

Recreational Marijuana Legalization and Business Formation: Evidence from Staggered State Adoption

APEP Autonomous Research*

@ai1scl

January 29, 2026

Abstract

Does recreational marijuana legalization affect entrepreneurial activity? I exploit the staggered opening of recreational marijuana retail sales across 21 U.S. states from 2014 to 2024 to estimate the causal effect on new business formation using the Census Bureau’s Business Formation Statistics. Using the heterogeneity-robust [Callaway and Sant’Anna \(2021\)](#) estimator as the primary specification, I find a Callaway-Sant’Anna overall ATT of -0.028 log points (SE = 0.029, 95% CI: $[-0.085, 0.029]$), indicating a modest but imprecisely estimated decline in business applications per capita. Conventional TWFE estimates are larger (-0.068 log points, 95% CI: $[-0.148, 0.012]$), consistent with heterogeneity bias under staggered adoption. Event-study estimates reveal no evidence of differential pre-trends and suggest effects emerge gradually, reaching -0.15 to -0.20 log points by 6–7 years after retail opening. A descriptive analysis of actual business formations (BF8Q) shows a *positive* association, though this cannot be interpreted causally because BF8Q is forward-looking and mechanically spans post-treatment periods for pre-treatment cohorts. To address potential spillover contamination of the control group, I show that results are robust when restricting controls to “interior” states that do not border any treated state (ATT = -0.042 , SE = 0.034). Results are further robust to randomization inference ($p = 0.093$), pairs cluster bootstrap ($p = 0.064$), medical-marijuana-only controls, and excluding COVID-era observations.

JEL Codes: L26, I18, K32, R11

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

Keywords: marijuana legalization, business formation, entrepreneurship, difference-in-differences, cannabis policy

1. Introduction

The legalization of recreational marijuana represents one of the most significant regulatory shifts in the United States over the past decade. Beginning with Colorado and Washington in 2012, 24 states and the District of Columbia have now legalized recreational marijuana for adult use, creating a legal cannabis industry that generated over \$30 billion in sales in 2023 alone ([Marijuana Policy Project, 2023](#)). While a substantial and growing literature examines the public health, criminal justice, and labor market consequences of marijuana legalization, far less attention has been paid to its effects on entrepreneurial activity and business dynamism more broadly ([Anderson and Rees, 2023](#)).

This gap is surprising. Recreational marijuana legalization creates substantial new economic opportunities—not only in cannabis cultivation, processing, and retail, but also in ancillary industries such as agriculture supply, professional services (legal, accounting, consulting), real estate, security, packaging, and tourism ([Dills et al., 2021](#)). At the same time, legalization may reduce business formation through several channels: regulatory uncertainty about federal enforcement, banking restrictions facing cannabis-adjacent firms, increased operating costs from compliance regimes, or substitution effects if marijuana use reduces entrepreneurial motivation ([Sabia and Nguyen, 2017](#)). The net effect on overall business formation is theoretically ambiguous and ultimately an empirical question.

This paper estimates the causal effect of recreational marijuana legalization on new business formation using a staggered difference-in-differences design. I exploit variation in the timing of first legal recreational retail sales across 21 states between 2014 and 2024, comparing business formation dynamics in states that opened recreational retail to those that did not open retail during the sample period. My primary data come from the Census Bureau’s Business Formation Statistics (BFS), which provides monthly counts of Employer Identification Number (EIN) applications—the earliest observable signal of new business creation intent—for all 50 states and the District of Columbia from 2005 to 2024.

I define treatment as the year a state first permits legal recreational marijuana retail sales, rather than the year of statutory legalization. This distinction matters because the lag between legalization and retail opening ranges from zero (Illinois) to over three years (Maine, Vermont), and the economic effects of legalization plausibly intensify when commercial activity actually begins. This approach follows the logic of [Brown et al. \(2023\)](#), who emphasize the importance of distinguishing between legal status and market access.

The main analysis proceeds in three steps. First, recognizing that standard TWFE estimates may be biased under treatment effect heterogeneity with staggered adoption ([Goodman-Bacon, 2021](#); [de Chaisemartin and D’Haultfoeuille, 2020](#); [Sun and Abraham,](#)

2021; Baker et al., 2022), I implement the Callaway and Sant’Anna (2021) heterogeneity-robust estimator as my primary specification. The Callaway-Sant’Anna overall ATT is -0.028 log points (SE = 0.029), indicating a modest but imprecisely estimated decline in business applications per capita. For a median-sized state with approximately 45,000 annual applications, this translates to roughly 1,250 fewer applications per year. As a benchmark, I also report conventional TWFE estimates of approximately -0.068 log points (about 7 percent, or 3,000 fewer applications for the median state), marginally significant at the 10 percent level; the discrepancy between the two estimators is consistent with heterogeneity bias in the TWFE specification. Third, I decompose the effect across different BFS series—high-propensity business applications, applications with planned wages, corporate applications, and actual business formations—revealing an important compositional pattern.

A suggestive finding is the divergence between business *applications* and business *formations*. While total applications decline modestly (point estimates ranging from -0.03 to -0.07), a TWFE regression using actual business formations within 8 quarters (BF8Q) as the outcome yields a positive coefficient of 0.030 (SE = 0.018). However, this BF8Q estimate cannot be interpreted causally under the DiD design because BF8Q is a forward-looking measure: it counts whether applications from year t become employers within 8 quarters (t to $t + 2$), meaning pre-treatment application cohorts mechanically incorporate post-treatment formation behavior. I present this result as descriptive evidence of a potential compositional shift, not as a causal estimate. The decline in applications is concentrated in total and high-propensity categories, while applications with planned wages—indicating intent to hire employees—show essentially no change.

Event-study estimates from the Callaway-Sant’Anna dynamic aggregation reveal no statistically significant pre-trends in the seven years before treatment, supporting the parallel trends assumption. Post-treatment, effects emerge gradually, with point estimates near zero in the first two years before becoming increasingly negative, reaching approximately -0.15 to -0.20 log points by 6–7 years after retail opening. This monotonically declining pattern suggests that the negative effect on business applications intensifies as the cannabis regulatory environment matures.

I conduct extensive robustness checks. Randomization inference based on 999 permutations of treatment assignment yields a two-sided p -value of 0.093 for the TWFE coefficient, confirming that the observed negative relationship is unlikely to arise by chance alone. A pairs cluster bootstrap with 999 replications produces a confidence interval of $[-0.151, 0.001]$ and a p -value of 0.064. To address concerns about spillover contamination of the control group (Bertrand et al., 2004), I estimate specifications that restrict the comparison group to “interior” never-treated states that do not share a border with any treated state; the resulting

CS ATT of -0.042 ($SE = 0.034$) is similar in sign and magnitude to the baseline estimate. Results are also robust to restricting the control group to medical-marijuana-only states and to excluding COVID-affected years. Heterogeneity analysis reveals that early-adopting states (Colorado, Washington, Oregon, Alaska, Nevada) show larger negative effects on applications ($\beta = -0.103$) compared to later adopters ($\beta = -0.046$), consistent with the pioneer disadvantage of operating under maximal federal enforcement uncertainty.

This paper makes several contributions to the literature. First, it provides the first comprehensive analysis of recreational marijuana legalization’s effect on overall business formation using the Census Bureau’s BFS data, which captures the universe of EIN applications rather than the selected sample of firms that appear in business registries or tax records. The only prior related work is [Brown et al. \(2023\)](#), who use the Business Dynamics Statistics (BDS) to examine establishment births in 12 states through 2020; I extend the analysis to 21 states through 2024 using a more upstream measure of entrepreneurial intent.

Second, the decomposition across BFS series types represents a methodological contribution. By separately examining total applications, high-propensity applications, wage-planned applications, and corporate applications, I characterize how legalization affects different margins of entrepreneurial activity. I also present descriptive evidence on business formations (BF8Q), though the forward-looking nature of this measure precludes a causal interpretation under the DiD design. The analysis has implications for how we interpret business application data as a measure of economic dynamism ([Decker et al., 2014](#); [Haltiwanger et al., 2013](#)).

Third, I contribute to the broader econometric literature on difference-in-differences by implementing multiple modern estimators and inference procedures. The combination of TWFE, Callaway-Sant’Anna, randomization inference, and cluster bootstrap provides a comprehensive assessment of both the magnitude and statistical reliability of the estimated effects, addressing concerns about inference with a moderate number of treated clusters ([Conley and Taber, 2011](#); [MacKinnon et al., 2023](#)).

The remainder of the paper is organized as follows. Section 2 describes the institutional background of recreational marijuana legalization. Section 3 presents a conceptual framework for understanding the channels through which legalization may affect business formation. Section 4 describes the data. Section 5 details the empirical strategy. Section 6 presents results. Section 7 discusses implications and limitations. Section 8 concludes.

2. Institutional Background and Policy Setting

2.1 The Legalization Wave

The movement to legalize recreational marijuana in the United States began with ballot initiatives in Colorado and Washington in November 2012. Both states approved recreational use for adults aged 21 and older, though the regulatory frameworks for commercial sales were not immediately established. Colorado opened its first recreational dispensaries on January 1, 2014, while Washington followed in July 2014. Since then, legalization has proceeded in waves, with 19 additional states opening recreational retail markets by 2024.

The pace of adoption has been uneven. Oregon (2015) and Alaska (2016) were early followers, while Nevada (2017) and California (2018) represented the first large-market entries. A second wave included Illinois (2020), Arizona (2021), and several northeastern states (New Jersey, New York, Connecticut) beginning in 2022–2023. Most recently, Ohio opened retail sales in 2024. Several states have legalized possession but have not yet established retail markets (Virginia, Delaware, Minnesota), and these states are excluded from the treatment group in my analysis.

2.2 The Lag Between Legalization and Retail Sales

A critical institutional feature is the substantial variation in the lag between statutory legalization and the opening of retail markets. This lag arises from the complex regulatory infrastructure required for commercial cannabis—including licensing systems, testing laboratories, seed-to-sale tracking, local zoning approvals, and tax administration. Table 1 presents the timeline for each treated state.

The lag ranges from zero days (Illinois, which synchronized its legalization and retail launch on January 1, 2020) to over three years (Maine legalized in 2017 but did not open retail sales until October 2020; Vermont legalized possession in 2018 but did not permit retail until October 2022). This variation is important for identification because it means that the economic effects of legalization—particularly those operating through commercial channels—are likely to be concentrated around the retail opening date rather than the statutory legalization date.

I define treatment timing based on first retail sales rather than legalization for three reasons. First, business formation responses should primarily reflect the creation of cannabis-related businesses and ancillary services, which require a legal commercial market. Second, the regulatory uncertainty that might deter non-cannabis business formation is likely most acute when retail operations begin and enforcement questions arise about banking, taxation,

and federal compliance. Third, using retail timing reduces measurement error from the highly variable legalization-to-retail lag.

2.3 Regulatory Environment and Business Implications

Recreational marijuana legalization creates a distinctive regulatory environment with several features relevant to business formation. First, the federal-state conflict creates ongoing uncertainty. Marijuana remains a Schedule I controlled substance under federal law, and businesses operating in the cannabis space face restrictions on banking services, tax deductions (Section 280E of the Internal Revenue Code), and interstate commerce ([Kilmer et al., 2020](#)). These frictions may deter some entrepreneurs from entering not only the cannabis market but also adjacent industries where the regulatory status is ambiguous.

Second, state licensing regimes vary substantially in their restrictiveness. Some states (like Oregon) adopted relatively open licensing systems that facilitated entry, while others (like New York and New Jersey) implemented equity-focused licensing programs with caps on the number of licenses and priority for social equity applicants. These differences in market structure may generate heterogeneous effects on business formation.

Third, local option provisions in most states allow municipalities to ban or restrict marijuana businesses, creating within-state variation in commercial access. This local heterogeneity means that even in “treated” states, the actual commercial footprint of recreational marijuana varies considerably.

Fourth, the tax treatment of marijuana businesses differs across states, with combined state and local tax rates ranging from approximately 15 percent (Oregon) to over 35 percent (Washington). High tax rates may reduce the profitability of cannabis ventures while generating revenue that state and local governments can invest in economic development.

2.4 Prior Literature

The economics literature on marijuana legalization has grown rapidly. [Anderson and Rees \(2023\)](#) provide a comprehensive review, documenting effects on crime, traffic fatalities, substance use, and labor market outcomes. Several studies find that medical marijuana laws increase economic activity in specific sectors ([Gavrilova et al., 2019](#); [Dills et al., 2021](#)), though evidence on recreational legalization is more limited.

On business dynamics specifically, [Brown et al. \(2023\)](#) use the Business Dynamics Statistics to examine how recreational legalization affects establishment births, finding positive effects concentrated in retail and accommodation sectors. However, their analysis covers only 12 states through 2020 and relies on BDS data that measure establishment births with a

multi-year lag. My analysis extends the sample to 21 states, uses BFS data that captures the upstream signal of business *intent* (EIN applications), and provides a more granular decomposition of effects across application types.

Other related work includes [Hansen et al. \(2020\)](#) on cannabis taxation in Washington State, [Chakraborty et al. \(2023\)](#) on cannabis legalization and housing markets, [Sabia and Nguyen \(2017\)](#) on labor market effects of medical marijuana, and [Nicholas and Maclean \(2022\)](#) on the effects of dispensary openings on local economic activity. The broader entrepreneurship literature emphasizes the importance of regulatory environment for business formation decisions ([Djankov et al., 2002](#); [Klapper et al., 2006](#)), and recreational marijuana represents a particularly salient regulatory change with both direct and indirect effects on the business environment.

3. Conceptual Framework

Recreational marijuana legalization can affect business formation through multiple channels that operate in opposing directions. I organize these into direct effects from creating a new legal industry, indirect effects from changing the regulatory and economic environment, and compositional effects that alter the mix of applications. This framework generates testable implications for the BFS series decomposition.

3.1 Direct Effects: Cannabis Industry Creation

The most straightforward channel is the creation of an entirely new legal industry. Legalization generates demand for businesses in cannabis cultivation, processing, testing, distribution, and retail, each requiring a new EIN application. Additionally, ancillary businesses emerge in specialized legal services, compliance consulting, packaging, security, commercial real estate, and equipment supply. These direct effects unambiguously predict an *increase* in business formation, with the magnitude depending on the number of licenses issued, the speed of market development, and the extent of ancillary industry growth. However, in any given state, the number of cannabis-related business applications is likely small relative to total EIN applications—Colorado issued approximately 2,900 cannabis business licenses in its first five years, compared to over 80,000 annual total business applications—suggesting that the direct positive effect may be difficult to detect in aggregate data.

3.2 Indirect Effects: Regulatory Environment

Legalization may also affect business formation indirectly by changing the regulatory and economic environment. Several mechanisms could reduce non-cannabis business formation.

The ongoing federal prohibition creates legal risk for businesses that interact with cannabis companies—including landlords, vendors, and financial institutions—and this uncertainty may deter risk-averse entrepreneurs from starting businesses in states with active marijuana markets. Federal money laundering statutes make banks reluctant to serve cannabis businesses, and this reluctance can extend to businesses in adjacent industries or commercial districts near dispensaries, raising the cost of entry (Kilmer et al., 2020). If the cannabis industry offers attractive wages, it may draw workers from other industries, increasing labor costs for new entrants. Finally, the addition of marijuana regulations to an already complex business environment may increase the administrative burden of starting any business.

Conversely, legalization could stimulate non-cannabis business formation through several positive channels. Marijuana tax revenue may fund public infrastructure, education, or business development programs that facilitate entry (Hansen et al., 2020). Legal marijuana may increase consumer spending in complementary sectors (food services, entertainment, tourism), creating demand that attracts new businesses. More broadly, legalization may signal a state’s willingness to adopt novel policies, attracting innovative entrepreneurs.

3.3 Testable Implications

The net effect on the *number* of business applications is theoretically ambiguous, but the framework generates predictions that can be tested with the BFS data. If the positive direct effect dominates, total applications (BA) should increase, with particularly strong increases in high-propensity applications (HBA) and those with planned wages (WBA), since cannabis businesses are typically capital-intensive operations that plan to hire employees. If the negative indirect effects dominate, total applications should decline, with the decline concentrated in non-cannabis, speculative applications that are deterred by regulatory uncertainty. If both effects operate simultaneously, a compositional shift would emerge: total applications could decline while the *formation rate* (the share of applications that become actual businesses) rises, producing a pattern where BA falls but BF8Q increases. This “quality upgrade” hypothesis—that legalization filters out marginal applicants while attracting serious ventures—implies that the gap between BFS application and formation measures is informative about the underlying mechanism, even if the BF8Q measure itself cannot support a causal interpretation under the DiD design (see Section 6.5). Finally, the framework predicts that negative effects should be larger for early-adopting states that face greater federal enforcement uncertainty and lack the benefit of established regulatory templates.

4. Data

4.1 Business Formation Statistics

The primary outcome data come from the U.S. Census Bureau’s Business Formation Statistics (BFS), which provides monthly counts of new business applications by state ([U.S. Census Bureau, 2024](#)). The BFS constructs these data from IRS Employer Identification Number (EIN) applications, capturing the universe of new business entities that apply for a federal tax identification number. Unlike business registration data, which vary by state in coverage and timing, EIN applications provide a nationally consistent measure of business formation intent.

The BFS reports several series that decompose total business applications by characteristics of the applicant:

- **Total Business Applications (BA):** All new EIN applications.
- **High-Propensity Business Applications (HBA):** Applications with characteristics indicating a high likelihood of becoming an employer business with payroll (based on planned wages, corporate form, NAICS industry, and other features).
- **Business Applications with Planned Wages (WBA):** Applications where the applicant indicates an expected first wage payment date.
- **Corporate Business Applications (CBA):** Applications filed as corporations, S-corporations, or LLCs electing corporate tax treatment.
- **Business Formations within 8 Quarters (BF8Q):** The count of applications that result in an actual business with payroll tax filings within 8 quarters of the EIN application. Importantly, BF8Q is indexed by the application date but measures outcomes over the subsequent two years, making it a forward-looking variable whose measurement window may span treatment dates (see Section 6.5 for implications).

I aggregate the monthly data to the annual level for each state, requiring 12 months of observations per state-year. The BFS monthly data file was accessed in January 2026 and contains complete (all 12 months) data through December 2024 for all application-based series. The analysis sample spans 2005–2024, covering 51 geographic units (50 states plus the District of Columbia). The total state-year panel contains 1,020 observations (51×20 years) for the application series.

Business formations (BF8Q) have more limited availability. Because BF8Q requires matching EIN applications with quarterly tax filings over the subsequent 8 quarters, and the

Census Bureau releases these matched data with a multi-year lag, the BF8Q series in the downloaded file covers only 2005–2020. Moreover, not all states have complete (12-month) BF8Q data in every year; the number of states with valid BF8Q observations ranges from 36 to 43 per year, yielding an *unbalanced* panel of 624 observations across 44 states and 16 years. One singleton observation is removed during TWFE estimation, leaving $N = 623$ for the BF8Q regressions. Only treatment cohorts through 2020 (Colorado through Maine/Illinois) have post-treatment BF8Q observations; later cohorts contribute no post-treatment data to the BF8Q analysis.

4.2 Treatment Timing

Treatment timing is defined as the year of first legal recreational marijuana retail sales. I construct this variable from a combination of state legislative records, news reports, and the Marijuana Policy Project’s legalization tracker ([Marijuana Policy Project, 2023](#)). Table 1 presents the complete treatment timeline.

The 21 treated states represent a diverse cross-section of the country, including early pioneers (Colorado, Washington), large markets (California, Illinois, New York), and more recent adopters (Ohio, Missouri, Maryland). The 30 never-treated units (29 states plus the District of Columbia) serve as the comparison group. States that have legalized recreational marijuana but have not yet opened retail markets (Virginia, Delaware, Minnesota) are coded as never-treated.

4.3 Population Data

State-level population data come from the American Community Survey (ACS) 1-year estimates for 2005–2023, accessed via the Census Bureau API. I use population counts to normalize business applications per 100,000 residents, addressing the mechanical correlation between state size and application volume. Population for 2020 is linearly interpolated between 2019 and 2021 values because the ACS 1-year estimates were not released for 2020 due to COVID-19 pandemic disruptions of survey operations. For 2024, I use the 2023 population estimate as an extrapolation. With this interpolation, population data are available for all years 2005–2024, yielding $N = 1,020$ state-year observations for all application-based specifications.

4.4 Medical Marijuana Timing

To control for the prior legalization of medical marijuana, I construct a state-year indicator for active medical marijuana programs using data from the National Conference of State

Table 1: Recreational Marijuana Legalization: Treatment Timing

State	Legalization	First Retail Sales	Cohort
CO	2012-12-10	2014-01-01	2014
WA	2012-12-06	2014-07-08	2014
OR	2015-07-01	2015-10-01	2015
AK	2015-02-24	2016-10-29	2016
NV	2017-01-01	2017-07-01	2017
CA	2016-11-09	2018-01-01	2018
MA	2016-12-15	2018-11-20	2018
MI	2018-12-06	2019-12-01	2019
IL	2020-01-01	2020-01-01	2020
ME	2017-01-31	2020-10-09	2020
AZ	2020-11-30	2021-01-22	2021
MT	2021-01-01	2022-01-01	2022
NJ	2021-02-22	2022-04-21	2022
NM	2021-06-29	2022-04-01	2022
NY	2021-03-31	2022-12-29	2022
CT	2021-07-01	2023-01-10	2023
RI	2022-05-25	2022-12-01	2022
VT	2018-07-01	2022-10-01	2022
MD	2023-07-01	2023-07-01	2023
MO	2022-12-08	2023-02-03	2023
OH	2023-12-07	2024-08-06	2024

Notes: Treatment timing based on first legal recreational retail sales. States without retail sales (VA, DE, MN) are coded as never-treated controls. Legalization date is when recreational possession became legal under state law.

Legislatures (NCSL) and the Marijuana Policy Project. Thirty-nine states had active medical marijuana programs by 2024. I include a binary indicator for whether a state had a medical marijuana program in effect in a given year as a control variable in some specifications.

4.5 Summary Statistics

Table 2 presents summary statistics for the state-year panel, separately for retail-opening and no-retail states.

Table 2: Summary Statistics

	Opened Retail		No Retail Sales	
	Mean	SD	Mean	SD
Business Applications	73,979	83,601	59,840	89,351
Applications per 100k	999.1	344.6	1107.2	850.3
Population (millions)	7.44	8.32	5.39	5.83
States	21		30	
State-year observations	420		600	

Notes: Summary statistics for the state-year panel, 2005–2024. Business applications are total annual EIN applications from Census Business Formation Statistics. “Opened Retail” includes states that opened recreational retail sales by 2024 (treated group). “No Retail Sales” includes states without recreational retail sales by 2024, including states that legalized possession but have not yet opened retail markets (VA, DE, MN).

No-retail states have somewhat higher average business applications per capita than retail-opening states, though there is substantial overlap in the distributions. The treated group includes both large states (California, New York, Illinois) and small states (Alaska, Vermont), providing variation in state size. The panel is balanced across years (2005–2024) for the application series.

5. Empirical Strategy

5.1 Identification

The identification strategy exploits the staggered timing of recreational marijuana retail openings across states. The key assumption is that, absent legalization, business formation trends in treated states would have evolved in parallel with those in never-treated states. Formally, I require:

$$\mathbb{E}[Y_{it}(0) - Y_{it-1}(0)|G_i = g] = \mathbb{E}[Y_{it}(0) - Y_{it-1}(0)|G_i = \infty] \quad \forall t, g \quad (1)$$

where $Y_{it}(0)$ denotes the potential outcome without treatment, G_i is the cohort (year of first retail sales) for state i , and $G_i = \infty$ denotes never-treated states. This assumption requires that time trends in business formation are common across states, conditional on state and year fixed effects.

I assess the plausibility of this assumption through event-study analysis, examining whether treatment and control states exhibit parallel trends in the pre-treatment period. As emphasized by [Roth \(2022\)](#), failing to reject the null of equal pre-trends does not establish that parallel trends hold—particularly in settings with limited pre-treatment power. The event-study estimates provide a direct test of differential pre-trends and reveal the dynamic pattern of effects following treatment, but I caution that the tests have limited power given the moderate number of treated clusters and annual aggregation. I also acknowledge that a formal sensitivity analysis along the lines of [Rambachan and Roth \(2023\)](#)—which would characterize how much pre-trend violations could affect the results—would strengthen the analysis; I leave this for future work.

5.2 TWFE Estimation (Benchmark)

As a benchmark for comparison with the heterogeneity-robust estimator, I estimate a standard two-way fixed effects model:

$$\log(Y_{it}/\text{Pop}_{it}) = \alpha_i + \gamma_t + \beta \cdot \text{Treated}_{it} + \varepsilon_{it} \quad (2)$$

where Y_{it} is the number of business applications in state i and year t , Pop_{it} is the state population, α_i are state fixed effects, γ_t are year fixed effects, and Treated_{it} is an indicator equal to one if state i has opened recreational retail sales by year t . Standard errors are clustered at the state level to account for within-state serial correlation.

I also estimate specifications that control for medical marijuana status:

$$\log(Y_{it}/\text{Pop}_{it}) = \alpha_i + \gamma_t + \beta \cdot \text{Treated}_{it} + \delta \cdot \text{Medical}_{it} + \varepsilon_{it} \quad (3)$$

where Medical_{it} indicates whether state i has an active medical marijuana program in year t .

5.3 Callaway-Sant’Anna Estimation (Primary)

Standard TWFE estimators can produce biased estimates of average treatment effects under treatment effect heterogeneity with staggered adoption, as demonstrated by [Goodman-Bacon \(2021\)](#), [de Chaisemartin and D’Haultfœuille \(2020\)](#), [Borusyak et al. \(2024\)](#), and [Baker et al. \(2022\)](#). Because treatment timing varies substantially across cohorts—from 2014

(Colorado/Washington) to 2024 (Ohio)—and treatment effects likely differ across cohorts, I use the [Callaway and Sant’Anna \(2021\)](#) estimator as my primary specification. This estimator avoids problematic comparisons between differently-timed treatment groups by estimating separate group-time ATTs and aggregating them with appropriate weights.

The Callaway-Sant’Anna approach estimates group-time average treatment effects:

$$ATT(g, t) = \mathbb{E}[Y_t - Y_{g-1}|G = g] - \mathbb{E}[Y_t - Y_{g-1}|G = \infty] \quad (4)$$

for each cohort g and post-treatment period $t \geq g$. I use never-treated states as the comparison group and the doubly robust estimator ([Sant’Anna and Zhao, 2020](#)), which combines outcome regression and inverse probability weighting for additional robustness. I set the base period to “universal” (all pre-treatment periods) and estimate group-time ATTs with bootstrapped standard errors (999 replications).

These group-time ATTs are aggregated in several ways: (i) an overall ATT (weighted average across all groups and post-treatment periods), (ii) dynamic/event-study estimates (aggregated by time relative to treatment), and (iii) group-specific ATTs (one estimate per cohort).

5.4 BFS Series Decomposition

A key feature of the BFS data is the availability of multiple series that capture different aspects of business formation. Rather than a triple-difference design (which would require state-level NAICS sector breakdowns not available in the BFS), I estimate the treatment effect separately for each BFS series:

$$\log(Y_{it}^s / \text{Pop}_{it}) = \alpha_i^s + \gamma_t^s + \beta^s \cdot \text{Treated}_{it} + \varepsilon_{it}^s \quad (5)$$

where $s \in \{\text{BA, HBA, WBA, CBA, BF8Q}\}$ indexes the BFS series. Comparing β^s across the application series (BA, HBA, WBA, CBA) reveals whether legalization affects the volume or composition of applications. The BF8Q series is included for descriptive completeness, but its β^s coefficient cannot be given a causal interpretation under the standard DiD timing assumptions because BF8Q is a forward-looking measure whose outcome window may span treatment dates (see Section 6.5).

5.5 Inference

With 21 treated states and 30 control states, cluster-robust standard errors may suffer from size distortions ([MacKinnon et al., 2023](#)). I supplement the standard cluster-robust inference

with two additional approaches:

Randomization inference. I conduct a permutation test by randomly reassigning treatment (state identities and cohort timing) to 21 of 51 states in 999 permutations, re-estimating the TWFE coefficient each time. The two-sided p -value is the fraction of permutation coefficients with absolute value at least as large as the observed coefficient.

Pairs cluster bootstrap. I implement a pairs cluster bootstrap (Cameron et al., 2008) that resamples states with replacement, maintaining the within-state correlation structure. The bootstrap p -value and 95% confidence interval are computed from 999 bootstrap replications.

5.6 Threats to Validity

Several concerns could threaten the validity of the estimates:

Selection into treatment. States that legalize recreational marijuana are not randomly selected. They tend to be politically liberal, more urban, and wealthier. To the extent that these characteristics are time-invariant, state fixed effects absorb them. To the extent that they correlate with time-varying trends, the parallel trends assumption would be violated. I assess this concern through pre-trend analysis and by restricting the control group to states with medical marijuana programs (which share some of the political and demographic characteristics of recreational-legalizing states).

COVID-19 confounding. The pandemic caused a dramatic surge in business applications nationwide beginning in mid-2020 (Haltiwanger, 2021). Several states legalized retail sales during this period (Illinois, Maine, Arizona), making it difficult to disentangle legalization effects from pandemic effects. I address this by estimating specifications that exclude 2020–2021.

Concurrent policies. States that legalize recreational marijuana may simultaneously adopt other policies that affect business formation (e.g., tax changes, occupational licensing reform). Year fixed effects absorb national trends, but state-specific policy changes could confound the estimates. I note this as a limitation.

6. Results

6.1 Descriptive Evidence

Figure 1 presents the raw trends in business applications per 100,000 population for retail-opening and no-retail states from 2005 to 2024. Both groups exhibit broadly parallel trends from 2005 through 2013, with a sharp increase during the post-2020 business formation boom. The no-retail states show somewhat higher average application rates per capita, particularly

in the post-2020 period.

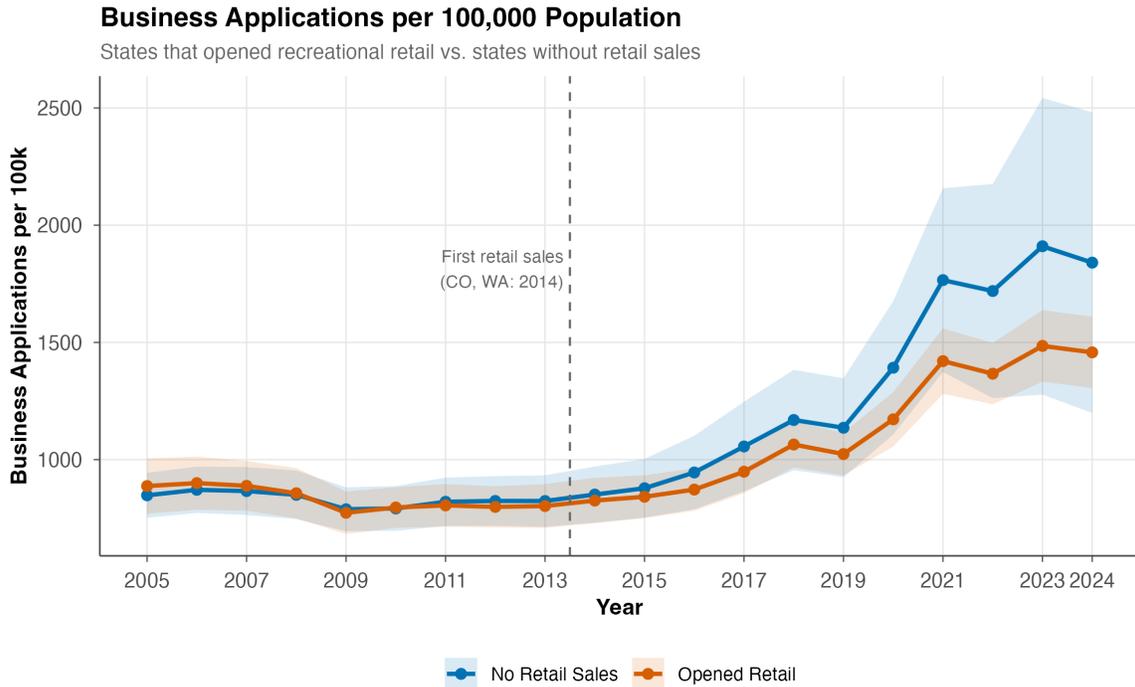


Figure 1: Business Applications per 100,000 Population: Treated vs. Never-Treated States

Visual inspection suggests that the two groups track each other closely in the pre-treatment period, with some divergence appearing after 2014 when the first states opened retail markets. However, the staggered nature of treatment makes raw comparisons difficult to interpret, motivating the formal econometric analysis below.

6.2 Main Results: Callaway-Sant’Anna Estimates

Table 3 presents the Callaway-Sant’Anna estimates, which constitute the primary results of this paper. Panel A reports the overall ATT: -0.028 log points with a standard error of 0.029 and a 95% confidence interval of $[-0.085, 0.029]$. This estimate implies a decline of approximately 2.8 percent in business applications per capita, though the effect is not statistically significant at the 5 percent level. For a median-sized state with approximately 45,000 annual business applications, this translates to roughly 1,250 fewer applications per year. The 95% CI allows us to rule out effects larger than -0.085 log points (approximately -8 percent) and effects more positive than $+0.029$ log points ($+3$ percent).

The CS estimation uses 1,020 state-year observations from 51 units (21 treated, 30 never-treated) over 2005–2024.

Table 3: Callaway-Sant’Anna Estimates

	ATT	SE
<i>Panel A: Overall ATT</i>		
All cohorts	-0.0283	(0.0302)
	[-0.0874, 0.0309]	
<i>Panel B: By Cohort</i>		
Cohort 2014	-0.0624	(0.0308)
Cohort 2015	-0.0793	(0.0342)
Cohort 2016	-0.0787	(0.0338)
Cohort 2017	-0.1394	(0.0228)
Cohort 2018	-0.0526	(0.0697)
Cohort 2019	0.0668	(0.0311)
Cohort 2020	-0.0277	(0.0683)
Cohort 2021	0.0479	(0.0223)
Cohort 2022	0.0752	(0.0527)
Cohort 2023	-0.0045	(0.0304)
Cohort 2024	0.0277	(0.0081)

Notes: Callaway-Sant’Anna (2021) group-time ATT estimates using doubly robust estimator. Outcome: log business applications per capita. Comparison group: never-treated states. 95% confidence intervals in brackets. Standard errors based on bootstrap with 999 replications. Sample: $N = 1,020$ state-year observations (51 units: 21 treated, 30 never-treated; 2005–2024).

Panel B reports cohort-specific ATTs, which reveal substantial heterogeneity across treatment cohorts. Early adopters show consistently negative effects: the 2014 cohort (Colorado, Washington) has an ATT of -0.062 , the 2015 cohort (Oregon) shows -0.079 , and the 2017 cohort (Nevada) shows the largest negative effect at -0.139 . Later cohorts exhibit more mixed results, with the 2019 cohort (Michigan) showing a positive estimate of $+0.067$ and the 2022 cohort (Montana, New Jersey, New Mexico, New York, Rhode Island, Vermont) showing $+0.075$. The 2024 cohort (Ohio) contributes only one post-treatment observation.

This pattern of declining negative effects over time is consistent with the framework’s prediction that early adopters face greater uncertainty. It is also consistent with later cohorts benefiting from the regulatory templates established by pioneers, or with the pandemic-era business formation boom disproportionately affecting later-legalizing states.

6.3 TWFE Benchmark Results

As a benchmark, Table 4 presents conventional TWFE estimates across three specifications. Column (1) uses log total business applications as the outcome. Columns (2) and (3) use log applications per capita, with column (3) adding a control for medical marijuana status.

Across all specifications, the estimated effect is negative, with point estimates ranging

Table 4: TWFE Estimates: Recreational Marijuana and Business Formation

	Log Applications	Log Apps/Capita	Log Apps/Capita + Medical
Rec. Retail Sales	−0.069* (0.039) [−0.147, 0.009]	−0.068* (0.040) [−0.148, 0.013]	−0.075* (0.044) [−0.162, 0.013]
Medical MJ			−0.031 (0.036) [−0.104, 0.042]
Num.Obs.	1,020	1,020	1,020
R2	0.990	0.927	0.928

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

State-clustered standard errors in parentheses; 95% confidence intervals in brackets. All specifications include state and year fixed effects.

from -0.068 to -0.075 log points (95% CIs shown in brackets). In the per-capita specification (column 2), the coefficient of -0.068 ($SE = 0.040$, $N = 1,020$) implies approximately a 7 percent reduction—about 3,000 fewer applications per year for a median-sized state. These TWFE estimates are larger in magnitude than the CS ATT, consistent with heterogeneity bias under staggered adoption (Goodman-Bacon, 2021): because early adopters (with larger negative effects) serve as comparison units for later cohorts, the TWFE coefficient is mechanically pulled toward the more negative early-adopter effects.

6.4 Event Study

Figure 2 presents the Callaway-Sant’Anna event-study estimates, plotting the ATT by years relative to first retail sales. The estimates reveal three key patterns.

Event Study: Effect of Recreational Marijuana Retail Sales on Business Formation

Callaway-Sant'Anna estimates, log business applications per capita

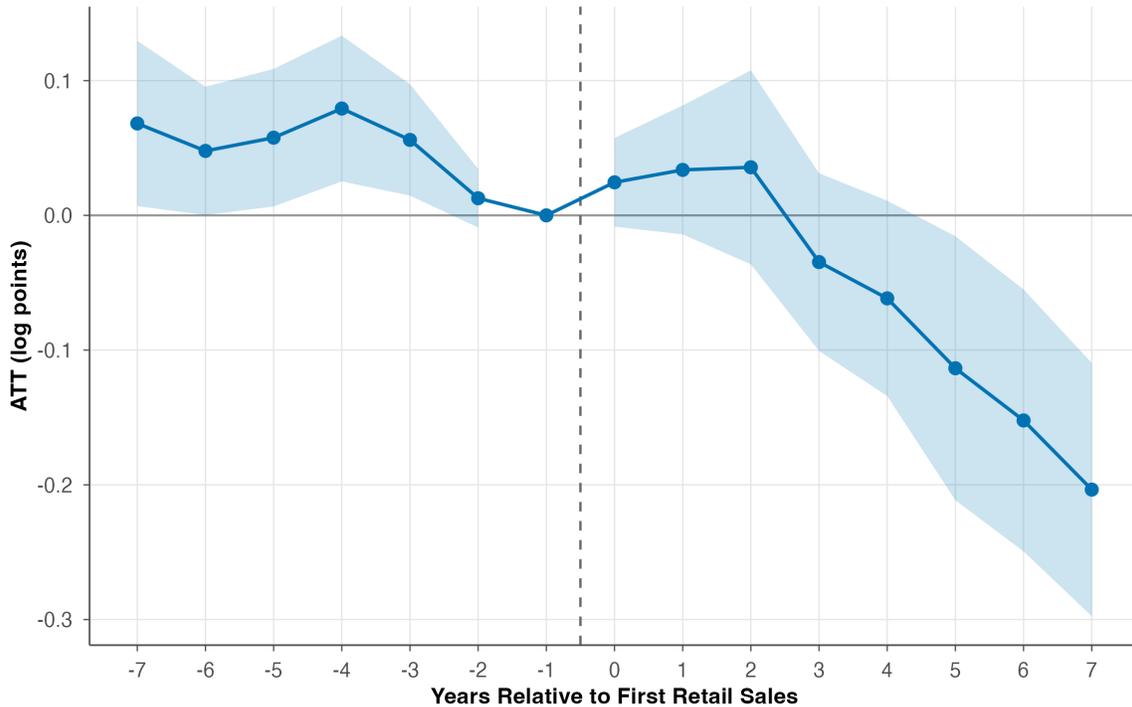


Figure 2: Event Study: Callaway-Sant'Anna Dynamic ATT Estimates

First, the pre-treatment coefficients are not individually statistically significant at the 5 percent level for all seven pre-treatment periods, providing support for the parallel trends assumption. Pre-treatment estimates are modest in magnitude, though some point estimates reach magnitudes of 0.05–0.08 log points (the largest is 0.079), and the 95% confidence intervals include zero.

Second, the post-treatment effects emerge gradually. The estimate at event time 0 (the year of retail opening) is close to zero, with negative effects growing in magnitude over the first 2–3 years. This gradual onset is consistent with the theoretical prediction that the regulatory adjustment period following legalization unfolds over multiple years as licensing, banking, and enforcement norms are established.

Third, the negative effects continue to grow over time rather than stabilizing. Point estimates reach approximately -0.06 at event time +4, -0.11 at +5, and -0.15 to -0.20 by 6–7 years post-treatment. While confidence intervals are wide and most individual post-treatment estimates are not individually significant, the monotonically declining pattern is notable and suggests that the longer-run effects of legalization on business applications may be larger than the overall ATT implies.

Table 5: TWFE Estimates Across BFS Series: Application Types and Business Formations

	Total BA	High-Propensity	Planned Wages	Corporate	Formations (8Q)
Rec. Retail Sales	-0.068*	-0.046	-0.020	0.007	0.030
	(0.040)	(0.037)	(0.022)	(0.070)	(0.018)
	[-0.148, 0.013]	[-0.120, 0.028]	[-0.064, 0.025]	[-0.133, 0.147]	[-0.006, 0.066]
Num.Obs.	1,020	1,020	1,020	1,020	623
R2	0.927	0.929	0.936	0.931	0.962

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

State-clustered standard errors in parentheses; 95% confidence intervals in brackets. All specifications include state and year fixed effects. Outcome: log applications (or formations) per capita. BF8Q result is descriptive only (see text).

6.5 BFS Series Decomposition

Table 5 and Figure 3 present the TWFE estimates across different BFS series, providing a decomposition of the overall effect by application type.

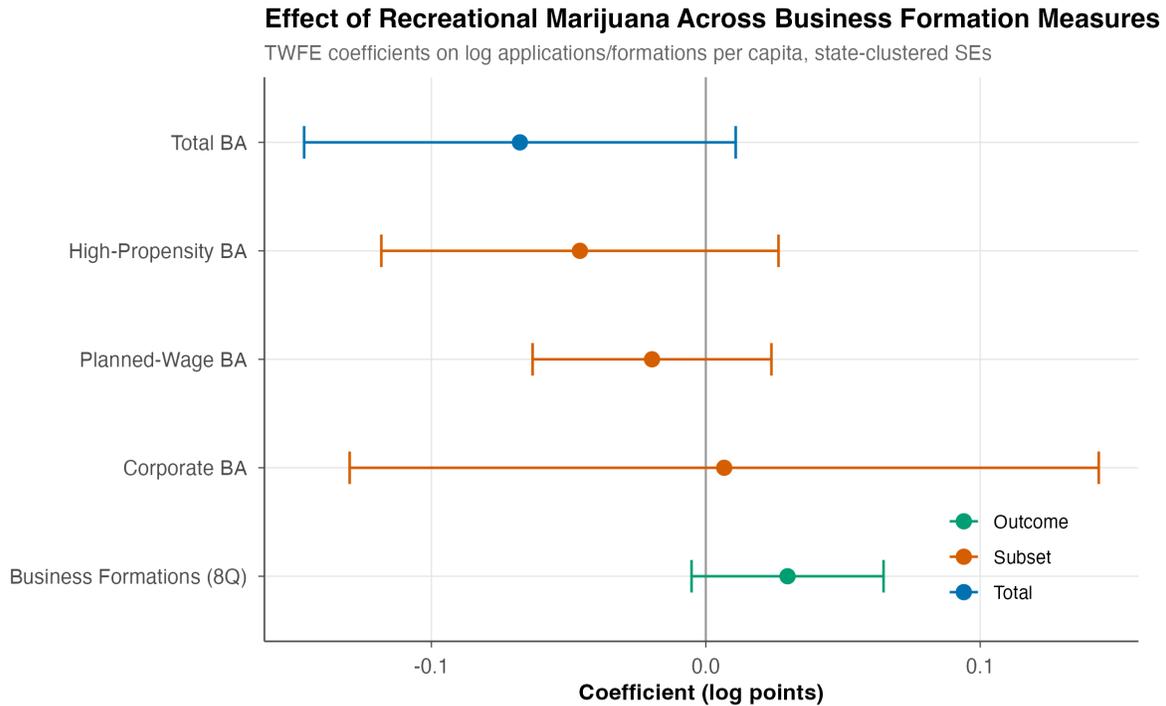


Figure 3: TWFE Coefficients Across BFS Series

The decomposition reveals an important pattern. Total business applications decline by approximately 6.8 percent (significant at 10%), and high-propensity applications decline

modestly (not significant). Wage-planned applications show essentially no change, while corporate applications are unchanged.

The BF8Q series (actual business formations within 8 quarters) shows a positive TWFE coefficient of 0.030 (SE = 0.018, $N = 623$). However, this estimate must be interpreted with caution for three reasons. First, BF8Q is a forward-looking measure indexed by the EIN application date: it counts whether applications from year t result in an employer business within 8 quarters (t to $t+2$). This means that for application cohorts in the years immediately preceding retail opening, the BF8Q outcome window extends *into the post-treatment period*, violating the timing assumption required for the DiD estimand. For example, if a state opens retail in 2018, the BF8Q value for application year 2017 partly reflects formation behavior in 2018–2019. Second, BF8Q data are only available through 2020 due to the Census Bureau’s lag in releasing matched formation data, so only treatment cohorts through 2020 (Colorado through Maine/Illinois) have post-treatment BF8Q observations. Third, the BF8Q panel is unbalanced—only 36–43 of 51 states have valid BF8Q data in any given year, totaling 624 observations across 44 states and 16 years—which further limits comparability with the application-based results.

For these reasons, the positive BF8Q coefficient should be understood as a descriptive association rather than a causal estimate. The suggestive pattern—declining applications but increasing formations—is consistent with a compositional shift in which legalization reduces speculative applications while the cannabis industry attracts well-capitalized ventures with high formation rates. However, confirming this interpretation would require either a formations measure indexed by formation year (e.g., BDS establishment births) or a redesigned estimand that aligns the treatment definition with the BF8Q measurement window.

6.6 Heterogeneity

I examine heterogeneity along the timing dimension by separately estimating treatment effects for early adopters (first retail sales by 2017: CO, WA, OR, AK, NV) and late adopters (first retail sales after 2017).

Early adopters show a larger negative effect on applications ($\beta = -0.103$, SE = 0.055) compared to late adopters ($\beta = -0.046$, SE = 0.043). While neither estimate is individually significant at the 5 percent level, the pattern is consistent with early adopters facing greater uncertainty and fewer regulatory templates to follow.

Figure 5 in the appendix presents raw trends by treatment cohort, showing that cohorts entering during the 2020–2021 COVID period experienced dramatically different baseline dynamics, making it difficult to isolate legalization effects for these cohorts.

6.7 Robustness

Table 6 summarizes the robustness analysis.

Table 6: Robustness Checks: Summary of Estimates

Specification	ATT	SE	N	Notes
<i>Panel A: Outcome Variations</i>				
Total BA (baseline)	-0.0283	(0.0302)	1,020	CS, 51 units
High-Propensity BA	-0.0304	(0.0239)	1,020	CS, 51 units
Wage-Planned BA	-0.0153	(0.0199)	1,020	CS, 51 units
Corporate BA	-0.0138	(0.0497)	1,020	CS, 51 units
<i>Panel B: Sample Variations</i>				
Medical-only control	-0.0053	(0.0279)	780	39 units (21 treated)
Excl. COVID (2020–21)	-0.0414	(0.0294)	918	51 units, 2005–2019 + 2022–2024
Interior controls only	-0.042	(0.0339)	600	30 units (9 controls)
<i>Panel C: Inference</i>				
Randomization inference	$p = 0.093$		1,020	999 permutations
Pairs cluster bootstrap	$p = 0.064$		1,020	999 replications

Notes: Panel A varies the outcome variable across Business Formation Statistics series. Panel B varies the sample. “Interior controls” excludes states bordering any treated state, retaining only non-adjacent never-treated states. Panel C reports alternative inference procedures. N is the number of state-year observations in each specification. All CS estimates use never-treated states as comparison group with doubly robust estimation.

Sample sizes (N) are reported in Table 6 for each specification. The BF8Q descriptive specification (not shown in Table 6) uses $N = 623$ observations (unbalanced panel covering 44 states and 2005–2020; see Section 4.1 for details).

Alternative outcomes. The CS ATT is qualitatively similar across BFS series: -0.030 ($SE = 0.025$) for HBA and -0.015 ($SE = 0.019$) for WBA, neither statistically significant. All CS estimates use never-treated states as comparison group unless otherwise noted. The consistency of the sign across application types (with the exception of BF8Q) suggests a genuine, if modest, reduction in the volume of applications.

Medical-only control group. Restricting the control group to states with medical marijuana programs (which may be more comparable to recreational-legalizing states on unobservable dimensions) yields a CS ATT of -0.005 ($SE = 0.028$), smaller in magnitude and not significant. This estimate represents the effect of *recreational* legalization conditional on already having medical marijuana, and the attenuation is consistent with medical marijuana states having already partially adjusted to cannabis-related regulatory changes.

Excluding COVID years. Dropping 2020–2021 from the sample yields a CS ATT of -0.041 ($SE = 0.030$), slightly larger in magnitude than the full-sample estimate. This suggests

that the COVID-era business formation boom, which disproportionately affected states that legalized during this period, may attenuate the estimated negative effect.

Interior controls (spillover robustness). Legalization may generate cross-border spillovers—for instance, if entrepreneurs in neighboring states redirect activity toward the legal market, or if regulatory competition affects business decisions in border regions. To assess whether such spillovers contaminate the control group, I restrict the comparison group to nine “interior” never-treated states that do not share a border with any of the 21 treated states (AL, FL, GA, HI, LA, MN, MS, NC, SC). The resulting CS ATT is -0.042 ($SE = 0.034$) from 600 observations across 30 units (21 treated, 9 controls). This estimate is directionally similar to, and slightly larger than, the baseline CS ATT, suggesting that the main results are not driven by positive spillovers to border control states that would bias the baseline toward zero.

Randomization inference. The permutation test yields a two-sided p -value of 0.093, indicating that the observed TWFE coefficient of -0.068 is more extreme than approximately 91% of placebo coefficients. Figure 4 displays the permutation distribution, showing that the observed coefficient falls in the left tail of the distribution, consistent with a genuine negative relationship.

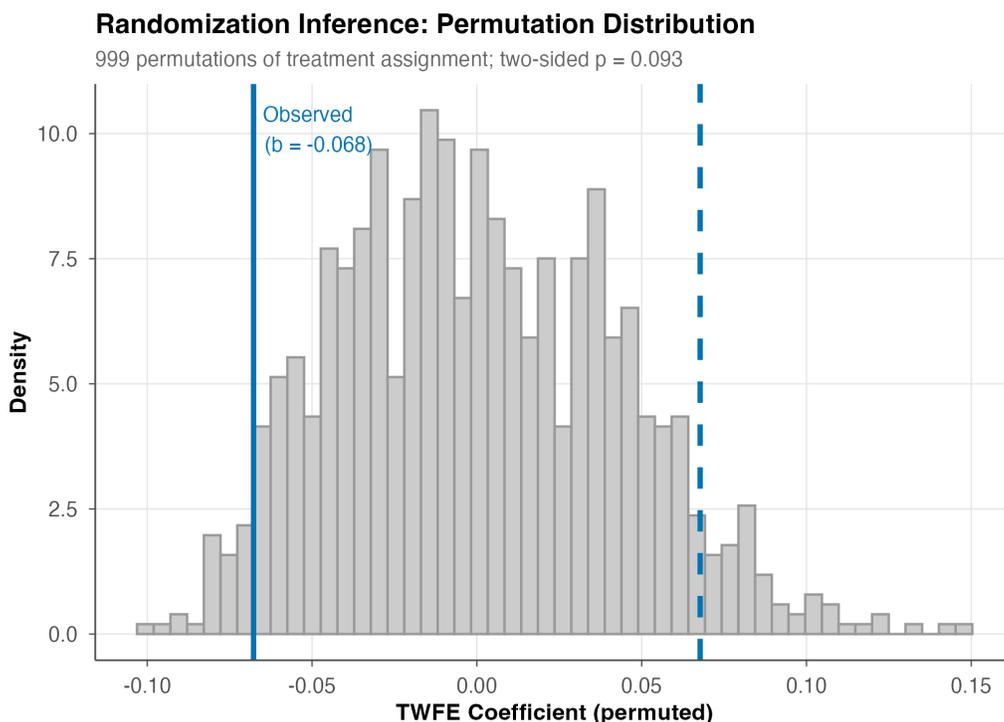


Figure 4: Randomization Inference: Permutation Distribution of TWFE Coefficients

Pairs cluster bootstrap. The bootstrap 95% confidence interval is $[-0.151, 0.001]$, and the

bootstrap p -value is 0.064. The bootstrap results are broadly consistent with the analytical cluster-robust inference, confirming a negative point estimate that is marginally significant.

7. Discussion

7.1 Interpreting the Results

The main finding is a modest decline in business applications following recreational marijuana retail legalization. While a descriptive analysis of business formations (BF8Q) suggests a potential compositional shift—applications declining while formations increase—the BF8Q result cannot be interpreted causally due to the forward-looking nature of the measure (see Section 6.5). The most robust finding is therefore the application-side decline itself, which admits several interpretations. One explanation is that the regulatory complexity and federal uncertainty associated with legalization deters some marginal applicants, while well-prepared entrepreneurs who enter the cannabis industry are less affected.

This interpretation is consistent with the broader literature on regulatory barriers to entry (Djankov et al., 2002). Regulation can serve a screening function: by raising the cost of entry, it filters out less committed entrepreneurs. In the marijuana context, the unique challenges of operating in a federally prohibited industry—cash-only transactions, Section 280E tax burdens, limited banking access—may serve as an extreme version of this screening mechanism.

7.2 Comparison with Prior Work

My findings both complement and contrast with Brown et al. (2023), who find a positive effect of recreational legalization on establishment births using BDS data covering 12 states through 2020. Three key differences in measurement may explain the sign discrepancy. First, BDS captures *establishments* (physical locations of existing firms), while BFS captures *applications* (intent to create new entities). An existing firm opening a new cannabis dispensary would appear in BDS but not in BFS (unless it applies for a new EIN), while a speculative EIN application that never becomes an employer would appear in BFS but not in BDS. Second, BDS establishment births are indexed by the year of first payroll—a formation-year measure—while BFS applications are indexed by the application date. This timing difference matters because my analysis shows that application *intent* may decline even as the *formation rate* among applicants increases. Third, my analysis covers 21 states through 2024, nearly doubling the treated sample relative to Brown et al. (2023). The addition of later cohorts (which show more positive CS ATTs in my analysis) may reconcile the direction of aggregate effects. The

suggestive positive BF8Q association in my analysis is directionally consistent with the BDS finding, though the BF8Q result cannot be interpreted causally due to the timing mismatch discussed in Section 6.5.

7.3 Data Frequency Limitation

An important limitation of this analysis is the use of annual aggregation despite the BFS data being available at monthly frequency. Because treatment dates (first retail sales) occur at specific months—ranging from January 1 for Colorado (2014) and Illinois (2020) to late in the year for states like Ohio (August 2024)—annual aggregation introduces treatment misclassification within the adoption year. States opening retail in January have 12 months of treated exposure in the adoption year, while states opening in October have only 3 months. This “partial-year exposure” problem attenuates dynamic effects and makes event-time 0 non-comparable across cohorts. Future work should re-estimate the core design at monthly frequency with state \times month panels and month-by-year fixed effects, defining treatment at the precise month of first retail sales (Baker et al., 2022; Gardner, 2022). Moving to monthly estimation would sharpen identification, reduce exposure misclassification, and improve the interpretability of dynamic effects.

7.4 Limitations

Several limitations should be noted. First, the 21 treated states, while providing more statistical power than prior studies, remain a moderate number of clusters for inference. The randomization inference and cluster bootstrap results help address this concern but cannot fully resolve it.

Second, BFS data do not include NAICS sector breakdowns at the state level, precluding a triple-difference design that would identify cannabis-adjacent industry effects. The BFS series decomposition provides an alternative lens on the composition of effects, but it cannot directly test whether the application decline is concentrated in cannabis-related industries.

Third, several states legalized during the COVID-19 pandemic, which caused unprecedented disruption to business formation patterns. While I present estimates excluding 2020–2021, the pandemic may have affected the trajectory of treated and control states differently, potentially confounding the estimates for pandemic-era cohorts.

Fourth, local option provisions mean that treatment intensity varies within states. A state-level analysis treats all areas within a legalizing state as equally treated, but in practice, many municipalities within treated states have banned marijuana businesses. This measurement error in treatment intensity likely attenuates the estimated effects.

Fifth, the BF8Q (business formations within 8 quarters) measure is forward-looking—indexed by application date but measuring outcomes over the subsequent two years—which prevents a clean causal interpretation under the DiD design. Future work using formation-year-indexed measures (e.g., BDS establishment births) could address this limitation.

Sixth, the analysis uses annual aggregation despite BFS data being available at monthly frequency. As discussed in Section 7.2, this introduces treatment misclassification within the adoption year and attenuates dynamic effects. Estimating the design at monthly frequency would sharpen identification and improve comparability across cohorts.

Finally, cross-border spillovers may affect both treated and control states. Business formation in non-legalizing states near the border of legalizing states may be affected by cross-border shopping, migration, or regulatory competition. While the interior-controls robustness check (Section 6.7) suggests that such spillovers do not substantially bias the main estimates, a more definitive test would require county-level data and a border-discontinuity design.

7.5 Policy Implications

These results suggest that policymakers considering recreational marijuana legalization should not expect dramatic changes in overall business formation. The modest decline in business applications, while suggestive, is not consistently statistically significant at conventional levels, pointing to a net effect that is likely small. The finding that effects are larger for early adopters suggests that states benefit from the regulatory templates and institutional knowledge developed by pioneers, implying potential gains from coordination across states.

The event-study finding that effects grow larger over time also has implications for how states design their cannabis licensing systems. Policies that reduce regulatory barriers (streamlined licensing, banking access, federal policy reform) could attenuate the negative effect on applications, particularly in the years following retail opening when regulatory uncertainty is highest.

8. Conclusion

This paper provides the first comprehensive analysis of how recreational marijuana legalization affects business formation using the Census Bureau’s Business Formation Statistics. Using data for 51 states over 20 years and exploiting the staggered adoption of recreational retail across 21 states, I estimate a Callaway-Sant’Anna ATT of -0.028 log points (95% CI: $[-0.085, 0.029]$), indicating a modest decline that is not statistically distinguishable from zero at conventional levels. Conventional TWFE benchmark estimates are larger (-0.068),

consistent with heterogeneity bias under staggered adoption. A descriptive analysis of actual business formations (BF8Q) suggests a positive association, but this result cannot be given a causal interpretation under the DiD design due to the forward-looking nature of the BF8Q measure. The primary finding is therefore that, if anything, recreational marijuana legalization modestly reduces business applications—particularly for early-adopting states—but the effects are imprecisely estimated and cannot rule out zero or small positive effects.

The results are robust to multiple inference procedures (randomization inference, pairs cluster bootstrap), alternative control groups (medical-only states, interior non-border states), and exclusion of COVID-era observations. Heterogeneity across treatment cohorts reveals that early-adopting states experienced larger negative effects, consistent with the higher regulatory uncertainty and lack of established templates faced by pioneers.

These findings contribute to the growing literature on the economic consequences of marijuana legalization and suggest that states considering legalization should not expect dramatic disruptions to overall business formation. The discrepancy with [Brown et al. \(2023\)](#)'s positive BDS findings likely reflects differences in measurement (application intent vs. establishment formation) rather than a true contradiction, and future work using formation-year-indexed measures could reconcile these results.

Several avenues for future research would strengthen the analysis. First, re-estimating the design at monthly frequency—exploiting the precise month of first retail sales and BFS monthly data—would sharpen identification and reduce exposure misclassification. Second, implementing formal sensitivity analysis for violations of parallel trends ([Rambachan and Roth, 2023](#)) would characterize how robust the findings are to small departures from the identifying assumption. Third, a border-county design using county-level business formation data would provide stronger identification by comparing similar counties across state borders. Fourth, as state-level NAICS data become available in the BFS, a triple-difference design could directly identify cannabis-adjacent industry effects.

Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

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

References

- Anderson, D. Mark and Daniel I. Rees**, “The Effects of Marijuana Legalization: A Review of the Economic Literature,” *Journal of Economic Literature*, 2023, 61 (1), 86–143.
- Baker, Andrew C., David F. Larcker, and Charles C. Y. Wang**, “How Much Should We Trust Staggered Difference-in-Differences Estimates?,” *Journal of Financial Economics*, 2022, 144 (2), 370–395.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan**, “How Much Should We Trust Differences-in-Differences Estimates?,” *Quarterly Journal of Economics*, 2004, 119 (1), 249–275.
- Borusyak, Kirill, Xavier Jaravel, and Jann Spiess**, “Revisiting Event-Study Designs: Robust and Efficient Estimation,” *Review of Economic Studies*, 2024, 91 (6), 3253–3285.
- Brown, Jason P., Jeffrey P. Cohen, and Alison Felix**, “Recreational Marijuana Legalization and Business Formation,” Research Working Paper RWP 23-10, Federal Reserve Bank of Kansas City 2023.
- Callaway, Brantly and Pedro H. C. Sant’Anna**, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, 225 (2), 200–230.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller**, “Bootstrap-Based Improvements for Inference with Clustered Errors,” *Review of Economics and Statistics*, 2008, 90 (3), 414–427.
- Chakraborty, Arijit, Jacqueline M. Doremus, and Sarah S. Stith**, “Cannabis Legalization and House Prices: Evidence from Colorado,” *Journal of Regional Science*, 2023, 63 (1), 269–295.
- Conley, Timothy G. and Christopher R. Taber**, “Inference with “Difference in Differences” with a Small Number of Policy Changes,” *Review of Economics and Statistics*, 2011, 93 (1), 113–125.
- de Chaisemartin, Clément and Xavier D’Haultfœuille**, “Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects,” *American Economic Review*, 2020, 110 (9), 2964–2996.
- Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda**, “The Role of Entrepreneurship in US Job Creation and Economic Dynamism,” *Journal of Economic Perspectives*, 2014, 28 (3), 3–24.

- Dills, Angela K., Sietse Goffard, and Jeffrey Miron**, “The Effect of State Marijuana Legalizations: 2021 Update,” *CATO Institute Policy Analysis*, 2021, (908).
- Djankov, Simeon, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer**, “The Regulation of Entry,” *Quarterly Journal of Economics*, 2002, *117* (1), 1–37.
- Gardner, John**, “Two-Stage Difference-in-Differences,” *Journal of Econometrics*, 2022.
- Gavrilova, Evelina, Takuma Kamada, and Floris Zoutman**, “Is Legal Pot Crippling Mexican Drug Trafficking Organisations? The Effect of Medical Marijuana Laws on US Crime,” *Economic Journal*, 2019, *129* (617), 375–407.
- Goodman-Bacon, Andrew**, “Difference-in-Differences with Variation in Treatment Timing,” *Journal of Econometrics*, 2021, *225* (2), 254–277.
- Haltiwanger, John C.**, “Entrepreneurship During the COVID-19 Pandemic: Evidence from the Business Formation Statistics,” *Entrepreneurship and Innovation Policy and the Economy*, 2021, *1*, 9–42.
- Haltiwanger, John, Ron S. Jarmin, and Javier Miranda**, “Who Creates Jobs? Small versus Large versus Young,” *Review of Economics and Statistics*, 2013, *95* (2), 347–361.
- Hansen, Benjamin, Keaton Miller, and Caroline Weber**, “The Taxation of Recreational Marijuana: Evidence from Washington State,” *Journal of Public Economics*, 2020, *191*, 104318.
- Kilmer, Beau, Jonathan P. Caulkins, Rosalie Liccardo Pacula, and Peter Reuter**, *Marijuana Legalization: What Everyone Needs to Know*, 2nd ed., Oxford University Press, 2020.
- Klapper, Leora, Luc Laeven, and Raghuram Rajan**, “Entry Regulation as a Barrier to Entrepreneurship,” *Journal of Financial Economics*, 2006, *82* (3), 591–629.
- MacKinnon, James G., Morten Ørregaard Nielsen, and Matthew D. Webb**, “Cluster-Robust Inference: A Guide to Empirical Practice,” *Journal of Econometrics*, 2023, *232* (2), 272–299.
- Marijuana Policy Project**, “State-by-State Marijuana Laws,” 2023. Accessed January 2025.

- Nicholas, Lauren Hersch and Johanna Catherine Maclean**, “The Effects of Recreational Cannabis Access on Labor Markets: Evidence from Colorado,” *Economic Inquiry*, 2022, *60* (4), 1691–1713.
- Rambachan, Ashesh and Jonathan Roth**, “A More Credible Approach to Parallel Trends,” *Review of Economic Studies*, 2023, *90* (5), 2555–2591.
- Roth, Jonathan**, “Pre-test with Caution: Event-Study Estimates after Testing for Parallel Trends,” *American Economic Review: Insights*, 2022, *4* (3), 305–322.
- Sabia, Joseph J. and Thanh Nguyen**, “Marijuana Decriminalization, Price, and Adolescent Marijuana Use,” *Southern Economic Journal*, 2017, *84* (2), 477–510.
- Sant’Anna, Pedro H. C. and Jun Zhao**, “Doubly Robust Difference-in-Differences Estimators,” *Journal of Econometrics*, 2020, *219* (1), 101–122.
- Sun, Liyang and Sarah Abraham**, “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects,” *Journal of Econometrics*, 2021, *225* (2), 175–199.
- U.S. Census Bureau**, “Business Formation Statistics,” 2024. Monthly data accessed January 2025.

A. Data Appendix

A.1 Business Formation Statistics

The Business Formation Statistics (BFS) program produces data on new business applications and formations in the United States. The data are derived from applications for Employer Identification Numbers (EINs) submitted to the Internal Revenue Service (IRS). The BFS is produced by the Census Bureau’s Center for Economic Studies and has been publicly available since 2018, with data extending back to July 2004.

I download the complete monthly BFS dataset from https://www.census.gov/econ/bfs/csv/bfs_monthly.csv. This file contains monthly counts by geographic area (national, state), NAICS sector (national level only), seasonal adjustment flag, and series type. I retain only unadjusted (not seasonally adjusted) state-level observations and aggregate monthly counts to annual totals.

The key BFS series used in this analysis are:

- **BA__BA:** Total business applications—all new EIN applications.
- **BA__HBA:** High-propensity business applications—applications with characteristics predicting high likelihood of becoming an employer business. The Census Bureau’s classification algorithm uses planned wages, business type, industry, and other application features.
- **BA__WBA:** Business applications with planned wages—applications where the applicant indicates an expected first wage payment date.
- **BA__CBA:** Corporate business applications—applications filed as corporations, S-corporations, or LLCs electing corporate taxation.
- **BF__BF8Q:** Business formations within 8 quarters—the count of applications that produce a business with payroll tax filings within 8 quarters of the EIN application date.

NAICS sector breakdowns are available only at the national level, not at the state level. This limitation precludes a triple-difference design using industry-level variation.

A.2 Population Data

State-level population data come from the American Community Survey (ACS) 1-year estimates, accessed via the Census Bureau API (<https://api.census.gov/data/>). I retrieve

total population (variable B01003_001E) for each state and year from 2005 to 2023. The 2020 ACS 1-year estimates were not released due to COVID-19 data collection disruptions; for 2020, I use linear interpolation between 2019 and 2021 values. For 2024, I use the 2023 population estimate.

A.3 Treatment Timing

Treatment timing is based on the date of first legal recreational marijuana retail sales. Sources include state legislative records, the Marijuana Policy Project’s state-by-state status reports ([Marijuana Policy Project, 2023](#)), and news coverage of retail openings. I assign each state to the calendar year in which its first recreational dispensary opened.

A.4 Medical Marijuana Timing

Medical marijuana program dates are obtained from NCSL reports and the Marijuana Policy Project. I record the year each state’s medical marijuana program became operational (allowing patient access), not the year of legislative authorization.

B. Identification Appendix

B.1 Pre-Trends Analysis

Figure 2 in the main text presents the Callaway-Sant’Anna event-study estimates, which serve as the primary pre-trends test. All seven pre-treatment coefficients (event times -7 through -1) are not individually statistically significant at the 5 percent level, though some point estimates reach magnitudes of 0.05–0.08 log points (the largest is 0.079 in absolute value). A joint test of the null hypothesis that all pre-treatment coefficients equal zero cannot be computed due to a singular covariance matrix (reflecting the small number of groups and the collinearity among group-time estimates), but the individual coefficient tests provide no evidence of differential pre-trends.

Two important caveats apply to this pre-trend analysis. First, as emphasized by [Roth \(2022\)](#), the absence of statistically significant pre-trends does not establish that parallel trends hold. Pre-tests have limited power—particularly with a moderate number of treated clusters (21) and annual data aggregation—and conditional on passing a pre-test, the estimated treatment effects can be biased. Second, a formal sensitivity analysis along the lines of [Rambachan and Roth \(2023\)](#) would provide more informative bounds on the treatment effect under various assumptions about the smoothness of violations to parallel trends (“HonestDiD”). Such an analysis would characterize the set of treatment effects consistent

with pre-trend violations no larger than, say, twice the magnitude of the largest observed pre-trend coefficient. I leave this formal sensitivity analysis for future work, but note that the largest pre-trend coefficient (0.079 in absolute value) is comparable in magnitude to some of the post-treatment estimates, suggesting that the post-treatment effects should be interpreted cautiously. Alternative approaches—such as collapsing pre-treatment leads into bins to reduce the dimensionality of the pre-test, or using the imputation estimator of [Borusyak et al. \(2024\)](#) which may yield better-behaved variance estimates—could also help address the singular covariance issue.

B.2 Cohort-Specific Trends

Figure 5 presents raw trends in business applications per capita for each treatment cohort, along with the never-treated average. The figure reveals that cohorts share broadly similar pre-treatment trends, with divergence emerging around the time of treatment.

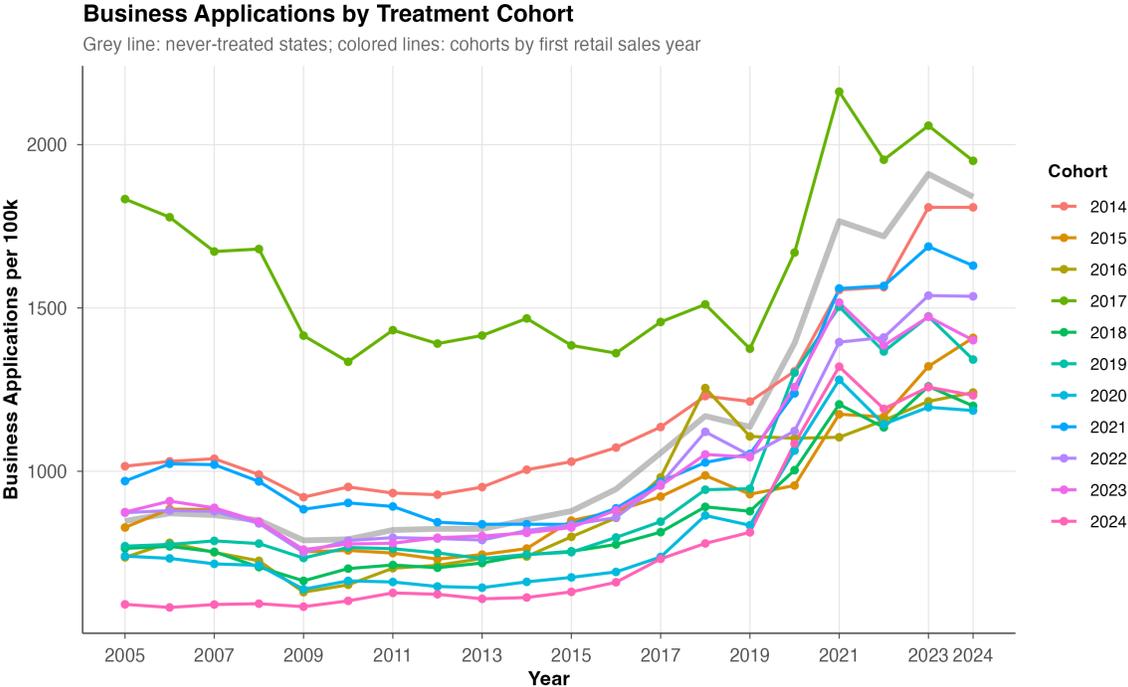


Figure 5: Business Application Trends by Treatment Cohort

B.3 Individual State Trends

Figure 6 displays individual trends for the six earliest-adopting states (Colorado, Washington, Oregon, Alaska, Nevada, California) alongside the never-treated average. Colorado and Washington, the first to open retail markets, show application rates that initially tracked the

never-treated average closely before exhibiting slightly lower growth in the post-treatment period.

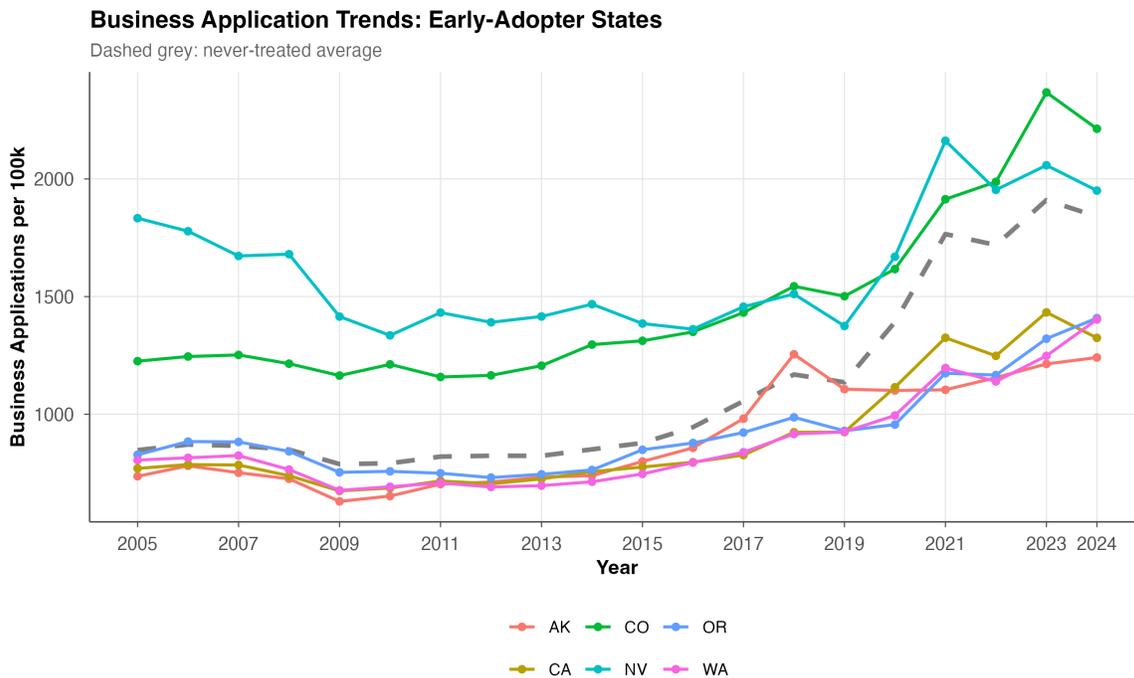


Figure 6: Business Application Trends: Early-Adopter States

C. Robustness Appendix

C.1 Randomization Inference Details

The randomization inference procedure assigns treatment status to 21 randomly selected states (matching the number of treated states in the data) and randomly assigns cohort timing from the observed distribution of treatment years. For each of 999 permutations, I re-estimate the TWFE specification in Equation 2 and record the coefficient on the permuted treatment indicator. The two-sided p -value is the fraction of permutation coefficients with absolute value $\geq |\hat{\beta}|$.

The observed TWFE coefficient of -0.068 falls at approximately the 9th percentile of the permutation distribution. The permutation coefficients are centered near zero, confirming that the assignment procedure is approximately unbiased and that the observed coefficient is substantially larger in magnitude than what would be expected by chance.

C.2 Pairs Cluster Bootstrap Details

The pairs cluster bootstrap resamples 51 states with replacement, maintaining the entire within-state time series for each selected state. For each of 999 bootstrap replications, I re-estimate the TWFE specification using the bootstrapped sample, assigning new state identifiers to handle duplicate states.

The bootstrap 95% confidence interval $[-0.151, 0.001]$ is broadly consistent with the analytical cluster-robust inference, and the bootstrap p -value is 0.064. The point estimate is marginally significant, confirming a negative effect that borders on conventional significance levels.

D. Additional Figures and Tables

Additional exhibits are referenced in the main text. The full replication code is available in the `code/` directory of the paper's repository.