Information System (NFLIS-Drug) to measure changes in the composition of drugs submitted to and analyzed by forensic laboratories. NFLIS-Drug collects information from federal, state, and local forensic laboratories on drug items seized by law enforcement, submitted for laboratory analysis, analyzed, and reported to NFLIS (Pitts et al., 2023). We aggregate reports to the state-by-half-year level and construct four measures of fentanyl market saturation: fentanyl and fentanyl-related reports per 100,000 population, non-fentanyl reports per 100,000 population, fentanyl and fentanyl-related reports as a share of all NFLIS drug reports, and fentanyl and fentanyl-related reports as a share of opioid-related reports.[10]

[9] Following standard CDC/NCHS definitions, we classify drug overdose deaths using underlying cause-of-death codes X40–X44 for unintentional drug poisoning, X60–X64 for intentional self-harm by drug poisoning, X85 for assault by drug poisoning, and Y10–Y14 for drug poisoning of undetermined intent. Multiple-cause-of-death codes T36–T50 identify poisoning by drugs, medicaments, and biological substances and allow us to characterize the drugs involved in overdose deaths. For example, opioid-involved deaths can be identified using T40 codes among deaths with an overdose underlying cause.

[10] We use fentanyl as shorthand for fentanyl and fentanyl-related drugs as categorized by NFLIS. Opioid-related reports are

A central concern in using NFLIS measures as conditioning variables is that they are generated by both drug markets and enforcement behavior. Changes in NFLIS reports may reflect changes in the underlying drug supply, but they may also reflect changes in arrests, case submission, laboratory testing, or reporting. This matters for share-based measures of fentanyl saturation. For example, fentanyl's share of all drug reports can increase because fentanyl reports increase, but it can also increase because non-fentanyl reports fall. If decriminalization reduces low-level drug enforcement, then the denominator of the fentanyl-share measure may fall mechanically. In that case, conditioning on fentanyl's share of all reports could overstate the relevant fentanyl shock and absorb part of the policy-induced change in enforcement.

To examine this enforcement channel directly, we use drug arrest data from the FBI Uniform Crime Reporting program. We restrict the sample to agencies that report in all 12 months of each year over the seven-year analysis window, aggregate arrests to the state-month level, and construct reported arrest rates per 100,000 residents covered by the balanced agency panel. We use these data to document whether reported drug arrests changed sharply after decriminalization in Oregon and Washington. We also use the arrest data to interpret the NFLIS measures: if reported drug arrests fall after decriminalization, then declines in non-fentanyl NFLIS reports are consistent with reduced enforcement and submission for the drug offenses captured by the arrest measure, rather than necessarily reflecting a decline in non-fentanyl drug availability. We therefore use NFLIS measures both diagnostically and as sensitivity controls, while our preferred specification does not condition on them.

3.2 Methods

Our primary approach in this paper is based on synthetic control. For each donor state (j $\in$ {2,..., J}), we find a donor weight w_j which minimizes the pre-treatment (t $\in \{1, ..., T^{pre}\}$) differences in the outcome Y_{jt} between the treatment state (j=1) and synthetic control group which is a weighted combination of all donor states. Mathematically, we find w_j that solves the following optimization problem.

defined as fentanyl and fentanyl-related compounds, narcotic analgesics, and literal heroin reports identified in NFLIS.

$$\min_{w_s} \sum_{t=1}^{T^{pre}} (Y_{1t} - \sum_{j=2}^{J} w_j Y_{jt})^2$$

$$\text{s.t.} \sum_{j=2}^{J} w_j = 1$$

$$w_j \geq 0 \quad \forall j$$

Once we solve for w_j, we use these donor weights to construct a synthetic counterfactual during the post-treatment period, which we assume to be the outcome of interest in the treatment states had neither Measure 110 nor The Blake Fix been enacted. The average post-treatment difference between the actual treatment state and the synthetic counterfactual is the causal effect of the policy.

When we include additional covariates, such as fentanyl measures, they enter the synthetic control procedure as additional predictors used to select the donor weights, as is standard in synthetic control applications (Abadie et al., 2010, Abadie, 2021). That is, the outcome is not residualized using these variables and the post-treatment estimates are not regression-adjusted for them. Instead, the covariates affect the composition of the synthetic comparison unit by improving balance on specified pre-treatment characteristics.

We conduct inference using permutation-based p-values. For each donor state $j \in \{2, \ldots, J\}$, we assign treatment to that donor state and re-estimate the synthetic control using the same procedure as for the treated state. We then calculate the ratio of post-treatment to pre-treatment mean squared prediction error (MSPE), where MSPE is the average squared difference between the observed outcome and its synthetic counterfactual. A larger post/pre MSPE ratio indicates a larger post-treatment gap relative to pre-treatment fit. We rank the treated state's ratio among the corresponding placebo ratios and report the share of placebo ratios that are at least as large as the treated-state ratio as the permutation p-value. We include all donor-state placebos in this ranking rather than excluding placebo units on the basis of pre-treatment fit.[11]

In addition to a traditional synthetic control model, we also explore sensitivity to modern synthetic control estimators, including augmented synthetic control (Ben-Michael et al., 2021), synthetic difference-

[11] For instance, if Oregon's MSPE ratio is the largest among 50 treated and placebo units, the permutation p-value is $1/50 = 0.02$.

in-differences (Arkhangelsky et al., 2021), generalized synthetic control methods (Xu, 2017), and synthetic control using multiple outcomes (Tian et al., 2026). Finally, we also estimate the matrix completion synthetic control estimator, which is the main estimator used in (Zoorob et al., 2024). Matrix completion imputes a low-rank approximation of the outcome matrix using untreated states (Athey et al., 2021). While the matrix completion method and traditional synthetic control both aim to estimate missing counterfactuals, one key difference is the constraints imposed on the weights. Synthetic control constructs counterfactuals as a convex combination of donor units, whereas matrix completion imposes no such constraint on weights, allowing for a more flexible factor structure. In addition, traditional synthetic control incorporates covariates (e.g., fentanyl) through matching to generate weights, whereas the matrix completion method assumes that covariates are entered linearly.

4 Results

4.1 Main Results

In Figure 1, we present trends for Washington and Oregon and their synthetic counterparts for drug possession arrests. Both Oregon and Washington show a decline in arrests in 2020, which is consistent with other broad decreases in law enforcement during the COVID-19 pandemic. Then, abruptly in 2021, drug arrests drop in both states (ATT=-11.86 for Oregon and ATT=-6.42 for Washington). Permutation-based tests suggest these estimated changes are significant at the 5 percent level. Most importantly, it suggests that although Washington only decriminalized drug possession for 5 months on paper, in reality drug possession was effectively decriminalized throughout the end of 2023. For both states, the estimated differences are significant at the lowest feasible level, suggesting that Oregon and Washington had the largest deviation from any potential placebo states after adjusting for model fit in the pre-period.

We present trends in overdoses in Oregon and Washington and their synthetic counterparts in Figure 2.[12] As the models match on the outcome variable, the synthetic counterparts match the trend and levels for total overdoses. Beginning in 2021, and then widening more in 2022 and 2023, overdoses depart from the synthetic control counterparts following decriminalization. The first 2 columns of Appendix Table 4

[12]Our estimated weights assigned to each donor state can be found in Appendix Figure 3.

outline the point estimates of the ATTs by year. The estimated mortality gap grows over the post-treatment period rather than being driven solely by the first post-policy months, with larger annualized effects in 2022 and 2023 than in 2021. Across nearly three years following drug decriminalization, monthly mortality rates increased by 0.806 per 100,000 in Oregon (p-value = 0.02) and 0.732 per 100,000 in Washington (p-value = 0.02). For Oregon, this number corresponds to approximately a 38 percent increase relative to the baseline mean, and an additional 1,186 deaths over the 35-month post-treatment period. These estimated 1,186 excess deaths represent approximately 29.0 percent of the 4,096 overdose deaths observed between February 2021 and December 2023. For Washington, this number corresponds to approximately a 30 percent increase relative to the baseline mean, and an additional 1,895 deaths over the 34-month post-treatment period. These estimated 1,895 excess deaths represent approximately 23.8 percent of the 7,962 overdose deaths observed between March 2021 and December 2023.

4.2 Robustness

We conduct several robustness tests to explore the sensitivity of the estimated impacts of drug decriminalization (Ferman et al., 2020).

First, we experiment with other extensions of the synthetic control model. In columns (1) to (5) of Appendix Table 5, we find that our estimates continue to show a significant increase in drug overdose rate when we use augmented synthetic control (Ben-Michael et al., 2021), synthetic difference-in-differences (Arkhangelsky et al., 2021), generalized synthetic control methods (Xu, 2017), matrix completion synthetic control (Zoorob et al., 2024), and synthetic control using multiple outcomes (Tian et al., 2026). In column (6), we find that our point estimates are qualitatively similar when we use a two-way fixed-effects difference-in-differences estimator. However, our estimates are more imprecise and have larger p-values.

Next, in Appendix Figure 4, we experiment with excluding one donor state at a time. We find that no particular control states are driving our increase in drug overdose rate. Our results from this exercise suggest that the drug overdose rate increased by at least 0.77 per 100,000 in Oregon and 0.71 per 100,000 in Washington.

In Appendix Table 6, we experiment with matching our Synthetic Control only up to February 2020,

allowing for our donor pool to be based on the pre-pandemic trend of drug-related fatalities. The main idea of this exercise is to ensure that our choice of the donor pool and, consequently, our results are not driven by state-specific dynamic effects of COVID-19 on substance use. Our findings remain qualitatively similar, suggesting that the COVID-19 pandemic's unique effects on drug use nationwide are not adding bias.

In Appendix Table 7, we estimate placebo synthetic control for California, Nevada, and Idaho. This addresses concerns that decriminalization coincided spuriously with the arrival of fentanyl from the East Coast to the West Coast. We find little evidence that drug overdose rates increased in these nearby states, suggesting that the coincidental arrival of fentanyl in Oregon does not drive our estimates. In addition, along the same lines, we experiment with three alternative tests to ensure fentanyl is not driving our results. First, we estimate a two-way fixed effects model using only these three states as our counterfactual. While inference is more challenging due to the small cluster problem, our estimated treatment effects suggest increases of 0.77 to 0.79 in Oregon and 0.76 to 0.78 in Washington, which are qualitatively similar to those in our main specification.

In the next subsection, we explore the timing of fentanyl's arrival in the West more fully.

4.3 Fentanyl's Geographic Spread and Decriminalization

Fentanyl substantially changed the geography of the overdose epidemic in the mid-2010s (Moore et al., 2023). The Northeast and parts of the Midwest saw fentanyl saturate their illicit markets sooner than the rest of the country. Its later arrival to the West poses a potential statistical concern, despite Oregon and Washington showing abnormal deviations from their synthetic control counterparts. Indeed, Oregon and Washington's donor pool is comprised of states from the East including Maryland, Virginia, and North Carolina. This could raise questions about whether these counterfactuals experienced similar trends in unobservables we aim to capture to estimate the effect of decriminalization in an unbiased way.

In Figure 3, we present the evolution of four different measures of fentanyl supply saturation, captured from the National Forensic Laboratory Information System (NFLIS). Panel A presents fentanyl reports relative to all drug reports, the preferred control for Zoorob et al. (2024). Panel B contains per capita fentanyl reports, while Panel C contains per capita non-fentanyl reports. Panel D presents the fentanyl share

of opioid-related reports, which may better capture the substitution between fentanyl and other opioids as it saturates markets.

We first assess the measurement concern underlying the fentanyl share used by Zoorob et al. (2024). In Appendix Table 8, we estimate a synthetic control model using alternative NFLIS drug-report measures as outcomes. We find that NFLIS drug reports fall following decriminalization (columns 2 to 4), which is consistent with the large decline in drug arrests documented in Figure 1. Because the denominator of the fentanyl share measure includes all NFLIS drug reports, a decline in non-fentanyl reports will mechanically increase fentanyl's measured share even if fentanyl reports do not increase proportionally. This finding suggests that conditioning on fentanyl's share of all drug reports may introduce post-treatment bias when decriminalization itself changes enforcement procedures and priorities, and thereby influences laboratory submissions.

To examine the timing of fentanyl's arrival in Oregon and Washington, we estimate single-break dates for each NFLIS measure using sup-Wald structural break tests. This approach is closely related to the changepoint exercise in Zoorob et al. (2024), but instead of relying on the at-most-one-change procedure used in their application, we search over admissible half-year break dates and select the date with the largest structural-break test statistic. We apply this procedure to three measures: the fentanyl share of all drug reports, fentanyl reports per 100,000 population, and the fentanyl share of opioid-related drug reports. Figure 3 shows the underlying time series. The resulting break dates allow us to assess whether the inferred timing of fentanyl's saturation is robust to alternative measures of fentanyl exposure, and whether changes in non-fentanyl reporting may mechanically affect the share-based measure used by Zoorob et al. (2024).

We next perform sup-Wald and Bai-Perron tests on each univariate time series to estimate structural breaks, assuming one exists. Figure 4 shows results for the first set of tests, and Appendix Figure 5 provides results for the second. For the set of sup-Wald tests, we allow a single structural break for each state by performing a Chow test on each possible break date and identifying that with the highest F-statistic. The Bai-Perron test allows for multiple breaks, and we provide results for both mean and trend. When we allow for multiple breaks in the Bai-Perron testing, we concur with Zoorob et al. (2024)'s estimate of a late-2020 break for both Oregon and Washington. Since these breakpoints also coincide with COVID-19, a national

shock to overdose (Friedman and Akre, 2021), we also test for multiple structural breaks on each univariate time series demeaned by the national mean of the same measure. This removes common national shocks from Oregon and Washington's state-level structure. Appendix Figure 6 shows the results for a break in both mean and trend.

In Table 1, we assess the robustness of key modeling choices made by Zoorob et al. (2024): the use of fentanyl saturation control and the extension of the analysis period to include one additional post-treatment year.[13] To make estimates comparable across different estimation windows and temporal aggregations, we annualize the estimated ATTs on mortality rates. To evaluate the robustness of the fentanyl control, we compare models without any fentanyl control to four specifications: one using the original control employed by the authors (i.e., the share of fentanyl reports relative to all drug reports, denoted as "Share"), a second using fentanyl reports per capita, a third using the share of fentanyl reports relative to all opioid-related reports, and a fourth restricting the donor pool to states hit by fentanyl later, as measured by Zoorob et al. (2024). We note that Spencer (2023) and Joshi et al. (2023) restricted their analysis to a period that was later in the fentanyl wave, and consequently, the states receiving positive weights picked by either of their models actually have similar estimated structural break dates based on the structural break tests in Figure 4.

We include the per-capita measure as an alternative proxy for fentanyl supply because it captures absolute fentanyl exposure normalized by population while avoiding the endogenous all-drug-report denominator. This measure is also close to the approach used in prior work on the geographic spread of fentanyl, which measures fentanyl exposure using the log of per-capita fentanyl reports (Zoorob, 2019). The primary justification for using the per-capita measure, rather than raw counts, is that it reduces mechanical variation in NFLIS reports attributable to state population size (Zibbell et al., 2023). However, while our per-capita measure is potentially influenced by changes in police behavior, we argue that such behavioral changes are themselves a consequence of Measure 110. For this reason, our preferred specification omits any control variables due to changing behavior of those associated with the criminal justice system. Furthermore, if Zoorob et al. (2024)'s intent in controlling fentanyl supply was to isolate shifts in law enforcement activity,

[13] In Appendix Table 9, we explore the robustness of beginning the sample window to 2014, when the fentanyl crisis started to pick up.

then any regime change, as previously discussed, would itself induce endogenous changes in police behavior, as enforcement efforts are likely to focus on drugs most associated with overdose mortality. Indeed, in-depth qualitative interviews found police de-prioritized enforcement activities around most drugs, but continued to focus on investigations and supply interdiction related to fentanyl. Although these often did not result in low-level arrests, the heightened enforcement efforts around fentanyl did result in seizures of drugs that were later tested in state labs (hence appearing as reports in NFLIS) (Henderson et al., 2023).

In Panel I columns (1) and (2) of Table 1, we reproduce the patterns reported in Figure 3 of Zoorob et al. (2024).[14] Without a fentanyl control, the estimated effect of Measure 110 is positive and statistically significant (column 1). When we include Zoorob et al. (2024)'s preferred control, the fentanyl share of all drug reports, the estimate becomes small and statistically insignificant (column 2). Replacing this share-based control with fentanyl reports per capita restores a positive and statistically significant estimated relationship (column 3). Using fentanyl's share of opioid-related reports also produces a positive but statistically insignificant estimate over the original 2008–2022 sample (column 4), suggesting that the attenuation is sensitive to how fentanyl exposure is measured. Finally, when we restrict the donor pool to states where fentanyl saturation occurred in 2019 or later based on the sup-Wald structural break analysis, the estimated effect remains positive and statistically significant (column 5). Moreover, we also note that states receiving positive weight for earlier synthetic control analyses generally are estimated to have later fentanyl saturation dates in the structural break analysis.

Panel I, columns 6 through 10 show a similar pattern for Washington. Models using fentanyl's share of all drug reports or opioid-related reports attenuate the estimates and are not statistically significant at the 5 percent level. By contrast, using fentanyl reports per capita or restricting the donor pool to late-arriving fentanyl states yields positive and statistically significant estimates, with annualized ATTs of 4.34 and 5.07 deaths per 100,000 residents, respectively.

In Panel II of Table 1, we extend Zoorob et al. (2024)'s estimation strategy by adding 2023 (one additional year of post-treatment data). Strikingly, in columns (2) and (7), we find that even using Zoorob et al. (2024)'s preferred control for fentanyl saturation the estimated effects are consistent: drug decriminalization

[14] Our estimates using the NFLIS fentanyl-share control in column (2) cannot be replicated exactly because NFLIS subsequently updated or backfilled some reports, leading to slight discrepancies between the measure used by Zoorob et al. (2024), which was collected before 2024, and our measure, which was collected in 2026.

led to a statistically significant increase in drug overdose deaths, albeit the estimated magnitude is slightly smaller. The estimated effect of decriminalization on overdoses is of similar magnitude when we control fentanyl's saturation using fentanyl's share of opioid-related reports rather than not controlling for saturation or restricting the donor pool to later-saturating candidates.

Finally, in Panel III, we explore the sensitivity of main synthetic control analyses (which are similar to the models of Spencer (2023) and Joshi et al. (2023)), which use monthly data and a narrower pre-period starting in 2018, and explore the sensitivity to various controls for fentanyl market saturation. Notably, we document that the attenuation from using Zoorob et al. (2024)'s control disappears; across all specifications, the estimated effects of Measure 110 on overdose mortality remain positive and statistically significant, regardless of the fentanyl market saturation control used. Moreover, our findings are robust to whether we restrict our donor pool to "late" fentanyl arriving states. This likely reflects that restricting the sample period to 2018 and later omits early fentanyl arrival from consideration.

To summarize, fentanyl substantially changed the geography of the opioid epidemic. We replicate earlier analyses and also show it saturated markets in the Northeast earlier, and then spread throughout most of the rest of the country rapidly between 2019 and 2020. We revisit Zoorob et al. (2024)'s conclusion that controlling for fentanyl's saturation substantially alters the estimated relationship between decriminalization and overdoses. We first document that their preferred control may be flawed due to its structural relationship with police enforcement priorities directly related to decriminalization. We then find that extending the sample period through 2023, while adjusting for market saturation timing, has little impact on the estimated relationship between decriminalization and overdoses. We also find adjusting for fentanyl market saturation has little impact on the overall estimated effects of decriminalization when using a window between 2018 and 2023 and monthly overdose data, the modeling choices of Spencer (2023) and Joshi et al. (2023).

5 Conclusion

The overdose epidemic is among the most devastating public health crises in the United States over the past century. In an effort to reduce overdose deaths, many policymakers are experimenting with the balance between access to resources to reduce the harms of drug use and abuse. In February 2021, Oregon

became the first state to implement such a policy through Measure 110. Following Oregon, in March 2021, Washington also adopted a similar drug decriminalization law through a court order, "Blake Fix". In this paper, we estimate the longer-run impacts of drug decriminalization on overdose mortality. We find evidence that drug decriminalization led to approximately 0.8 and 0.7 per 100,000 increase in drug mortality in Oregon and Washington, respectively. In magnitude, these estimates imply 1,186 excess deaths in Oregon and 1,895 excess deaths in Washington, accounting for approximately 29.0 percent and 23.8 percent of observed post-policy overdose deaths, respectively. These estimates should be interpreted as reduced-form effects of the policy regimes that followed decriminalization. Our design does not separately identify whether mortality changed through drug use, market composition, enforcement behavior, treatment access, or other implementation channels.

While the magnitude of the estimated effect may be striking, we acknowledge several limitations in our empirical approach. Most notably, the pre-treatment period overlaps with the COVID-19 pandemic, meaning that donor states that served as suitable counterfactuals for Oregon pre-pandemic may not have remained so afterward. To address this, we exclude the COVID-19 period from our pre-treatment window and find that the estimated effect remains positive and statistically significant, with an ATT of 0.88 to 0.89.

Additionally, concerns have been raised that the fentanyl wave of the opioid epidemic may confound estimates of Measure 110's impact (Zoorob et al., 2024). We replicate, extend, and conduct sensitivity analyses on the findings in Zoorob et al. (2024) to assess (a) whether the saturation of the illicit drug market with fentanyl biased estimates of the impact of Measure 110, and (b) whether their half-year specification was robust to alternative pre-treatment periods. We find that their preferred proxy, fentanyl's share of all drug reports, creates a sharp increase in the share of reports related to fentanyl due to a *decrease* in non-fentanyl reports. This proxy is valid only if the decrease in non-fentanyl reports reflects a real decline in non-fentanyl drug presence rather than a post-policy change in enforcement, submissions, or testing; otherwise, the share can mechanically overstate fentanyl saturation. As an alternative, we control for fentanyl reports per capita, acknowledging that this measure is also influenced by enforcement intensity, and observe that the attenuation effect found using the fentanyl share disappears. We also explore fentanyl's share of opioid-related reports as an alternative proxy. Beyond sensitivity to the fentanyl measure itself, we also test the robustness of these

results to variations in the length of the pre-treatment window. Across shorter pre-treatment periods, the estimated effect remains positive and statistically significant.

We contribute to the harm reduction and substance use disorder intervention literature by providing credible evidence of the net mortality impacts of decriminalization policies. However, it is important to acknowledge that Measure 110 included many components beyond decriminalization, including several that are still ongoing, such as expanded funding for treatment services. We also draw attention to the so-called fentanyl loophole, which was not closed until June 2023. While our findings suggest that Measure 110 increased Oregon's overdose mortality rate, further research is needed to identify the specific mechanisms driving this increase. Finally, our analysis is necessarily limited to the unique context of Oregon's implementation of decriminalization, including its imperfections. Our results reflect the policy as it was enacted in Oregon, rather than decriminalization in the abstract.

Any evaluation of Oregon's outcomes must also consider parallel evidence from Washington's Blake Fix. Our analysis indicates that while Washington may have had a brief period of decriminalization on paper, it actually saw a sustained decrease in enforcement, making it comparable to Oregon's more publicized decriminalization. Washington also saw sustained increases in drug overdoses from 2021 through 2023. Future research, both quantitative and qualitative, is needed to fully understand the similarities in decriminalization (both states saw sustained declines in drug-related arrests) as well as the differences. For instance, Oregon created new treatment pathways, funding, and a Behavioral Health Resource Network, which remains in place to this day. This is just one important difference between the two decriminalization contexts, but it also took years to get funding allocated to treatment providers.

Now that drugs have been recriminalized in Oregon and Washington, it can be tempting to try to estimate new short-run analyses to measure whether recriminalization is already impacting overdoses. However, the results counsel caution in interpreting short post-policy windows, as our main estimates suggest the net effect of decriminalization grew substantially larger in years two and three. Moreover, recriminalization was accompanied by drug diversion and deflection strategies which vary across local law enforcement agencies. More research is needed to investigate and understand the complexity surrounding these new local policy choices.

Further research is also needed to assess the broader impacts of Measure 110 beyond overdose mortality. As Netherland et al. (2022) note: "The success or failure of Measure 110 has the potential to shape drug policy in the USA for decades to come, but 'success' or 'failure' is entirely dependent on what outcomes are being measured, how the data are gathered, and whether the findings are understood within the broader context of what is happening on the ground in Oregon." A comprehensive evaluation of the policy must therefore move beyond a narrow focus on mortality and engage with the full spectrum of ways in which substance use intersects with health, criminal justice, and human well-being.

References

A. B. C. News. Oregon officials declare state of emergency to address fentanyl crisis in Portland, January 2024. URL https://abcnews.go.com/US/oregon-officials-declare-state-emergency-address-fentanyl-crisis/story?id=106826829.

Alberto Abadie. Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, 59(2):391–425, June 2021. ISSN 0022-0515. doi: 10.1257/jel.20191450. URL https://www.aeaweb.org/articles?id=10.1257/jel.20191450.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490):493–505, June 2010. ISSN 0162-1459. doi: 10.1198/jasa.2009.ap08746. URL https://doi.org/10.1198/jasa.2009.ap08746. _eprint: https://doi.org/10.1198/jasa.2009.ap08746.

Amanda Agan and Sonja Starr. Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment*. *The Quarterly Journal of Economics*, 133(1):191–235, February 2018. ISSN 0033-5533. doi: 10.1093/qje/qjx028. URL https://doi.org/10.1093/qje/qjx028.

Anthony Johnson, Haven Wheelock, and Janie Gullickson. Drug Addiction Treatment and Recovery Act, August 2019. URL https://egov.sos.state.or.us/elec/web_irr_search.record_detail?p_reference=20200044..LSCYYYDRUG.

Dmitry Arkhangelsky, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager. Synthetic Difference-in-Differences. *American Economic Review*, 111(12):4088–4118, December 2021. ISSN 0002-8282. doi: 10.1257/aer.20190159. URL https://www.aeaweb.org/articles?id=10.1257/aer.20190159.

Susan Athey, Mohsen Bayati, Nikolay Doudchenko, Guido Imbens, and Khashayar Khosravi. Matrix Completion Methods for Causal Panel Data Models. *Journal of the American Statistical Association*, 116(536):1716–1730, October 2021. ISSN 0162-1459. doi: 10.1080/01621459.2021.1891924. URL https://doi.org/10.1080/01621459.2021.1891924. _eprint: https://doi.org/10.1080/01621459.2021.1891924.

Gary S. Becker. Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2):169–217, March 1968. ISSN 0022-3808. doi: 10.1086/259394. URL https://www.journals.uchicago.edu/doi/abs/10.1086/259394.

Gary S. Becker, Kevin M. Murphy, and Michael Grossman. The Economic Theory of Illegal Goods: The Case of Drugs, December 2004. URL https://www.nber.org/papers/w10976.

Eli Ben-Michael, Avi Feller, and Jesse Rothstein. The Augmented Synthetic Control Method. *Journal of the American Statistical Association*, 116(536):1789–1803, October 2021. ISSN 0162-1459. doi: 10.1080/01621459.2021.1929245. URL https://doi.org/10.1080/01621459.2021.1929245. _eprint: https://doi.org/10.1080/01621459.2021.1929245.

Hannah Byles, Navid Sedaghat, Nathan Rider, William Rioux, Alexandra Loverock, Boogyung Seo, Avnit Dhanoa, Taylor Orr, Nicole Dunnewold, Lisa Tjosvold, and S Monty Ghosh. Barriers to calling emergency services amongst people who use substances in the event of overdose: A scoping review. *International Journal of Drug Policy*, 132:104559, October 2024. ISSN 0955-3959. doi: 10.1016/j.drugpo.2024.104559. URL https://www.sciencedirect.com/science/article/pii/S0955395924002433.

Carolina Caetano, Brantly Callaway, Stroud Payne, and Hugo Sant'Anna Rodrigues. Difference in Differences with Time-Varying Covariates, June 2024. URL http://arxiv.org/abs/2202.02903. arXiv:2202.02903.

Colin F. Camerer, Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan, Emma Heikensten, Felix Holzmeister, Taisuke Imai, Siri Isaksson, Gideon Nave, Thomas Pfeiffer, Michael Razen, and Hang Wu. Evaluating replicability of laboratory experiments in economics. *Science*, 351(6280):1433–1436, March 2016. doi: 10.1126/science.aaf0918. URL https://www.science.org/doi/10.1126/science.aaf0918.

Matias D. Cattaneo, Yingjie Feng, and Rocio Titiunik. Prediction Intervals for Synthetic Control Methods. *Journal of the American Statistical Association*, 116(536):1865–1880, October 2021. ISSN 0162-1459, 1537-274X. doi: 10.1080/01621459.2021.1979561. URL https://www.tandfonline.com/doi/full/10.1080/01621459.2021.1979561.

Center for Disease Control and Prevention, National Center for Health Statistics. U.S. Overdose Deaths In 2021 Increased Half as Much as in 2020 - But Are Still Up 15%, May 2022. URL https://www.cdc.gov/nchs/pressroom/nchs_press_releases/2022/202205.htm.

Andrew C. Chang and Phillip Li. Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Usually Not". *Finance and Economics Discussion Series*, 2015(83):1–26, October 2015. ISSN 19362854. doi: 10.17016/FEDS.2015.083. URL https://www.federalreserve.gov/econres/feds/is-economics-research-replicable-sixty-published-papers-from-thirteen-journals-say-quotusually-notqu
htm.

Victor Chernozhukov, Kaspar Wuthrich, and Yinchu Zhu. A t-test for synthetic controls, January 2024. URL http://arxiv.org/abs/1812.10820. arXiv:1812.10820 [econ].

Daniel Ciccarone. The Triple Wave Epidemic: Supply and Demand Drivers of the US Opioid Overdose Crisis. *The International journal on drug policy*, 71:183–188, September 2019. ISSN 0955-3959. doi: 10.1016/j.drugpo.2019.01.010. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6675668/.

Corey S. Davis, Spruha Joshi, Bianca D. Rivera, and Magdalena Cerdá. Changes in arrests following decriminalization of low-level drug possession in Oregon and Washington. *International Journal of Drug Policy*, 119:104155, September 2023. ISSN 0955-3959. doi: 10.1016/j.drugpo.2023.104155. URL https://www.sciencedirect.com/science/article/pii/S0955395923002025.

William N. Evans, Ethan Lieber, and Patrick Power. How the Reformulation of OxyContin Ignited the Heroin Epidemic, April 2018. URL https://www.nber.org/papers/w24475.

Bruno Ferman, Cristine Pinto, and Vitor Possebom. Cherry Picking with Synthetic Controls. *Journal of Policy Analysis and Management*, 39(2):510–532, 2020. ISSN 1520-6688. doi: 10.1002/pam.22206. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/pam.22206. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pam.22206.

Joseph Friedman and Samir Akre. COVID-19 and the Drug Overdose Crisis: Uncovering the Deadliest Months in the United States, January July 2020. *American Journal of Public Health*, 111(7):1284–1291, July 2021. ISSN 0090-0036. doi: 10.2105/AJPH.2021.306256. URL https://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2021.306256.

Sónia Félix, Pedro Portugal, and Ana Sofia Tavares. Going after the Addiction, Not the Addicted: The Impact of Drug Decriminalization in Portugal, July 2017. URL https://papers.ssrn.com/abstract=3010673.

Catalina Gaitán. Oregon's Measure 110 drug treatment hotline has served only 577, gets new operator. *Oregon Live*, December 2023. URL https://www.msn.com/en-us/health/other/oregon-s-measure-110-drug-treatment-hotline-has-served-only-577-gets-new-operator/ar-AA1lt98x.

Kelsey Henderson, Christopher Campbell, and Brian Renauer. Impacts of Successive Drug Legislation Shifts: Qualitative Observations from Oregon Law Enforcement [Interim Report: Year One]. *Criminology and Criminal Justice Faculty Publications and Presentations*, January 2023. URL https://pdxscholar.library.pdx.edu/ccj_fac/114.

Frank G. Hillary and John D. Medaglia. What the replication crisis means for intervention science. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, 154:3–5, August 2020. ISSN 0167-8760. doi: 10.1016/j.ijpsycho.2019.05.006. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6842660/.

Harry J. Holzer, Steven Raphael, and Michael A. Stoll. Perceived Criminality, Criminal Background Checks, and the Racial Hiring Practices of Employers. *The Journal of Law and Economics*, 49(2):451–480, October 2006. ISSN 0022-2186. doi: 10.1086/501089. URL https://www.journals.uchicago.edu/doi/abs/10.1086/501089.

Spruha Joshi, Bianca D. Rivera, Magdalena Cerdá, Gery P. Guy, Jr, Andrea Strahan, Haven Wheelock, and Corey S. Davis. One-Year Association of Drug Possession Law Change With Fatal Drug Overdose in Oregon and Washington. *JAMA Psychiatry*, September 2023. ISSN 2168-622X. doi: 10.1001/jamapsychiatry.2023.3416. URL https://doi.org/10.1001/jamapsychiatry.2023.3416.

Johanna Catherine Maclean, Justine Mallatt, Christopher J. Ruhm, and Kosali Simon. Economic Studies on the Opioid Crisis: A Review, November 2020. URL https://www.nber.org/papers/w28067.

Dhingra Manka. Washington SB5476 | 2021-2022 | Regular Session, 2021. URL https://legiscan.com/WA/text/SB5476/id/2403634.

Timothy J. Moore, William W. Olney, and Benjamin Hansen. Importing the Opioid Crisis? International Trade and Fentanyl Overdoses, November 2023. URL https://www.nber.org/papers/w31885.

Jules Netherland, Alex H. Kral, Danielle C. Ompad, Corey S. Davis, Ricky N. Bluthenthal, Nabarun Dasgupta, Michael Gilbert, Riona Morgan, and Haven Wheelock. Principles and Metrics for Evaluating Oregon's Innovative Drug Decriminalization Measure. *Journal of Urban Health*, 99(2):328–331, April 2022. ISSN 1468-2869. doi: 10.1007/s11524-022-00606-w. URL https://doi.org/10.1007/s11524-022-00606-w.

Oregon Health Justice Recovery Alliance. Oregon Health Justice Recovery Alliance, June 2021. URL https://web.archive.org/web/20210609033613/https://healthjusticerecovery.org/.

Analisa Packham. Syringe exchange programs and harm reduction: New evidence in the wake of the opioid epidemic. *Journal of Public Economics*, 215:104733, November 2022. ISSN 0047-2727. doi: 10.1016/j.jpubeco.2022.104733. URL https://www.sciencedirect.com/science/article/pii/S0047272722001359.

Wayne J. Pitts, David Heller, Hope Smiley-McDonald, BeLinda Weimer, Megan Grabenauer, Katherine Bollinger, Jeri Ropero-Miller, and DeMia Pressley. Understanding research methods, limitations, and applications of drug data collected by the National Forensic Laboratory Information System (NFLIS-Drug). *Journal of Forensic Sciences*, 68(4):1335–1342, 2023. ISSN 1556-4029. doi: 10.1111/1556-4029.15269. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1556-4029.15269. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1556-4029.15269.

Representative Jason Kropf and Senator Kate Lieber. House Bill 4002, April 2024. URL https://olis.oregonlegislature.gov/liz/2024R1/Measures/Overview/HB4002.

Kellen Russoniello, Sheila P. Vakharia, Jules Netherland, Theshia Naidoo, Haven Wheelock, Tera Hurst, and Saba Rouhani. Decriminalization of drug possession in Oregon: Analysis and early lessons. *Drug Science, Policy and Law*, 9:20503245231167407, January 2023. ISSN 2050-3245. doi: 10.1177/20503245231167407. URL https://doi.org/10.1177/20503245231167407.

Chelsea L. Shover, Titilola O. Falasinnu, Candice L. Dwyer, Nayelie Benitez Santos, Nicole J. Cunningham, Rohan B. Freedman, Noel A. Vest, and Keith Humphreys. Steep increases in fentanyl-related mortality west of the Mississippi River: Recent evidence from county and state surveillance. *Drug and Alcohol Dependence*, 216:108314, November 2020. ISSN 0376-8716. doi: 10.1016/j.drugalcdep.2020.108314. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7521591/.

Hope M. Smiley-McDonald, Peyton R. Attaway, Lynn D. Wenger, Kathryn Greenwell, Barrot H. Lambdin, and Alex H. Kral. "All carrots and no stick": Perceived impacts, changes in practices, and attitudes among law enforcement following drug decriminalization in Oregon State, USA. *International Journal of Drug Policy*, 118:104100, August 2023. ISSN 0955-3959. doi: 10.1016/j.drugpo.2023.104100. URL https://www.sciencedirect.com/science/article/pii/S0955395923001470.

Adam Soliman. Disrupting Drug Markets: The Effects of Crackdowns on Rogue Opioid Suppliers. *American Economic Journal: Economic Policy*, 17(4):165–191, November 2025. ISSN 1945-7731. doi: 10.1257/pol.20230640. URL https://www.aeaweb.org/articles?id=10.1257/pol.20230640.

Noah Spencer. Does drug decriminalization increase unintentional drug overdose deaths?: Early evidence from Oregon Measure 110. *Journal of Health Economics*, 91:102798, September 2023. ISSN 1879-1646. doi: 10.1016/j.jhealeco.2023.102798.

State of Washington. State v. Blake, April 2021. URL https://www.co.pacific.wa.us/courts/state-v-blake.html.

Wei Tian, Seojeong Lee, and Valentyn Panchenko. Synthetic Controls with Multiple Outcomes. *The Econometrics Journal*, page utag005, March 2026. ISSN 1368-4221. doi: 10.1093/ectj/utag005. URL https://doi.org/10.1093/ectj/utag005.

Vote No on Measure 110. News | Vote No on Measure 110!, 2020. URL https://web.archive.org/web/20210126193538/https://votenoon110.com/news/.

Kelsa L. West, Karla Lindquist, and Luke N. Rodda. Fentanyl epidemic hits the U.S. West Coast: Opioid-related deaths in San Francisco from 2009–2019. *International Journal of Drug Policy*, 95:103402, September 2021. ISSN 0955-3959. doi: 10.1016/j.drugpo.2021.103402. URL https://www.sciencedirect.com/science/article/pii/S0955395921003078.

Yiqing Xu. Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. *Political Analysis*, 25(1):57–76, January 2017. ISSN 1047-1987, 1476-4989. doi: 10.1017/pan.2016.2. URL https://www.cambridge.org/core/journals/political-analysis/article/generalized-synthetic-control-method-causal-inference-with-interactive-fixed-effects-models/B63A8BD7C239DD4141C67DA10CD0E4F3.

Jon E. Zibbell, Arnie Aldridge, Megan Grabenauer, David Heller, Sarah Duhart Clarke, DeMia Pressley, and Hope Smiley McDonald. Associations between opioid overdose deaths and drugs confiscated by law enforcement and submitted to crime laboratories for analysis, United States, 2014–2019: an observational study. *The Lancet Regional Health – Americas*, 25, September 2023. ISSN 2667-193X. doi: 10.1016/j.lana.2023.100569. URL https://www.thelancet.com/journals/lanam/article/PIIS2667-193X(23)00143-6/fulltext?uuid=uuid%3Ab14feb6b-9726-4eb0-a6e9-07f5f90622f7.

Michael Zoorob. Fentanyl shock: The changing geography of overdose in the United States. *International Journal of Drug Policy*, 70:40–46, August 2019. ISSN 0955-3959. doi: 10.1016/j.drugpo.2019.04.010. URL https://www.sciencedirect.com/science/article/pii/S0955395919301136.

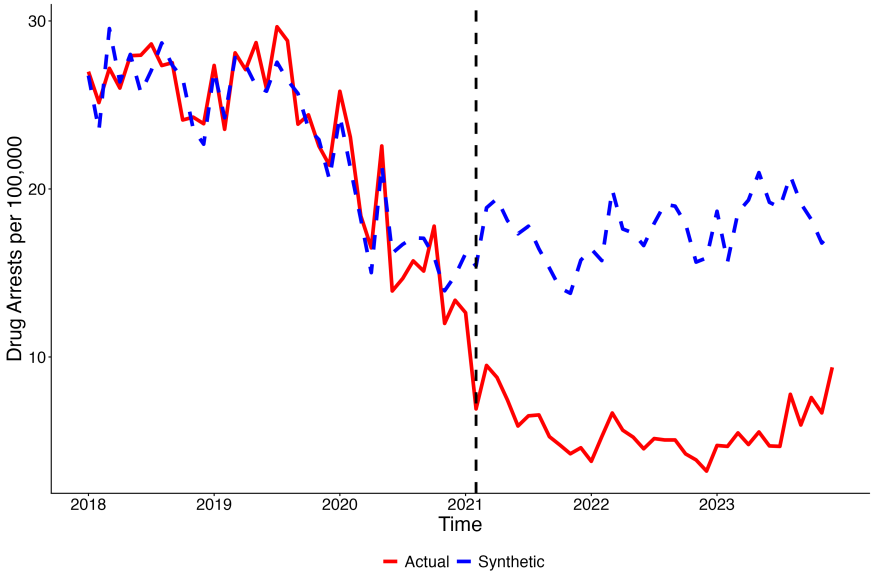
Michael J. Zoorob, Ju Nyeong Park, Alex H. Kral, Barrot H. Lambdin, and Brandon del Pozo. Drug Decriminalization, Fentanyl, and Fatal Overdoses in Oregon. *JAMA Network Open*, 7(9):e2431612, September 2024. ISSN 2574-3805. doi: 10.1001/jamanetworkopen.2024.31612. URL https://doi.org/10.1001/jamanetworkopen.2024.31612.

6 Tables & Figures

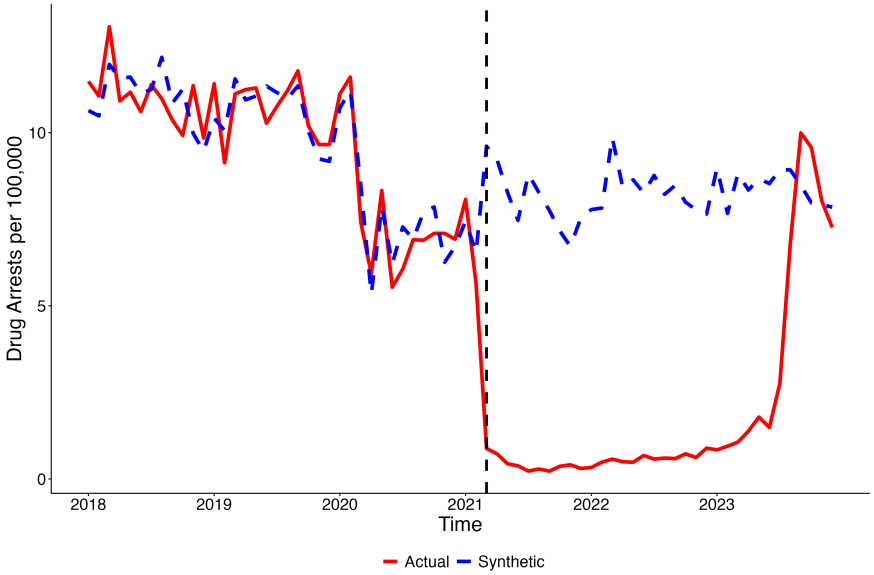
6.1 Figures

Figure 1: UCR Estimate, Arrests

(a) Oregon (ATT=-11.86, p-value=0.021)



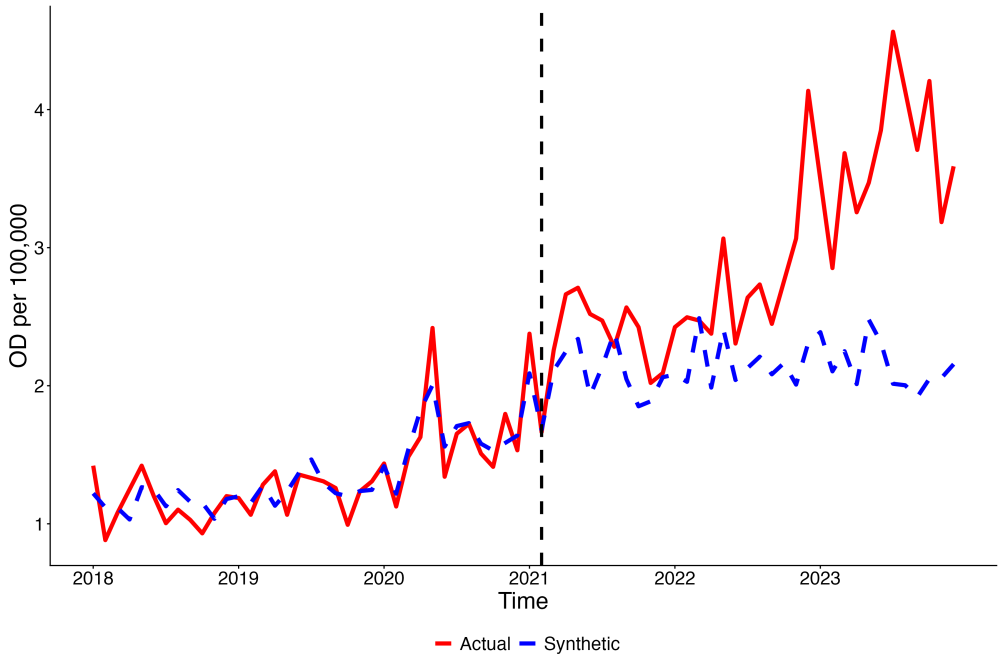
(b) Washington (ATT=-6.42, p-value=0.021)



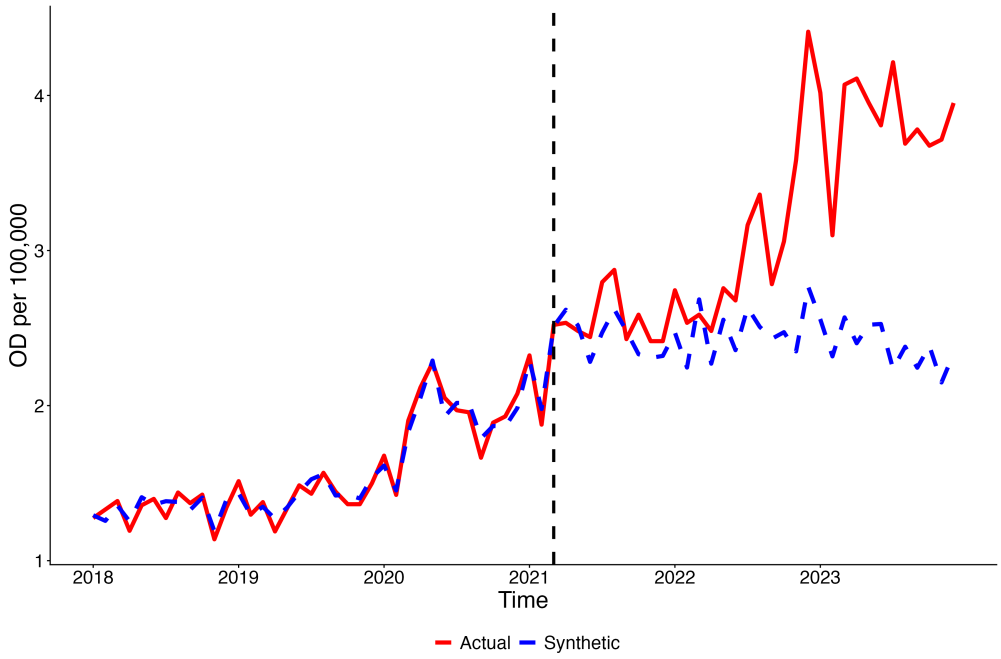
Notes: This figure presents synthetic control estimates for drug possession arrests using the UCR data for Oregon and Washington. Despite the recriminalization of possession in Washington, arrest rates do not revert to pre-treatment levels until late 2023, suggesting persistent changes in enforcement behavior that may help explain the observed patterns in overdose mortality.

Figure 2: Oregon and Washington Overdoses vs. Synthetic Counterfactuals

(a) Oregon (ATT=0.806, p-value=0.02)

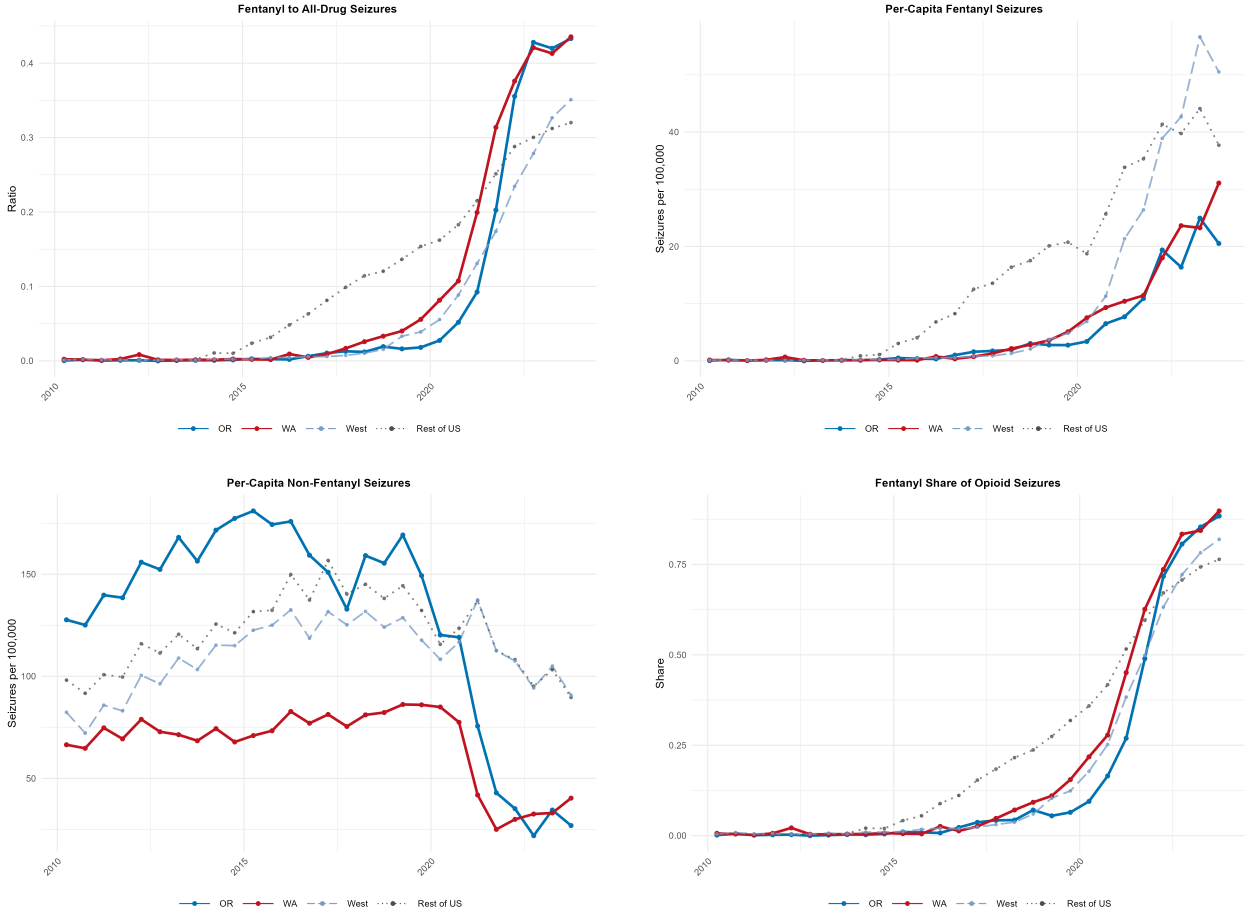


(b) Washington (ATT=0.732, p-value=0.02)



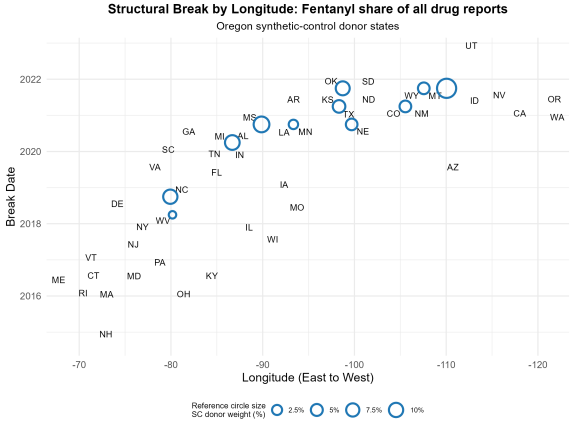
Notes: Synthetic control estimates for Oregon (panel a) and Washington (panel b) are shown. While both states present similar pre-treatment fit (MSPE = 0.0265 for Oregon and 0.0038 for Washington), a substantial divergence between each treated state and its synthetic control emerges after treatment and becomes pronounced in 2022–2023.

Figure 3: Fentanyl Supply Measures in Oregon, Washington, Western States, and the Rest of the United States

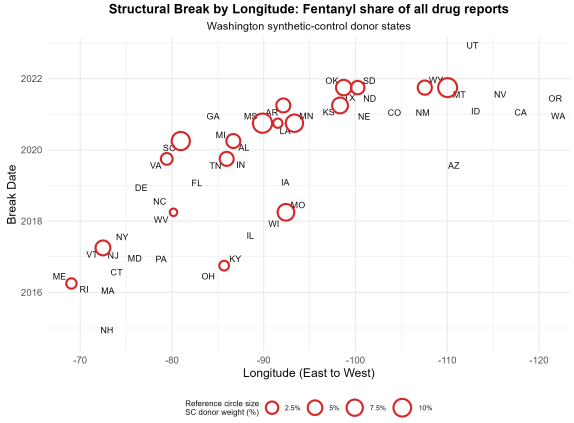


Notes: This figure uses state-level NFLIS drug report data aggregated to half-year periods. Observations are indexed by state and half-year, where the first half of each calendar year includes January–June and the second half includes July–December. The plotted series cover 2010 through 2023. The figure plots Oregon, Washington, the mean across Western comparison states, and the mean across all remaining U.S. states. Western comparison states are Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming, Alaska, and Hawaii. Panel A reports fentanyl and fentanyl-related reports as a share of all drug reports. Panels B and C report fentanyl and non-fentanyl reports per 100,000 population, respectively. Panel D reports fentanyl and fentanyl-related reports as a share of opioid-related reports. Opioid-related reports are defined as fentanyl and fentanyl-related compounds, narcotic analgesics, and heroin reports identified in NFLIS. Group comparison series are unweighted means across states within each group.

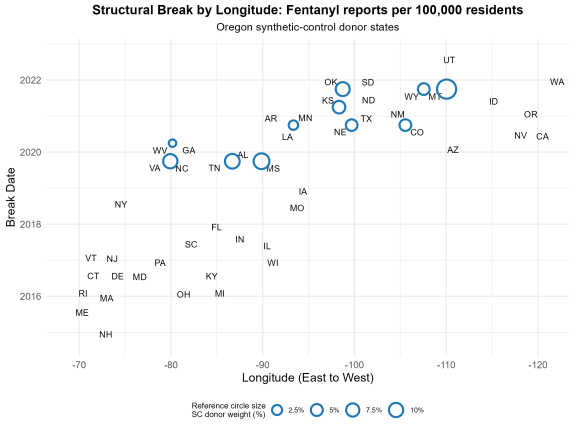
Figure 4: sup-Wald Single Break Tests and Synthetic-Control Donor Weights



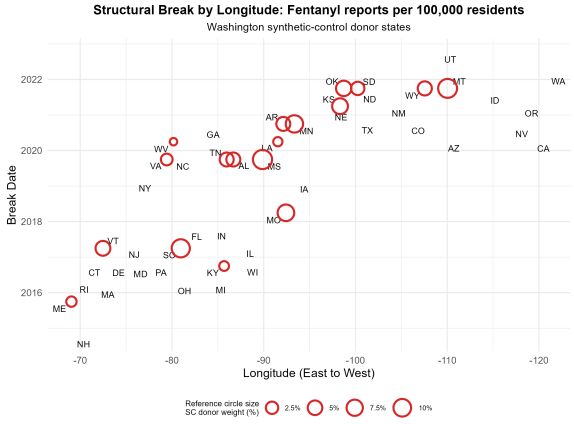
(a) Oregon: fentanyl share of all drug reports



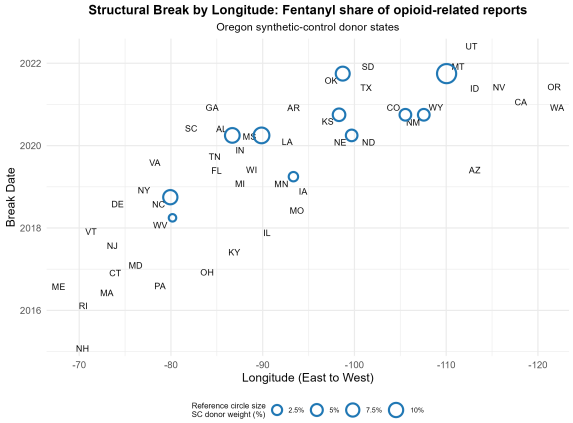
(b) Washington: fentanyl share of all drug reports



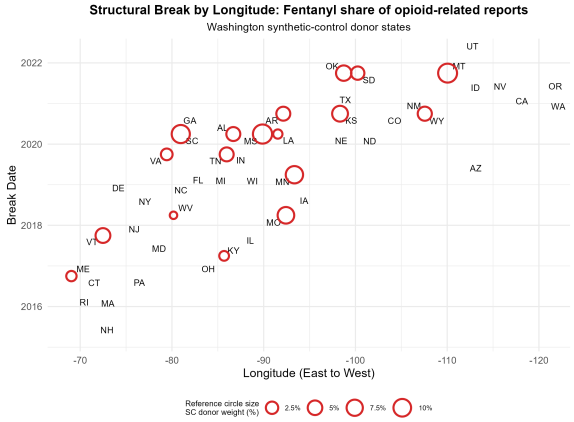
(c) Oregon: fentanyl reports per capita



(d) Washington: fentanyl reports per capita



(e) Oregon: fentanyl share of opioid-related reports



(f) Washington: fentanyl share of opioid-related reports

Notes: This figure presents sup-Wald structural break tests for three state-level NFLIS report measures aggregated to half-year periods. Each state label is plotted at the state's geographic centroid longitude and estimated single break date. Blue rings identify states receiving positive donor weight in the Oregon all-overdose synthetic-control specification, and red rings identify states receiving positive donor weight in the Washington all-overdose synthetic-control specification. Ring size is proportional to the donor weight. Opioid-related reports are defined as fentanyl and fentanyl-related compounds, narcotic analgesics, and heroin reports identified in NFLIS.

6.2 Tables

Table 1: Exploring Sensitivity to Alternative Fentanyl Controls

	Oregon					Washington				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel I: Matrix Completion SC Using Half-Year Data, 2008-2022									
ATT	3.661***	-0.634	4.937***	0.532	4.997***	2.682**	-2.135*	4.338***	-1.118	5.072***
p-value	{0.000}	{0.575}	{0.000}	{0.656}	{0.002}	{0.044}	{0.094}	{0.000}	{0.369}	{0.001}
	Panel II: Matrix Completion SC Using Half-Year Data, 2008-2023									
ATT	8.573***	4.027***	10.188***	5.327***	9.436***	7.855***	3.066***	9.265***	4.307***	9.123***
p-value	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.010}	{0.000}	{0.003}	{0.000}
	Panel III: Traditional SC Using Monthly Data, 2018-2023									
ATT	9.677**	10.376**	8.85**	11.892**	12.05**	8.785**	8.645**	8.244**	10.235**	9.794**
p-value	{0.02}	{0.02}	{0.02}	{0.02}	{0.038}	{0.02}	{0.02}	{0.02}	{0.02}	{0.038}
Fentanyl Control	None	Drug share	PC	Opioid share	None	None	Drug share	PC	Opioid share	None
Donor Pool Restriction?	No	No	No	No	Yes	No	No	No	No	Yes

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: Panel I estimates Zoorob et al. (2024)'s synthetic control method, which uses half-year data from 2008 to 2022. Panel II estimates Zoorob et al. (2024)'s synthetic control method, but with an additional year of post-treatment data. Panel III estimates our preferred synthetic control method, which uses monthly data from 2018 to 2023. For Panels I and II, the ATT is multiplied by 2, and for Panel III, the ATT is multiplied by 12 to report an average effect over the year. In columns (2) and (7), we include a control for the ratio of fentanyl reports to all drug reports, which is the covariate used in Zoorob et al. (2024). In columns (3) and (8), we use fentanyl reports per capita as our covariate, and in columns (4) and (9), we use the ratio of fentanyl reports to opioid-related reports as our covariate. In columns (5) and (10), we restrict our donor pool to states where the fentanyl crisis arrived in 2019 or later, as defined by Zoorob et al. (2024). In Panel III columns (1), (5), (6), and (10), we use the rate of drug overdose for all pre-treatment months to find our synthetic weights. In Panel III columns (2) to (4) and (7) to (9), we use half-year pre-treatment rates of drug overdose as well as half-year fentanyl controls from 2018-2020 to find our synthetic weights. The p-values generated using the jackknife method (Panels I and II) and permutation tests (Panel III) are reported inside the curly brackets.

Appendix

A Reconciliation of Joshi et al. (2023) and Spencer (2023)

In this appendix, we decompose key differences between the two papers that led to differences in their conclusions about the first-year impacts of decriminalization.

A.1 Definitions

Users of the CDC WONDER database can filter underlying causes of death using ICD (International Classification of Diseases) codes.[15] Spencer (2023) filters by unintentional overdoses (X40–X44) whereas Joshi et al. (2023) filter by all overdoses (X40–X44, X60–X64, X85, Y10–Y14, multiple-cause-of-death T36–T50).

As categorized by the World Health Organization, unintentional overdoses are external causes of morbidity and mortality. Specifically (X40-X49) is defined as "accidental overdose of drug, wrong drug given or taken in error, and drug taken inadvertently"; "accidents in the use of drugs, medicaments and biological substances in medical and surgical procedures;" and "(self-inflicted) poisoning, when not specified whether accidental or with intent to harm." The subset Spencer (2023) uses, X40-X44, is categorized under "Accidental poisoning by and exposure to noxious substances." X60–X64 includes intentional self-harm from poisoning by drugs. X85 is classified as homicidal poisoning. Y10–Y14 covers all drug poisonings with undetermined intent. Finally, T36–T50 is classified as "Poisoning by drugs, medicaments and biological substances" and would be included in a multiple-cause-of-death download of the series.[16]

The average difference between total overdoses nationally is about 692 per month. The average using Joshi et al. (2023)'s definition is 7710, and the average using Spencer (2023)'s definition is 7019. The standard deviation is 1539 and 1543, respectively. In the states of Oregon and Washington, the averages using Joshi et al. (2023)'s definition are 89 and 176, respectively. For Spencer (2023)'s definition, the monthly averages are 79 in Oregon and 162 in Washington. Appendix Figure 1 plots the national time trend for these different definitions. We note that the majority of drug deaths are due to accidental overdoses (Panel a), and these two outcomes follow a similar trend over time (Panel b).[17]

There are reasonable arguments for using different outcome measures in each study. Drug decriminalization may affect mortality through a demand-side channel—that is, by increasing the number of users and/or the frequency or dosage of use—rather than through increases in suicide or homicide. As a result, researchers may have more statistical power when focusing specifically on accidental overdoses. However, analyzing all drug-related mortality may help mitigate potential classification or measurement errors, since sometimes the data can misclassify unintentional overdoses as suicides or homicides.

A.2 Timing

The synthetic control method requires selecting a pre-treatment period over which to match outcomes to determine the optimal synthetic weights. As Abadie et al. (2010) notes: "In any research design that exploits time variation in the outcome variable to estimate the effect of an intervention, synthetic control estimators may be biased if forward-looking economic agents react in advance of the policy intervention under investigation." Thus, researchers need to use economic theory to justify their decisions on which periods to minimize the sum of squared prediction errors and determine the correct weights.

Spencer (2023) matches on 36 monthly pre-treatment periods between January 2018 and December 2020.[18] While Spencer (2023)'s pre-treatment sample window ends one month before the implementation of Measure 110, he argues that he chose this decision because he wanted consistency across outcomes. Some of his outcomes are at the annual level, meaning the synthetic control can only match until December and define treatment starting in January. Joshi et al. (2023) also begin their sample in January 2018 but utilize

[15] At this time, the tenth revision of these codes (ICD-10) is being used. Codes can be searched using the ICD-10-CM Browser Tool `https://www.cdc.gov/nchs/icd/icd10cm_browsertool.htm`

[16] Filtering by "multiple-cause-of-death" allows the user to subset even overdoses into a specific drug used.

[17] In Appendix Figure 2, we present an analogous state-level monthly comparison for Oregon and Washington.

[18] Spencer (2023) makes a note of the policy being implemented on February 1, 2021, in the paper, but chooses January as the treatment unit in coding and chart output. In his replication package, this is identified by unit 37 (t=37) being selected as the treatment period, which in a monthly dataset beginning (t=1) in January of 2018, corresponds to January of 2021.

all 37 pre-treatment periods, including January 2021, citing they "followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline for observational studies."

Beyond ensuring consistency across outcomes, the correct specification between the two models depends on whether anticipation effects are relevant. For instance, if there was anticipation prior to implementation (e.g., in January 2021), where overdoses had already begun to change, researchers should exclude the anticipatory window for matching. While we do not statistically verify whether and when anticipation occurred, our robustness checks compare specifications that end the matching period in December 2020 versus January 2021.

Finally, the authors differ in the number of observations in the post-treatment period. Spencer runs his analysis through the end of 2021, while Joshi et al. (2023) run their analysis through the end of March, 2022.

A.3 Data Set

For reasons of security and privacy, the CDC suppresses the number of deaths for an observation if the number of deaths is below ten. Spencer (2023) uses this suppressed version of the data and imputes integers from a random uniform distribution on the interval [0,9] to fill suppressed values. Joshi et al. (2023) have access to the unrestricted version of the data.

In Appendix Figure 8, we re-estimate our synthetic control 1,000 times using different imputed values and plot the distribution of possible point estimates. The findings from this exercise confirm that the point estimates reported by Spencer (2023) are reasonable and not driven by artifacts of imputed data.

A.4 Replication & Reconciliation Exercise

In column 1 of Panel I and column 4 of Panel II of Appendix Table 1, we successfully replicate both Joshi et al. (2023) and Spencer (2023) results for Measure 110 in Oregon. However, we note that our point estimates are slightly different from those reported by the papers. One explanation for this difference is the difference in the Stata command that we utilize.[19] Moreover, in fig. 3, we show the synthetic control weights that we used to replicate both papers. These weights are qualitatively similar to the previous studies.

In the remaining columns of Appendix Table 1, we experiment with changing the specification to be more similar to each other's specification, including changing the definitions of drug overdose (Panels I and II), changing the treatment start month (columns 1 and 3 vs. 2 and 4), and sample window (columns 1 and 2 vs. 3 and 4). We document a few findings. First, ceteris paribus, our estimated effects are smaller when we focus on all drug-related mortality instead of unintentional drug overdoses. Second, our point estimate and precision are sensitive to whether we begin our treatment in January or February. For instance, when we focus on unintentional drug overdoses and end our sample in March 2022 (Panel I, columns (3) and (4)), our estimated coefficient decreases by 0.057 deaths per 100,000 residents per month. Third, the results are robust to the inclusion of three more months of post-treatment data.

We repeat the same exercise in Appendix Table 2; however, this time we focus on Washington. Similar to Oregon, we find that the differences in the two studies' conclusions stem from differences in the definitions of drug overdoses and when the treatment date began (and when the synthetic control was matched until).

In Panel I of Appendix Table 3, we re-estimate Spencer (2023) and Joshi et al. (2023)'s preferred model and add 90 percent prediction intervals (Cattaneo et al., 2021). In Panel II, we also calculate the 90 percent confidence intervals using t-tests following Chernozhukov et al. (2024) and using K = 3 cross-fold validation. We note that our point estimates and statistical significance, reported in Appendix Table 3, are slightly different from those of our replication results. This discrepancy may be due to slight differences in methodologies and weights assigned to each donor state. Nevertheless, we note that the 90 percent confidence and prediction interval overlap between Spencer (2023)'s preferred specification and Joshi et al. (2023)'s main model. The results from this exercise confirm that the estimated treatment effects in the two papers may not statistically differ.

Finally, in Appendix Table 4, we compare Joshi et al. (2023) and Spencer (2023)'s preferred model using the longer-run post-treatment window covering up to 2023 (columns 1 and 2 versus columns 3 and 4). We find that their findings converge in the long run, suggesting that the short run disagreement is largely explained

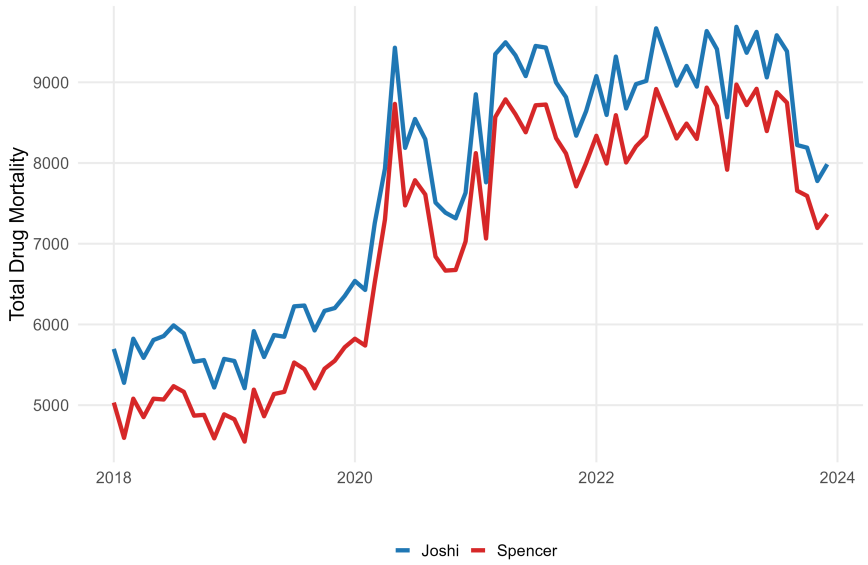
[19] We use Synth package, whereas Spencer (2023) uses Synth2 and Joshi et al. (2023) use synth_runner.

by outcome definitions, treatment timing, and limited follow-up rather than by irreconcilable differences in the synthetic-control designs.

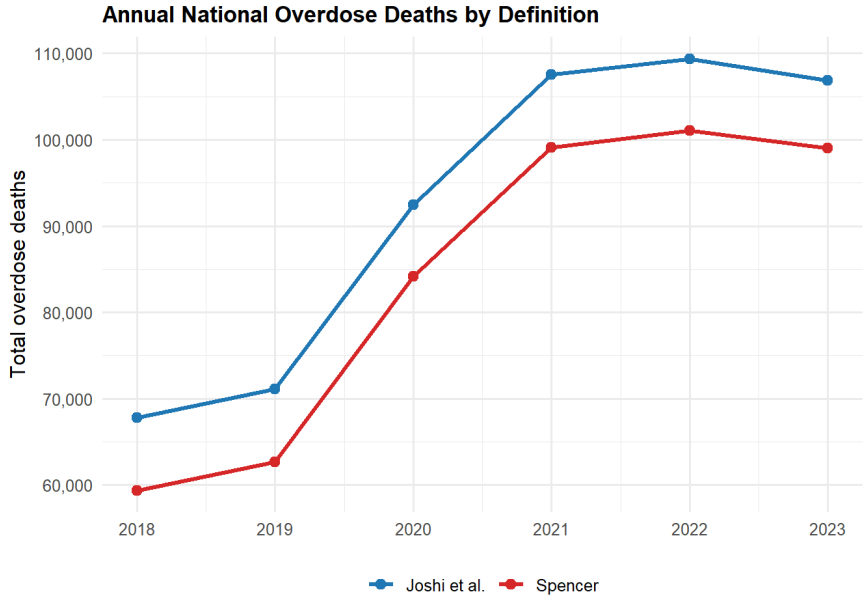
B Appendix Figures

Appendix Figure 1: National overdose deaths using all-overdose and unintentional-overdose outcome definitions.

(a) Monthly overdose deaths



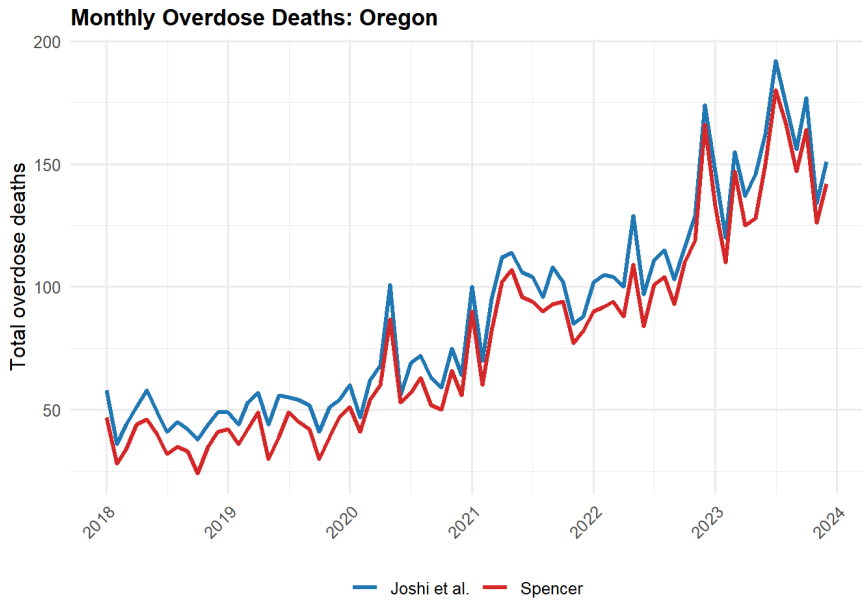
(b) Annual overdose deaths



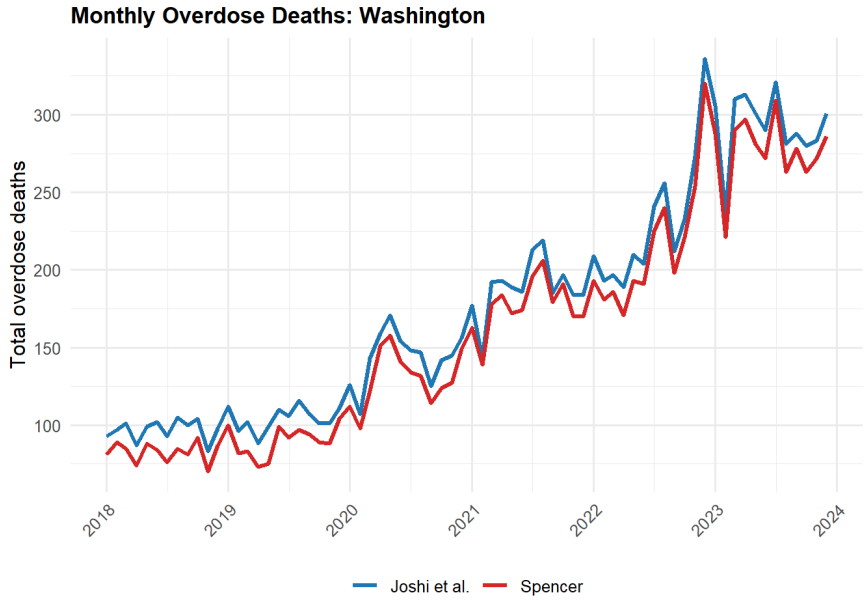
Notes: This figure plots national overdose deaths using alternative outcome definitions: all overdose deaths and unintentional overdose deaths. The all-overdose definition corresponds to the outcome used by Joshi et al. (2023), while the unintentional-overdose definition corresponds to the outcome used by Spencer (2023). While levels differ due to inclusion criteria, trends are highly similar over time, indicating that differences in definitions primarily affect scale rather than underlying dynamics.

Appendix Figure 2: Overdose deaths in Oregon and Washington using all-overdose and unintentional-overdose outcome definitions.

(a) Oregon



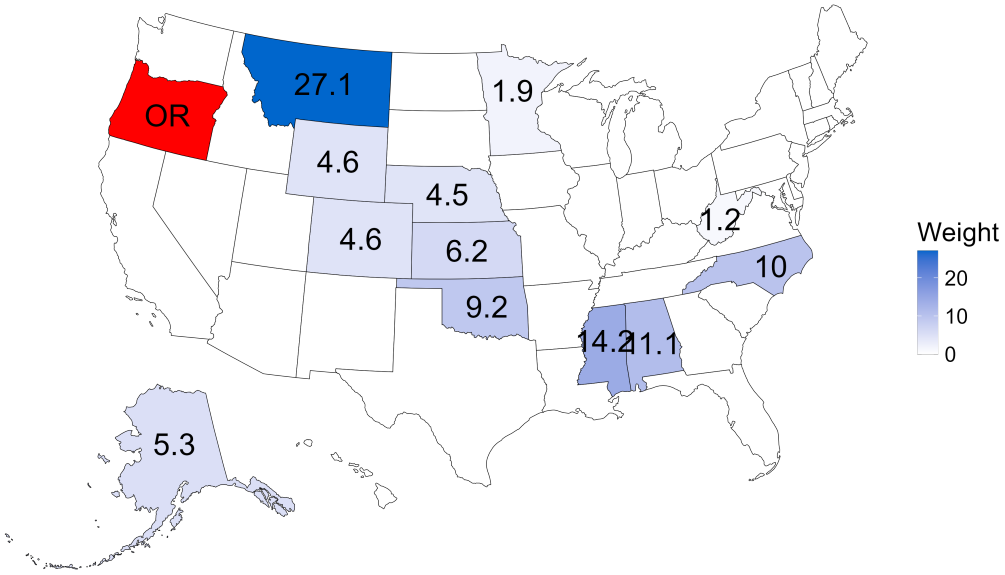
(b) Washington



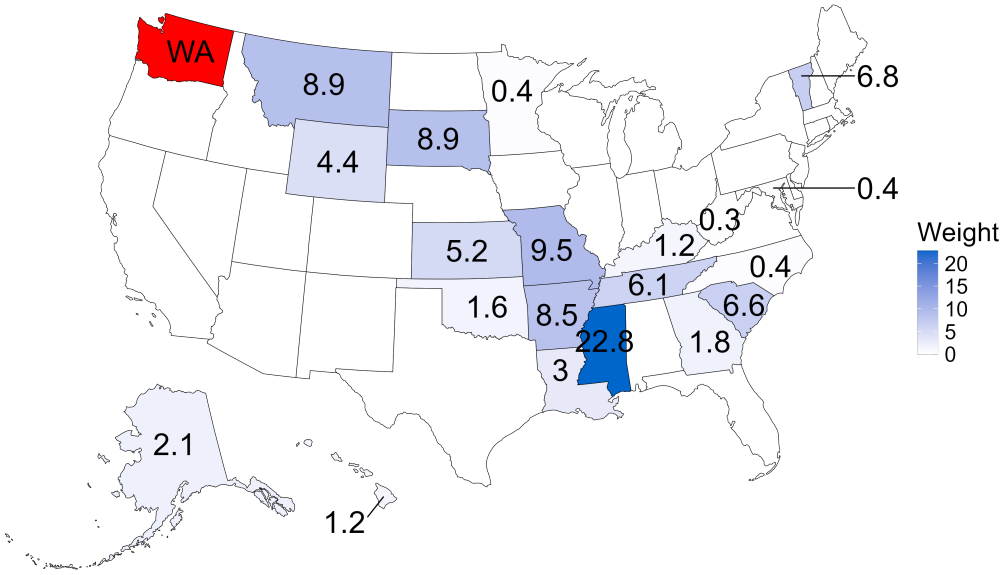
Notes: This figure plots overdose deaths in Oregon and Washington under alternative outcome definitions: all overdose deaths and unintentional overdose deaths. The all-overdose definition corresponds to the outcome used by Joshi et al. (2023), while the unintentional-overdose definition corresponds to the outcome used by Spencer (2023). While levels differ across definitions, trends are highly similar within each state, indicating that differences in outcome definitions primarily affect scale rather than the evolution of overdose mortality over time.

Appendix Figure 3: Replication Synthetic Control Donor Weights

(a) Oregon



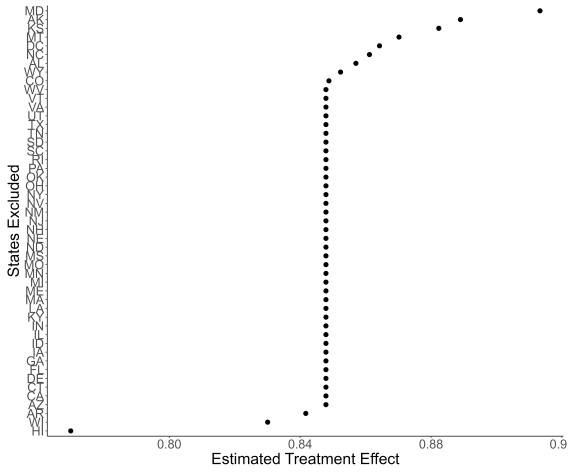
(b) Washington



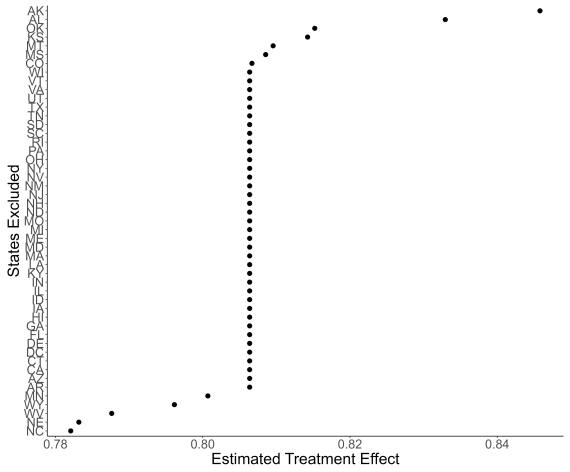
Notes: This figure maps the estimated donor weights used to construct synthetic Oregon and synthetic Washington in the all-overdose replication specification, corresponding to the specification used by Joshi et al. (2023).

Appendix Figure 4: Synthetic-Control Estimate: Dropping One Control Unit at a Time

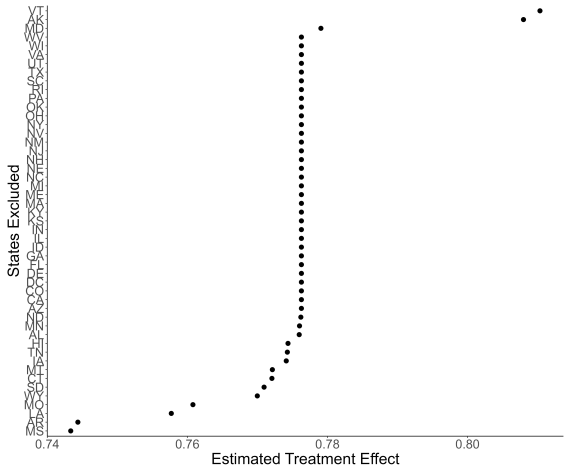
(a) Unintentional-overdose specification, Oregon



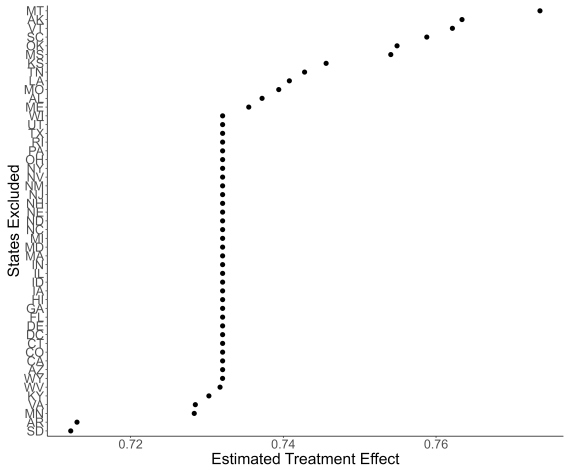
(b) All-overdose specification, Oregon



(c) Unintentional-overdose specification, Washington

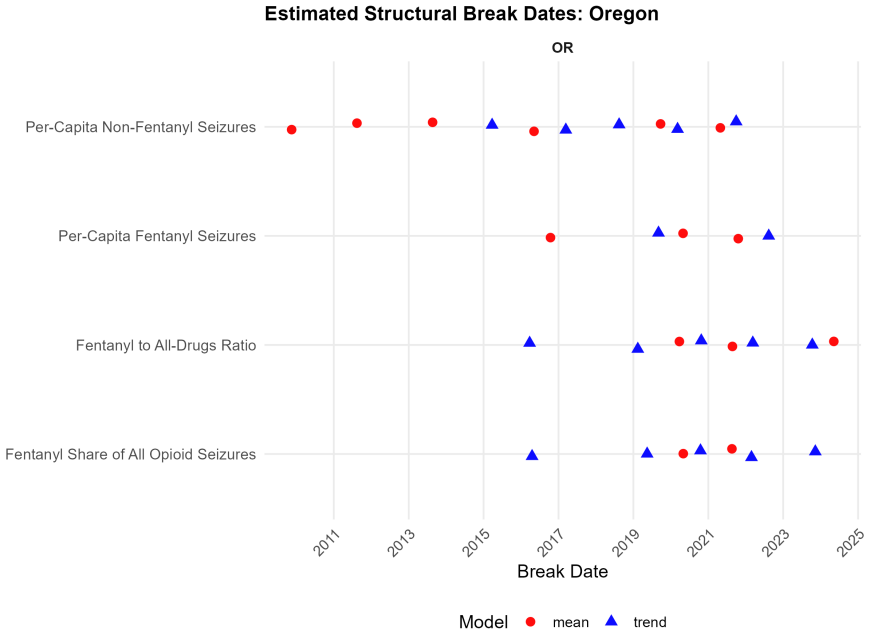


(d) All-overdose specification, Washington

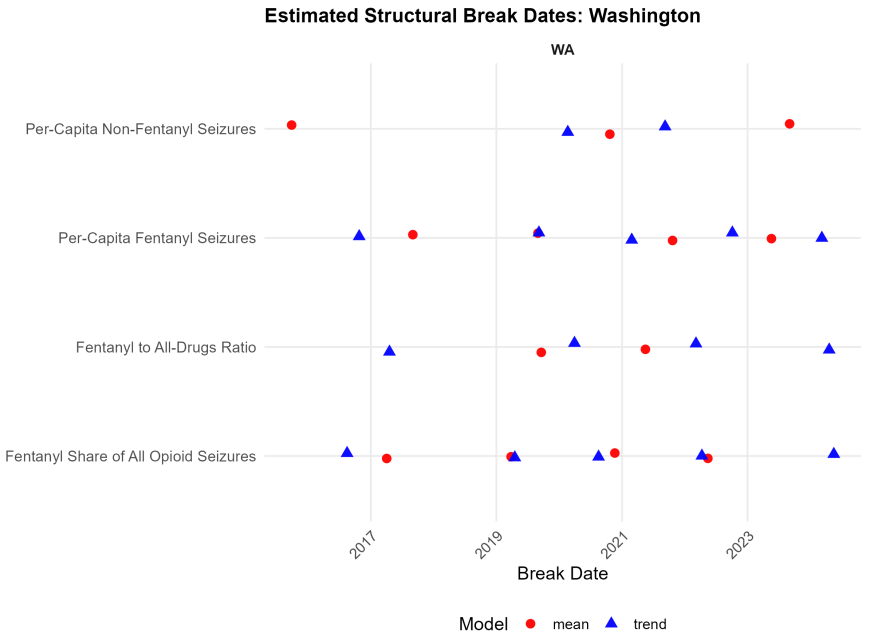


Notes: This figure reports synthetic control estimates obtained by iteratively excluding each donor state from the pool. The unintentional-overdose panels correspond to the outcome definition used by Spencer (2023), while the all-overdose panels correspond to the outcome definition used by Joshi et al. (2023). Results are stable across exclusions, indicating that no single donor state drives the estimated increase in overdose mortality.

Appendix Figure 5: Bai-Perron Multiple Break Tests for All Measures



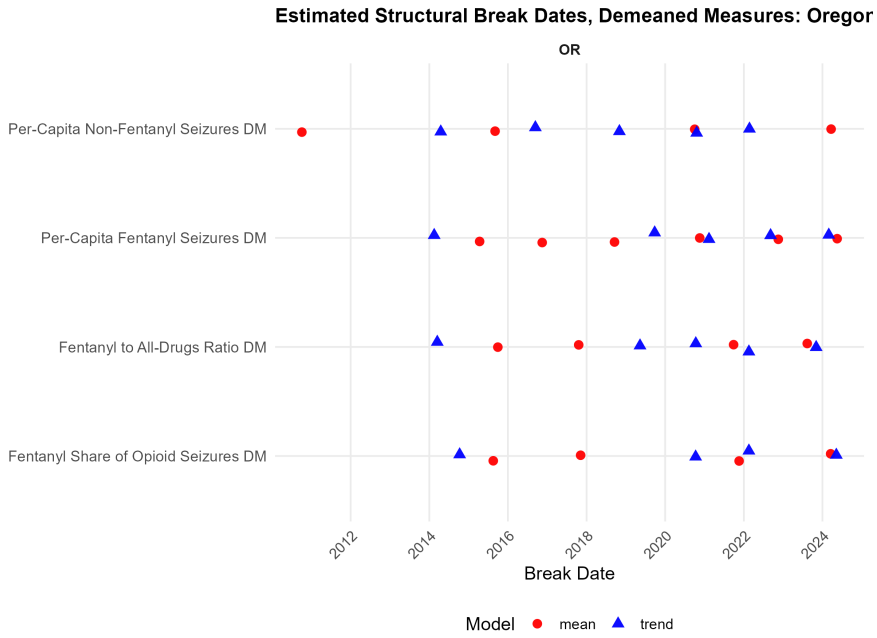
(a) Oregon



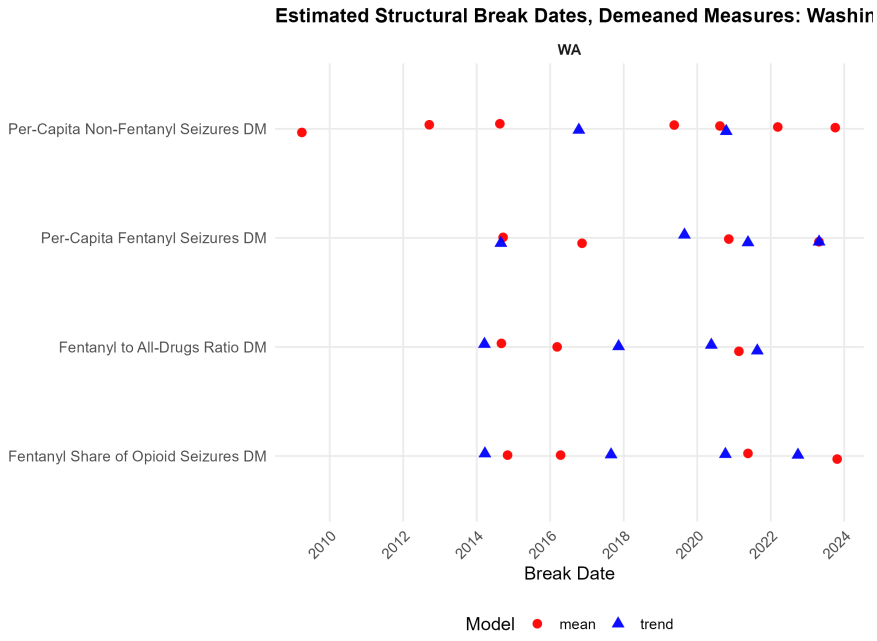
(b) Washington

Notes: This figure reports Bai–Perron multiple structural break tests for four state-level NFLIS report measures aggregated to half-year periods: fentanyl and fentanyl-related reports as a share of all drug reports, fentanyl and fentanyl-related reports per 100,000 population, non-fentanyl reports per 100,000 population, and fentanyl and fentanyl-related reports as a share of opioid-related reports. Opioid-related reports are defined as fentanyl and fentanyl-related compounds, narcotic analgesics, and literal heroin reports identified in NFLIS. Break dates vary across measures and specifications, indicating that the timing of structural changes is sensitive to how fentanyl exposure is measured.

Appendix Figure 6: Bai-Perron Multiple Break Tests for Demeaned Measures



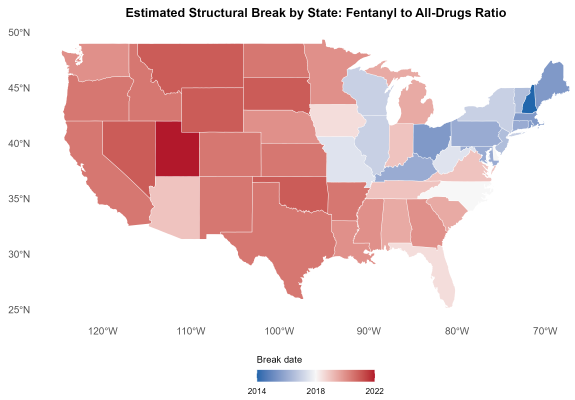
(a) Oregon



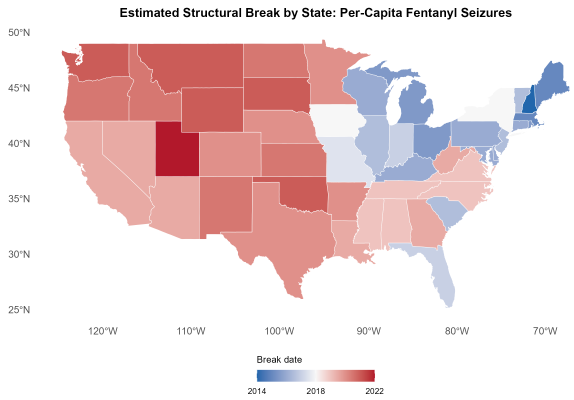
(b) Washington

Notes: This figure reports Bai–Perron multiple structural break tests for demeaned versions of the four NFLIS seizure measures shown in Appendix Figure 5. Each measure is demeaned by subtracting the national half-year mean of the same measure before estimating break dates. This adjustment removes common national shocks, including disruptions associated with the COVID-19 pandemic, so that estimated breaks reflect deviations in Oregon and Washington relative to contemporaneous national trends. The demeaned specification subtracts the contemporaneous unweighted mean across states from each state-half observation, so these estimates should be interpreted as removing common national period shocks, including pandemic-era disruptions, rather than as a population-weighted national adjustment.

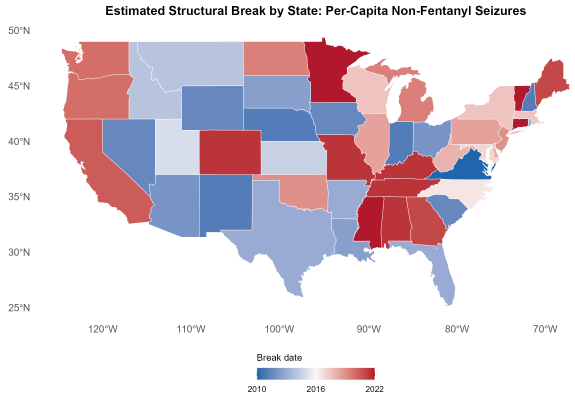
Appendix Figure 7: Estimated Structural Break Dates by State



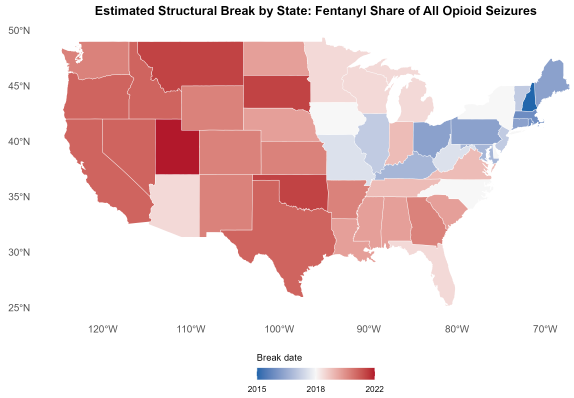
(a) Fentanyl share of all drug reports.



(b) Fentanyl reports per capita.



(c) Non-fentanyl reports per capita.

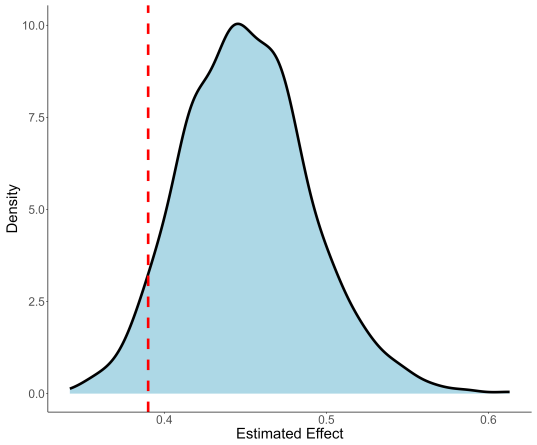


(d) Fentanyl share of opioid-related reports.

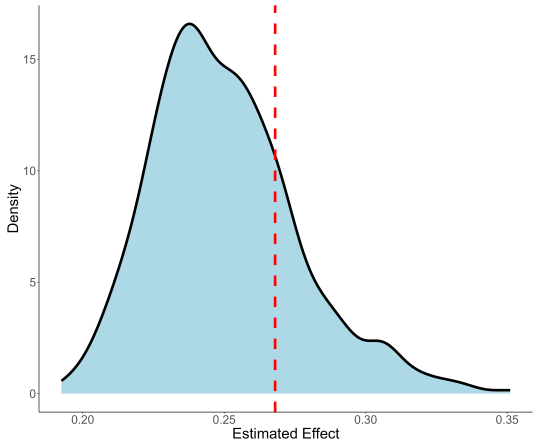
Notes: This figure maps estimated single structural break dates by state across alternative state-level NFLIS report measures aggregated to half-year periods. Opioid-related reports are defined as fentanyl and fentanyl-related compounds, narcotic analgesics, and heroin reports identified in NFLIS. Break dates vary substantially across measures and states, indicating that the geographic timing of fentanyl's spread is sensitive to measurement choice and not entirely consistent with a uniform east-to-west diffusion pattern.

Appendix Figure 8: Sensitivity to Random Imputation of Suppressed Overdose Counts

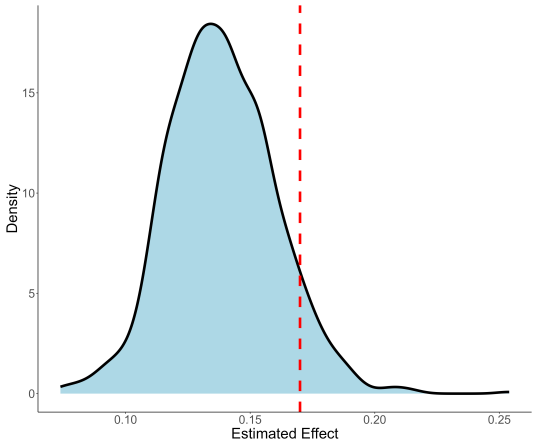
(a) Unintentional-overdose specification: Oregon



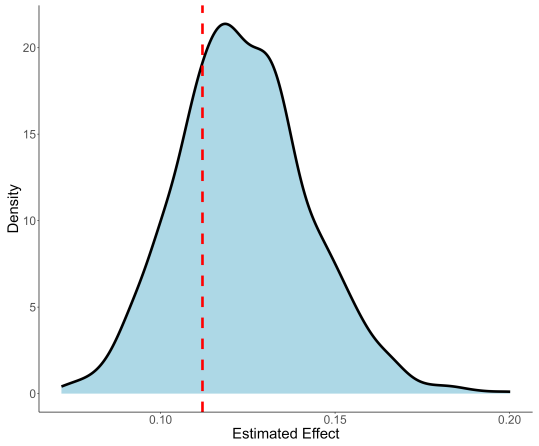
(b) All-overdose specification: Oregon



(c) Unintentional-overdose specification: Washington



(d) All-overdose specification: Washington



Notes: This figure presents simulated average treatment effects under 1,000 alternative imputations for state-month observations with suppressed counts in the restricted CDC data. The unintentional-overdose panels correspond to the outcome definition used by Spencer (2023), while the all-overdose panels correspond to the outcome definition used by Joshi et al. (2023). The distribution of estimates is centered near the reported point estimates, indicating that results are not driven by the imputation procedure used to address data suppression.

C Appendix Tables

Appendix Table 1: Replication: Synthetic-Control Estimate for Oregon

	Panel I: Unintentional Overdose Outcome			
ATT	0.434**	0.381**	0.403**	0.346**
p-value	{0.02}	{0.02}	{0.02}	{0.04}
	Panel II: All-Overdose Outcome			
ATT	0.329*	0.272	0.32	0.27
p-value	{0.08}	{0.16}	{0.16}	{0.26}
Sample	2018-21	2018-21	2018-22	2018-22
Matching Until	Dec	Jan	Dec	Jan

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: Synthetic control estimates for Oregon are shown under alternative specifications. Panel I uses unintentional overdose mortality (Spencer, 2023), while Panel II uses all overdose mortality (Joshi et al., 2023). Estimates vary with outcome definition, treatment timing, and sample window, with statistical significance sensitive to these choices. Notably, point estimates are similar in magnitude across specifications, suggesting that differences in statistical significance across studies are likely driven by limited statistical power in short post-treatment periods. Permutation-based p-values are reported in curly brackets.

Appendix Table 2: Replication: Synthetic-Control Estimate for Washington

	Panel I: Unintentional Overdose Outcome			
ATT	0.113**	0.155**	0.147**	0.184**
p-value	{0.02}	{0.02}	{0.02}	{0.02}
	Panel II: All-Overdose Outcome			
ATT	0.06**	0.103	0.087**	0.115*
p-value	{0.02}	{0.12}	{0.02}	{0.06}
Sample	2018-21	2018-21	2018-22	2018-22
Matching Until	Jan	Feb	Jan	Feb

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: Synthetic control estimates for Washington are shown under alternative specifications. Panel I uses unintentional overdose mortality following Spencer (2023), while Panel II uses all overdose mortality following Joshi et al. (2023). As with Oregon, estimates are sensitive to outcome definition and treatment timing, but point estimates are similar in magnitude across specifications. Statistical significance varies in early periods, suggesting that differences across studies are driven by limited statistical power rather than substantive disagreement. Permutation-based p-values are reported in curly brackets.

Appendix Table 3: Synthetic-Control Estimates with Prediction & Confidence Intervals, Replication

	Oregon		Washington	
	Panel I: Prediction Intervals			
ATT	0.406**	0.264	0.107*	0.098
	[0.13,0.68]	[-0.2,1.01]	[-0.01,0.23]	[-0.08,0.27]
	Panel II: Chernozhukov et al. T-Test Procedure			
ATT	0.429*	0.331	0.175	0.192
	[0.03,0.83]	[-0.01,0.67]	[-0.01,0.36]	[-0.1,0.48]
Outcome	Unintentional	All overdose	Unintentional	All overdose

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: This table reports synthetic control estimates with 90 percent prediction intervals and confidence intervals for the replication specifications. Intervals overlap across the unintentional-overdose and all-overdose specifications, suggesting that estimated effects are not statistically different across specifications despite differences in point estimates. Prediction intervals are reported in Panel I and confidence intervals in Panel II.

Appendix Table 4: Extension: Synthetic-Control Estimate, Overall & Yearly ATT

	All Overdose		Unintentional Overdose	
State	OR	WA	OR	WA
ATT Overall	0.806**	0.732**	0.848**	0.776**
p-value	{0.02}	{0.02}	{0.02}	{0.02}
ATT 2021	0.272	0.103	0.434**	0.113**
p-value	{0.16}	{0.12}	{0.02}	{0.02}
ATT 2022	0.582**	0.533**	0.644**	0.633**
p-value	{0.02}	{0.02}	{0.02}	{0.02}
ATT 2023	1.521**	1.455**	1.466**	1.528**
p-value	{0.02}	{0.02}	{0.02}	{0.02}

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: Synthetic control estimates are shown. Columns (1) and (2) use the all-overdose specification, which follows Joshi et al. (2023) in outcome definition and pre-treatment window, and columns (3) and (4) use the unintentional-overdose specification used by Spencer (2023). The p-values, calculated via a permutation test, are provided inside the curly brackets.

Appendix Table 5: Robustness: Alternative Estimators

	(1) Aug Synth	(2) SDiD	(3) GSynth	(4) Matrix Completion	(5) Multi-Outcome SC	(6) TWFE
			Panel A: Oregon			
ATT Overall	0.806**	0.81**	0.829***	0.838***	0.856*	0.883*
p-value	{0.021}	{0.04}	{0.000}	{0.000}	{0.065}	{0.094}
			Panel B: Washington			
ATT Overall	0.733***	0.859**	0.832***	0.836***	0.769***	0.888*
p-value	{0}	{0.036}	{0.000}	{0.000}	{0.004}	{0.093}

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: Column (1) reports results from augmented synthetic control; column (2) reports results from synthetic difference-in-differences method; column (3) reports results from generalized synthetic control method; column (4) reports results from matrix completion method; column (5) reports results from multiple outcome synthetic control; and column (6) reports results from two-way fixed effects estimator.

Appendix Table 6: Synthetic-Control Estimates Excluding the COVID-19 Period from Matching

	Oregon		Washington	
	(1)	(2)	(3)	(4)
ATT	0.878**	0.889**	0.809**	0.954**
p-value	{0.02}	{0.02}	{0.02}	{0.02}
Outcome	Unintentional	All overdose	Unintentional	All overdose

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: This table reports synthetic control estimates excluding the COVID-19 period from the pre-treatment window. Results remain positive and statistically significant across specifications, indicating that the estimated effects are not driven by pandemic-related changes in overdose trends. Permutation-based p-values are reported in curly brackets.

Appendix Table 7: Placebo Tests for Nearby Western States

	Unintentional-overdose outcome			All-overdose outcome		
	CA	ID	NV	CA	ID	NV
ATT	0.135	-0.089	0.418	0.176	-0.132	0.255
p-value	{0.408}	{0.796}	{0.367}	{0.327}	{0.776}	{0.429}

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: This table reports placebo synthetic control estimates for California, Idaho, and Nevada. Estimated effects are small and statistically insignificant, suggesting that the observed increases in overdose mortality in Oregon and Washington are not driven by broader regional trends. Permutation-based p-values are reported in curly brackets.

Appendix Table 8: Synthetic-Control Estimates for NFLIS Drug Reports

	Fentanyl	Non-fentanyl	Opioid	All reports	Fentanyl share	Opioid share
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel I: Oregon					
ATT	-15.048	-77.973**	-13.429**	-86.887**	17.95**	14.44*
p-value	{0.12}	{0.02}	{0.02}	{0.04}	{0.02}	{0.08}
	Panel II: Washington					
ATT	-19.451	-25.292**	-7.725**	-18.818**	15.045**	7.411
p-value	{0.16}	{0.02}	{0.02}	{0.04}	{0.04}	{0.24}

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: Synthetic control estimates are shown. All estimates use half-year data from 2010 to 2020 to find the synthetic weights. Columns (1)–(4) report changes in NFLIS reports per 100,000 residents; columns (5)–(6) report changes in fentanyl report shares. The p-values, calculated via a permutation test, are provided inside the curly brackets.

Appendix Table 9: Exploring Sensitivity to Alternative Fentanyl Controls – Starting Sample in 2008

	Oregon					Washington				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel I: Matrix Completion SC Using Half-Year Data, 2008-2022									
ATT	4.747***	1.672	6.138***	2.638**	5.814***	4.249***	0.321	5.55***	1.645	5.633***
p-value	{0.000}	{0.116}	{0.000}	{0.013}	{0.000}	{0.000}	{0.800}	{0.000}	{0.140}	{0.000}
	Panel II: Matrix Completion SC Using Half-Year Data, 2008-2023									
ATT	9.837***	5.996***	10.608***	7.009***	10.145***	9.257***	4.98***	9.958***	6.254***	9.724***
p-value	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}	{0.000}
Fentanyl Control	None	Drug share	PC	Opioid share	None	None	Drug share	PC	Opioid share	None
Donor Pool Restriction?	No	No	No	No	Yes	No	No	No	No	Yes

* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

Notes: Panel I estimates Zoorob et al. (2024)'s synthetic control method, which uses half-year data from 2008 to 2022. Panel II estimates Zoorob et al. (2024)'s synthetic control method, but with an additional year of post-treatment data. The ATT is multiplied by 2 to report an average effect over the year. In columns (2) and (7), we include a control for the ratio of fentanyl reports to all drug reports, which is the covariate used in Zoorob et al. (2024). In columns (3) and (8), we use fentanyl reports per capita as our covariate, and in columns (4) and (9), we use the ratio of fentanyl reports to opioid-related reports as our covariate. In columns (5) and (10), we restrict our donor pool to states where the fentanyl crisis arrived in 2019 or later, as defined by Zoorob et al. (2024). The p-values generated using the jackknife method (panels I and II) are reported inside the curly brackets.