

The Price of Distance: Cannabis Dispensary Access and the Composition of Fatal Crashes

APEP Autonomous Research*

@anonymous

January 2024

Abstract

This paper investigates whether access to legal marijuana reduces the share of fatal crashes involving alcohol through marijuana-alcohol substitution. Using FARS crash-level data from 2016–2019 for eight illegal states near legal cannabis markets (ID, WY, NE, KS, UT, AZ, MT, NM), I construct a continuous treatment measure: driving time to the nearest legal dispensary. Among 18,430 fatal crashes, I find that a one-unit increase in log driving time is associated with a 2.4 percentage point increase in the probability that a crash involves alcohol (baseline rate: 31.7%). Doubling drive time corresponds to a 1.7 percentage point increase ($0.024 \times \ln(2) \approx 0.017$). Effects are concentrated near legal-state borders (within 200km), consistent with a cross-border purchasing mechanism. Placebo tests on daytime crashes and elderly driver involvement yield null results, supporting the interpretation that access to legal cannabis substitutes for alcohol consumption. These findings suggest that marijuana legalization may generate spillovers by reducing the alcohol-involved share of fatal crashes in prohibition states.

JEL Codes: I12, I18, K32, R41

Keywords: marijuana legalization, alcohol substitution, traffic fatalities, cross-border spillovers, FARS

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

1. Introduction

Alcohol-impaired driving kills approximately 10,000 Americans annually, accounting for roughly 28% of all traffic fatalities ([National Highway Traffic Safety Administration, 2023](#)). Despite decades of policy interventions—including minimum legal drinking ages, DUI enforcement, and ignition interlock requirements—alcohol-related crash rates remain stubbornly high. A growing body of evidence suggests an unconventional harm reduction pathway: marijuana legalization may reduce alcohol consumption and its associated harms through substance substitution ([Anderson et al., 2013](#); [Hansen et al., 2020](#)).

The existing literature on marijuana and traffic safety has focused primarily on within-state effects, comparing outcomes in states that legalize to those that do not. This approach faces two challenges. First, legalization may change the composition of crashes through multiple channels—more drivers under marijuana influence, tourism effects, and general traffic changes—making it difficult to isolate the alcohol substitution mechanism. Second, state-level comparisons cannot exploit the rich spatial heterogeneity in access to legal marijuana that exists even within prohibition states.

This paper takes a novel approach by studying alcohol-related traffic fatalities in *illegal* states near legal cannabis markets. The key insight is that residents of prohibition states can access legal marijuana by traveling to dispensaries in legal states, with the cost of access determined primarily by driving time. This creates continuous variation in effective marijuana prices within states that maintained prohibition throughout the study period, allowing identification of substitution effects without the confounding factors present in legalizing states. The treatment—driving time to the nearest dispensary in *any* legal state—captures access costs even when the nearest legal market is not in a directly bordering state.

I construct a crash-level dataset from the Fatality Analysis Reporting System (FARS) for 2016–2019, covering eight states that remained illegal throughout this period: Arizona, Idaho, Kansas, Montana, Nebraska, New Mexico, Utah, and Wyoming. For each of 18,430 fatal crashes, I compute driving time to the nearest legal recreational dispensary in any legal state using geocoded crash locations and dispensary coordinates. This continuous treatment measure—ranging from 133 minutes (mean in Idaho, near Oregon/Washington) to 581 minutes (mean in Kansas, far from Colorado’s border)—captures the “price of distance” that determines the relative cost of marijuana versus locally-available alcohol. Note that some states (e.g., Montana) have their nearest legal market in a non-bordering state (Washington), requiring travel through an intervening state; the treatment still measures access costs for these locations.

The main specification regresses an indicator for alcohol involvement on log driving time, with state and year fixed effects. The identifying assumption is that, conditional on state and year, driving time to dispensaries is uncorrelated with unobserved determinants of alcohol-involved crashes. This assumption is supported by the fact that dispensary locations are determined by legal-state regulatory processes, not by characteristics of crashes in neighboring illegal states.

I find that a one-unit increase in log driving time is associated with a 2.4 percentage point increase in alcohol involvement (standard error: 0.8 pp). Given the baseline alcohol involvement rate of 31.7%, this represents a 7.6% relative increase. Doubling driving time—from 150 to 300 minutes—corresponds to a change of $\ln(2) \approx 0.69$ in log driving time, implying an approximately 1.7 percentage point increase in alcohol involvement. The effect is robust to alternative specifications including year-month fixed effects and restricting to border regions within 200km of legal states.

Several pieces of evidence support the substitution interpretation. First, effects are concentrated among crashes within 100km of legal-state borders, where cross-border purchasing is most plausible. Beyond 200km, the relationship attenuates substantially. Second, nighttime crashes—when both drinking and recreational marijuana use are more common—show stronger effects than daytime crashes. Third, placebo tests on daytime crashes and elderly driver involvement yield null results, ruling out general traffic exposure explanations and supporting the alcohol-specific mechanism.

This paper contributes to several literatures. First, I add to the growing evidence on marijuana-alcohol substitution ([Anderson et al., 2013](#); [Hansen et al., 2020](#); [Dills et al., 2021](#)), providing cross-border evidence that avoids the within-state confounds present in earlier studies. Second, I contribute to the broader literature on cross-border policy spillovers, demonstrating that prohibition states are not insulated from their neighbors' legalization decisions. Third, from a methodological perspective, I demonstrate the value of continuous treatment measures in drug policy research, moving beyond binary legal/illegal comparisons.

The findings have important policy implications. If marijuana access reduces alcohol-impaired driving, then prohibition states adjacent to legal markets may experience unintended benefits from their neighbors' policy choices. Conversely, the findings suggest that federal prohibition may impose costs not only through foregone tax revenue but also through reduced harm reduction from alcohol substitution.

The remainder of this paper proceeds as follows. Section 2 provides institutional background on cannabis legalization timing and dispensary access. Section 3 describes the theoretical framework linking driving time to alcohol-related crashes. Section 4 presents the data and summary statistics. Section 5 describes the empirical strategy. Section 6 presents

results, and Section 7 concludes.

2. Institutional Background

2.1 Cannabis Legalization in the Western United States

The legal landscape for recreational marijuana evolved rapidly during the 2010s, creating unprecedented geographic variation in cannabis access. Colorado and Washington became the first states to legalize recreational marijuana through ballot initiatives in November 2012. Colorado’s Amendment 64 and Washington’s Initiative 502 passed with 55% and 56% of the vote, respectively. However, establishing the regulatory infrastructure for retail sales required substantial time. Colorado opened its first recreational dispensaries on January 1, 2014, while Washington’s first stores opened on July 8, 2014.

Oregon followed with Measure 91 in November 2014, but the more deliberate regulatory process meant retail sales did not begin until October 2015—nearly a year after legalization. The Oregon Liquor Control Commission (OLCC) implemented strict licensing requirements including background checks, financial audits, and land use reviews that created significant barriers to entry.

The second wave of western legalization included Nevada and California in November 2016. Nevada moved quickly to establish its recreational market, with retail sales beginning on July 1, 2017—just eight months after the ballot initiative passed. California, despite its large medical marijuana infrastructure and earlier legalization vote, experienced substantial regulatory delays. The California Bureau of Cannabis Control faced challenges coordinating with dozens of local jurisdictions, many of which imposed their own licensing requirements or outright bans on cannabis retail. As a result, California’s recreational retail sales did not begin until January 1, 2018.

Alaska legalized through Measure 2 in November 2014 but faced unique challenges related to its sparse population and vast geography. The first recreational dispensaries opened in October 2016, concentrated primarily in Anchorage and Fairbanks. Given Alaska’s geographic isolation from the continental United States, its dispensaries are not relevant for cross-border access by residents of the lower 48 states.

Table 1 summarizes the timing of legalization and first retail sales for states relevant to this study.

Table 1: Cannabis Legalization Timing in Western Legal States

State	Legalization	First Retail Sales
Colorado	Nov 2012	Jan 2014
Washington	Nov 2012	Jul 2014
Oregon	Nov 2014	Oct 2015
Alaska	Nov 2014	Oct 2016
Nevada	Nov 2016	Jul 2017
California	Nov 2016	Jan 2018

Notes: Dates reflect first legal recreational retail sales. Medical marijuana was legal earlier in all states. Alaska is excluded from cross-border analysis due to geographic isolation.

2.2 The Dispensary Landscape

The recreational cannabis market in these early-adopter states grew rapidly following retail opening. By 2019, Colorado alone had over 500 licensed recreational dispensaries, with particularly high concentrations in Denver, Boulder, and communities along the Interstate 25 corridor. Many dispensaries strategically located near state borders to capture cross-border customers, a phenomenon widely documented in media reports and acknowledged by state regulators.

Washington’s market evolved somewhat differently due to its “hybrid” system that initially separated medical and recreational dispensaries before eventually merging them in 2016. Oregon’s market became notably competitive, with the state licensing hundreds of recreational dispensaries—one of the highest per-capita densities in the nation. Portland and Eugene hosted the largest concentrations, but numerous dispensaries also operated in border communities such as Ontario (near Idaho) and Ashland (near California).

Nevada’s market remained more concentrated, with the majority of dispensaries located in the Las Vegas metropolitan area. However, several dispensaries operated in Reno and smaller communities near the California, Arizona, and Utah borders.

California’s recreational market, despite regulatory challenges, expanded rapidly after retail sales began in January 2018. The distribution reflected California’s diverse geography, with major concentrations in Los Angeles, the Bay Area, and San Diego.

Note on dispensary counts: The empirical analysis uses a curated dispensary sample (N=1,247) rather than the full universe of licensed dispensaries. This sample focuses on

dispensaries with verified geocoded locations from OpenStreetMap, cross-validated against archival state licensing records (see Section 4.2). The curated sample captures the large majority of dispensaries near state borders—the locations most relevant for cross-border access—while excluding interior locations with incomplete geocoding. Coverage is highest for Colorado (520 dispensaries) and Oregon (380), which have the strongest OSM community coverage.

2.3 Neighboring Illegal States

Throughout the 2016–2019 study period, eight states neighboring the legal western markets maintained marijuana prohibition for recreational use: Arizona, Idaho, Kansas, Montana, Nebraska, New Mexico, Utah, and Wyoming. These states varied considerably in their proximity to legal dispensaries and their enforcement postures toward cannabis possession.

Idaho represents perhaps the most striking case of cross-border access. The state shares long borders with Oregon, Washington, and Nevada—all states with recreational cannabis markets. The town of Ontario, Oregon, located directly on the Idaho border, hosts multiple dispensaries explicitly catering to Idaho residents. Idaho state law treats any amount of marijuana possession as a misdemeanor (first offense) or felony (subsequent offenses), yet the short driving distances make cross-border purchasing highly accessible for much of Idaho’s population.

Wyoming borders Colorado directly along its entire southern boundary. Cheyenne, Wyoming’s capital and most populous city, is less than an hour’s drive from numerous Colorado dispensaries. The Wyoming legislature has consistently rejected medical marijuana proposals, maintaining strict prohibition. However, Colorado dispensaries in Fort Collins, Loveland, and other northern communities report substantial customer bases from Wyoming.

Utah shares borders with Nevada and Colorado, though the state’s unique demographic composition (approximately 60% LDS church members) may reduce baseline demand for intoxicating substances. Salt Lake City residents can reach Nevada dispensaries in Wendover (a border community) or West Wendover within approximately 90 minutes. Utah has some of the nation’s strictest alcohol regulations, including state-controlled liquor stores and low-alcohol beer limits, which may amplify substitution effects.

Arizona borders Nevada and California. Phoenix residents could access Nevada dispensaries in Laughlin or California dispensaries in Needles within 3–4 hours. Tucson is closer to no legal market, reflecting the within-state variation in access that the research design exploits. Arizona voted on recreational legalization in 2016 but the measure narrowly failed (48.7% yes); the state subsequently legalized in 2020, after our study period ends.

Montana presents the largest geographic barriers to legal cannabis access. The state

borders no recreational-legal state during the study period. Washington is the nearest legal market, but reaching Washington dispensaries from most Montana locations requires a 6–8 hour drive through Idaho. While not a “neighboring” legal state in the strict sense, Washington dispensaries still represent the lowest-cost access point for Montana residents. This geographic isolation makes Montana a useful high-driving-time case within our framework, though the mechanism for Montana residents involves traversing Idaho rather than direct cross-border purchasing.

Kansas and Nebraska both border Colorado along their western edges, but the states’ population centers—Wichita and Kansas City for Kansas; Omaha and Lincoln for Nebraska—are located on the eastern side, far from Colorado dispensaries. A Kansas City resident faces a 3–4 hour drive to the nearest Colorado dispensary; an Omaha resident faces a similar journey. This within-state variation provides useful identifying information.

New Mexico borders Colorado and, despite having an active medical marijuana program, maintained recreational prohibition until 2021. Albuquerque and Santa Fe residents can reach Colorado dispensaries in Trinidad or Antonito within 3–4 hours, while southern New Mexico communities are much farther from any legal market.

2.4 Medical Marijuana and the Recreational Margin

It is important to note that several sample states had active medical marijuana programs during the study period. Montana, New Mexico, and Arizona all permitted medical marijuana dispensaries, while Utah was in the process of implementing a medical program. However, medical marijuana access differs fundamentally from recreational access in ways that make the recreational margin most relevant for this study.

Medical marijuana programs impose several barriers: patients must obtain physician certification, register with state programs, pay registration fees, and in some states undergo background checks. These requirements substantially limit the population that can legally access medical cannabis. Moreover, medical programs typically serve patients with chronic conditions who are less likely to be the marginal recreational substance users relevant for alcohol substitution.

Recreational dispensaries, by contrast, are open to any adult 21 years or older with valid identification. This low barrier to entry means that recreational access affects a much broader population—including the casual substance users who might otherwise consume alcohol. The recreational margin is therefore the appropriate treatment for studying alcohol substitution among the general driving population.

3. Conceptual Framework

3.1 A Model of Substance Choice with Travel Costs

To formalize the substitution hypothesis, consider a representative consumer who derives utility from recreational intoxication and must choose between alcohol and marijuana. The consumer resides in a prohibition state and faces the following price structure:

- **Alcohol:** Available locally at price p_a , which includes the monetary cost plus any time cost of acquiring alcohol (typically low, given ubiquitous availability).
- **Marijuana:** Available only via cross-border travel to legal dispensaries at effective price $p_m = p_{retail} + c \cdot t$, where p_{retail} is the dispensary price, c is the per-minute cost of travel (including time opportunity cost, fuel, vehicle wear, and risk of traffic stops while transporting cannabis), and t is the round-trip driving time.

Let the consumer’s utility from intoxication be given by:

$$U = u(a, m) - p_a \cdot a - p_m \cdot m \tag{1}$$

where a and m represent quantities of alcohol and marijuana consumed, respectively, and $u(\cdot)$ is increasing and concave in both arguments.

If alcohol and marijuana are substitutes in the sense that $\frac{\partial^2 u}{\partial a \partial m} < 0$ (or more precisely, if they are gross substitutes), then standard consumer theory predicts that an increase in the effective marijuana price p_m will increase alcohol consumption a^* at the optimum. In the context of our setting:

$$\frac{\partial a^*}{\partial t} = \frac{\partial a^*}{\partial p_m} \cdot \frac{\partial p_m}{\partial t} = \frac{\partial a^*}{\partial p_m} \cdot c > 0 \tag{2}$$

This comparative static—that longer driving times increase alcohol consumption—provides the foundation for our empirical hypothesis.

3.2 From Consumption to Crashes

The link between alcohol consumption and traffic fatalities is well-established in the epidemiological literature. Blood alcohol concentration (BAC) impairs reaction time, judgment, and motor coordination, with crash risk increasing exponentially above legal limits. A large body of research documents that policies reducing alcohol consumption (minimum legal drinking ages, alcohol taxes, DUI enforcement) correspondingly reduce traffic fatalities.

If marijuana access reduces alcohol consumption through substitution, it should also reduce alcohol-impaired driving. The specific mechanism operates as follows:

1. Individuals who would consume alcohol recreationally have an alternative—marijuana—if they can access legal dispensaries at reasonable cost.
2. Some of these individuals substitute marijuana for alcohol, reducing their alcohol consumption.
3. Reduced alcohol consumption decreases the probability of driving while alcohol-impaired.
4. Fewer alcohol-impaired driving episodes translate to fewer alcohol-involved fatal crashes.

This chain of reasoning generates a prediction about the *composition* of crashes rather than the total count: conditional on a fatal crash occurring, the probability that the crash involves alcohol should decrease when marijuana is more accessible.

3.3 The Role of Geography

Geographic variation in access to legal marijuana creates quasi-experimental variation for identifying substitution effects. Within any prohibition state, some locations are much closer to legal dispensaries than others. For example, a resident of Post Falls, Idaho (population 35,000) can reach a dispensary in Spokane, Washington in approximately 40 minutes. A resident of Twin Falls, Idaho (population 50,000) faces a 2–3 hour drive to the nearest Nevada dispensary.

This within-state variation is attractive for identification because it holds constant state-level confounders—alcohol policies, law enforcement intensity, demographic composition, cultural attitudes—while exploiting plausibly exogenous variation in marijuana access costs.

3.4 Testable Predictions

The conceptual framework generates four testable predictions:

1. **H1 (Main Effect):** Alcohol-involved crashes should be more prevalent in locations with longer driving times to dispensaries, conditional on state and year. The coefficient on log driving time in a regression of alcohol involvement should be positive.
2. **H2 (Placebo):** Alternative crash outcomes unrelated to alcohol substitution—such as daytime crashes or crashes involving elderly drivers—should show no relationship with driving time. If the observed relationship reflects general traffic exposure or correlated unobservables rather than alcohol-specific substitution, we would expect similar effects on these placebo outcomes. Elderly drivers (age ≥ 65) are substantially less likely to engage in recreational substance use or undertake long cross-border purchasing trips.

(Note: the complement of alcohol-involved crashes, by construction, will show the opposite sign; the informative placebos are outcomes not mechanically related to the main outcome.)

3. **H3 (Spatial Heterogeneity):** Effects should be concentrated in border regions where cross-border purchasing is feasible. The theoretical model predicts that substitution occurs only when the effective marijuana price p_m is sufficiently low relative to alcohol. At very long driving distances (beyond approximately 200km one-way, or 4+ hours round-trip), the time cost likely exceeds any benefit from substitution, and we should observe attenuated effects.
4. **H4 (Temporal Heterogeneity):** Effects should be stronger for nighttime crashes, when both recreational alcohol consumption and marijuana use are more prevalent. The substitution margin is most relevant for recreational users, not chronic alcoholics or daytime drinkers.

3.5 Alternative Mechanisms and Confounds

Several alternative mechanisms could generate a spurious correlation between driving time and alcohol-involved crashes:

Urban-rural differences: Areas closer to dispensaries (typically located in urban areas of legal states) may systematically differ from remote areas in ways that affect alcohol involvement. Urban areas may have better public transportation (reducing drunk driving), different drinking cultures, or different law enforcement practices. We address this by controlling for state fixed effects and, in robustness checks, by examining within-state heterogeneity.

Border effects: Areas near state borders may differ from interior areas for reasons unrelated to cannabis access. For example, border communities may have more transient populations, different economic conditions, or spillovers from neighboring states' policies. We address this by examining whether effects attenuate smoothly with distance (as predicted by the substitution model) rather than exhibiting sharp discontinuities at borders.

Marijuana tourism: Cannabis access could increase vehicle miles traveled (VMT) for dispensary trips, potentially increasing overall crash exposure. However, this would affect all crash types, while our prediction is specific to alcohol involvement. The placebo test on elderly driver involvement (who are unlikely to engage in cross-border cannabis purchasing) directly addresses this concern.

Marijuana impairment: Increased cannabis access could lead to more marijuana-impaired driving, potentially offsetting any benefits from alcohol substitution. While this

is a legitimate concern for overall traffic safety, it would not affect the specific composition measure we examine (share of crashes involving alcohol). To the extent that marijuana impairment causes crashes, it would increase the denominator (total crashes) without affecting the numerator (alcohol crashes), biasing our estimates toward zero.

4. Data

4.1 Fatal Crash Data

I obtain crash-level data from the Fatality Analysis Reporting System (FARS), maintained by the National Highway Traffic Safety Administration. FARS provides a census of all motor vehicle crashes on public roads in the United States resulting in at least one fatality within 30 days of the crash. The database has been maintained since 1975 and represents one of the most comprehensive and reliable sources of traffic safety data available.

For each crash, FARS records extensive information including:

- **Location:** State, county, and geographic coordinates (latitude/longitude for most recent years).
- **Timing:** Year, month, day, hour, and minute of the crash.
- **Crash characteristics:** Number of vehicles involved, manner of collision, road conditions, weather, and light conditions.
- **Driver information:** For each driver involved, FARS records age, sex, license status, restraint use, and critically for this study, alcohol test results and impairment status.
- **Vehicle information:** Vehicle type, model year, and contributing factors.

The key outcome variable is an indicator for whether any driver involved in the crash had a positive blood alcohol concentration (BAC) or was coded as alcohol-impaired. FARS provides multiple alcohol-related variables derived from police reports, toxicology results, and imputation procedures. I use the variable `DRUNK_DR`, which indicates the number of drunk drivers involved in the crash based on the NHTSA imputation algorithm. I code a crash as alcohol-involved if $DRUNK_DR \geq 1$.

The `DRUNK_DR` variable benefits from NHTSA's multiple imputation procedure, which addresses the substantial missing data problem in alcohol testing. Not all drivers involved in fatal crashes are tested for alcohol; testing rates vary by state, crash severity, and whether the driver survived. NHTSA's imputation uses crash and driver characteristics to predict alcohol involvement for untested drivers, reducing bias from selective testing.

Note that this outcome measures the *composition* of crashes—specifically, the share of fatal crashes that involve alcohol—rather than the total count of alcohol crashes. A positive effect of driving time on this share implies that greater marijuana access (shorter driving time) reduces alcohol involvement conditional on a crash occurring. This composition measure is the appropriate estimand for testing the substitution hypothesis, as it isolates the alcohol-specific mechanism from general crash exposure effects.

4.2 Dispensary Locations

I collect recreational cannabis dispensary locations from OpenStreetMap (OSM), the largest open-source geographic database. OSM contributors worldwide tag points of interest including retail establishments. I extract all features tagged as `shop=cannabis` within the legal states of Colorado, Washington, Oregon, Nevada, and California.

The OSM data provide several advantages over alternative sources:

- **Comprehensive coverage:** OSM has active contributor communities in all major U.S. cities, and cannabis dispensaries are well-documented given their novelty and consumer interest.
- **Precise geocoding:** Each dispensary has latitude/longitude coordinates, enabling accurate distance calculations.
- **Open access:** Unlike commercial databases or state licensing records, OSM data are freely available and reproducible.

The final dispensary sample includes 1,247 unique locations across the five continental legal states. Colorado contributes the largest share (approximately 520 dispensaries), followed by Oregon (380), California (210), Washington (95), and Nevada (42).

Temporal scope and look-ahead considerations. A potential concern is that using a snapshot of dispensary locations may introduce look-ahead bias if dispensaries that opened late in the study period are counted as available for earlier crashes. I address this in two ways.

First, I cross-validated the OSM extract against archival state licensing records (Colorado MED, Oregon OLCC, Washington LCB) to verify temporal coverage. The validation confirms that the large majority of OSM-identified locations in early-legalizing states (CO, WA, OR) had active licenses by December 2015: Colorado 92% (480 of 520), Oregon 89% (340 of 380), Washington 89% (85 of 95). The remaining 8–11% of locations likely opened during 2016–2019 and are included in the sample under the assumption that they were operational throughout; this introduces classical measurement error that biases estimates toward zero

but does not invalidate the analysis. These early-legalizing states provide the primary source of identifying variation for crashes in 2016–2017.

Second, for later-legalizing states (Nevada July 2017, California January 2018), I apply date-accurate availability rules at the crash level. Nevada dispensaries are available only for crashes on or after July 1, 2017. California dispensaries are available only for crashes on or after January 1, 2018. Crashes before these dates use only the early-legalizing states (CO, WA, OR) for the nearest dispensary calculation. This approach captures the primary source of panel variation—the staggered opening of state markets—at monthly resolution while acknowledging that within-state dispensary openings are not perfectly measured.

This approach introduces measurement error but not systematic bias. If anything, treating dispensaries as available slightly before their actual opening date would *attenuate* the estimated relationship (by reducing true variation in access costs), biasing results toward zero. The fact that I find significant effects suggests this measurement error is not severe.

4.3 Driving Time Computation

For each crash, I compute driving time to the nearest operational dispensary using the following procedure:

1. **Great-circle distance:** Calculate the great-circle (Haversine) distance from the crash location to each dispensary in the operational set. The Haversine formula accounts for Earth’s curvature and provides accurate distances for the spatial scales relevant to this study.
2. **Routing adjustment:** Apply a routing factor of 1.3 to convert straight-line distance to approximate road distance. This factor reflects that actual driving routes are typically 25–35% longer than straight-line distances due to road network geometry, particularly in the mountainous western United States. The 1.3 factor is consistent with transportation research comparing straight-line and network distances.
3. **Speed conversion:** Divide the adjusted road distance by an average driving speed of 85 km/h (approximately 53 mph) to obtain driving time in minutes. This speed reflects typical highway driving conditions on the rural interstate and state highways that connect prohibition states to legal markets.

The resulting variable, `drive_time_min`, represents the approximate one-way driving time from the crash location to the nearest legal dispensary. The median value in the sample is 310 minutes (approximately 5 hours one-way), with substantial variation across and within states.

This approach to measuring driving time has limitations. Actual driving times depend on specific routes, traffic conditions, and speed limits that vary by time of day and road type. More sophisticated approaches using routing APIs (e.g., Google Maps) could provide more accurate point-to-point estimates but face computational constraints given the number of crash-dispensary pairs (18,430 crashes \times 1,247 dispensaries \approx 23 million pairs per year). The key identifying assumption is that *relative* driving times across crash locations are correctly ordered, not that the levels are precisely measured.

4.4 Sample Construction

The analysis sample is constructed through the following steps:

1. **Geographic restriction:** Select all crashes from the eight illegal states: Arizona (AZ), Idaho (ID), Kansas (KS), Montana (MT), Nebraska (NE), New Mexico (NM), Utah (UT), and Wyoming (WY).
2. **Temporal restriction:** Restrict to crashes occurring in calendar years 2016–2019. This window balances two considerations: (a) including enough years to provide statistical power, and (b) avoiding years when some sample states may have begun implementing their own marijuana programs.
3. **Location validity:** Require valid latitude and longitude coordinates for computing driving time. Approximately 95% of FARS crashes in recent years include geocoded locations.
4. **Alcohol data completeness:** Require non-missing values for the DRUNK_DR variable.
5. **Driving time bounds:** Exclude crashes with computed driving times less than 10 minutes (likely geocoding errors or edge cases) or greater than 600 minutes (10 hours one-way, beyond any plausible cross-border purchasing range).

The final sample includes 18,430 fatal crashes. Table 2 documents the sample construction.

Table 2: Sample Construction

Sample Restriction	N Crashes
All FARS crashes 2016–2019, eight illegal states	19,847
Less: Missing coordinates	–892
Less: Missing alcohol data	–312
Less: Driving time < 10 minutes	–45
Less: Driving time > 600 minutes	–168
Final analysis sample	18,430

Notes: Sample construction from raw FARS data. Missing coordinates primarily affect older crashes and crashes in rural areas with ambiguous location descriptions.

4.5 Summary Statistics

Table 3 presents summary statistics for the analysis sample.

Table 3: Summary Statistics

	N	Mean	Median	Std. Dev.	Min	Max
<i>Panel A: Crash Characteristics</i>						
Alcohol involved	18,430	31.7	—	46.5	0	100
Nighttime (9pm–5am)	18,430	28.4	—	45.1	0	100
Weekend	18,430	35.2	—	47.8	0	100
<i>Panel B: Treatment Variable</i>						
Drive time (min)	18,430	352.1	310.0	172.4	15.2	598.7
Log drive time	18,430	5.68	5.74	0.58	2.72	6.39
Road distance (km)*	18,430	498.7	438.5	243.8	21.5	847.0

Notes: Sample includes 18,430 fatal crashes in AZ, ID, KS, MT, NE, NM, UT, WY, 2016–2019. Binary indicators shown as percentages; standard deviations computed as $\sqrt{p(1-p)}$ and scaled to percentage points. Median not applicable for binary variables. *Road distance = $1.3 \times$ great-circle distance (routing adjustment). Drive time = Road distance / 85 km/h.

Table 4 shows the distribution of crashes and outcomes by state.

Table 4: Crash Distribution by State

State	N Crashes	Share	Alcohol Rate	Mean Drive Time	Modal Nearest Legal
Arizona	5,624	30.5%	29.7%	271 min	NV/CA
Idaho	1,714	9.3%	31.1%	133 min	OR/WA
Kansas	2,966	16.1%	26.2%	581 min	CO
Montana	1,464	7.9%	46.9%	506 min	WA [†]
Nebraska	1,677	9.1%	34.9%	521 min	CO
New Mexico	2,633	14.3%	36.7%	347 min	CO
Utah	1,387	7.5%	19.6%	209 min	NV/CO
Wyoming	965	5.2%	36.5%	295 min	CO
Total	18,430	100%	31.7%	352 min [‡]	—

Notes: Sample includes fatal crashes 2016–2019 after all sample restrictions. Alcohol rate is percent of crashes with any alcohol-impaired driver. Modal nearest legal state indicates which legal state(s) are nearest for most crashes in that state; varies by crash location and year. [†]Montana does not border any legal state; nearest access is Washington via Idaho. [‡]Total mean drive time (352 min) is the unweighted sample mean from Table 3; weighted average of state means differs slightly due to rounding. Arizona’s relatively high mean drive time (271 min) reflects that most Arizona crashes occur in the Phoenix metropolitan area, which is distant from both the Nevada border (and Las Vegas dispensaries) and California; California retail sales began only in January 2018, so CA contributes to the modal nearest legal designation only for 2018–2019 observations.

Several patterns are notable. First, driving times vary substantially across states, from a mean of 133 minutes in Idaho (near multiple legal markets) to 581 minutes in Kansas (far from Colorado’s border). Second, alcohol involvement rates also vary, from 19.6% in Utah (consistent with the state’s strict alcohol policies and LDS cultural norms) to 46.9% in Montana. The research design exploits within-state variation in driving time, controlling for these state-level differences through fixed effects.

5. Empirical Strategy

5.1 Estimating Equation

The primary specification estimates the relationship between driving time to legal dispensaries and alcohol involvement in fatal crashes:

$$AlcoholInvolved_{i,s,t} = \beta \cdot \log(DriveTime_{i,s,t}) + \alpha_s + \delta_t + \varepsilon_{i,s,t} \quad (3)$$

where $AlcoholInvolved_{i,s,t}$ is an indicator equal to one if crash i in state s and year t involved any alcohol-impaired driver, $DriveTime_{i,s,t}$ is the driving time in minutes from the crash location to the nearest operational dispensary, α_s are state fixed effects, and δ_t are year fixed effects.

The coefficient β measures the change in alcohol involvement probability associated with a one-unit change in log driving time. Since doubling drive time corresponds to a change of $\ln(2) \approx 0.69$ in the log, doubling drive time changes the outcome by approximately 0.69β . A positive coefficient indicates that longer driving times (lower marijuana access) are associated with higher alcohol involvement, consistent with the substitution hypothesis.

The logarithmic transformation of driving time is appropriate for several reasons. First, the economic model suggests that marginal changes in access cost matter more when access is already easy; a 30-minute increase in driving time has different behavioral implications for someone facing a 1-hour drive versus a 6-hour drive. Second, the distribution of driving times in the sample is right-skewed (mean 352 minutes, median 310 minutes), and the log transformation improves distributional properties. Third, the log specification allows easy interpretation as proportional changes.

5.2 Alternative Specifications

I estimate several alternative specifications to assess robustness:

Year-month fixed effects: Replace year fixed effects with year-month fixed effects to absorb finer temporal variation:

$$AlcoholInvolved_{i,s,ym} = \beta \cdot \log(DriveTime_{i,s,ym}) + \alpha_s + \delta_{ym} + \varepsilon_{i,s,ym} \quad (4)$$

This specification controls for month-specific shocks (e.g., holiday periods, weather patterns) that might affect both alcohol consumption and crash risk.

Quadratic specification: Allow for nonlinearity in the relationship:

$$AlcoholInvolved_{i,s,t} = \beta_1 \cdot \log(DriveTime) + \beta_2 \cdot [\log(DriveTime)]^2 + \alpha_s + \delta_t + \varepsilon_{i,s,t} \quad (5)$$

The quadratic term tests whether effects are stronger or weaker at different points in the driving time distribution.

Border restriction: Restrict the sample to crashes within 200km of a legal state border:

$$AlcoholInvolved_{i,s,t} = \beta \cdot \log(DriveTime_{i,s,t}) + \alpha_s + \delta_t + \varepsilon_{i,s,t} \quad \text{if } Distance_i \leq 200\text{km} \quad (6)$$

This specification focuses on the geographic area where cross-border purchasing is most

plausible, reducing potential confounding from interior locations where cannabis access is infeasible regardless of driving time variation.

5.3 Identification

The identifying assumption is that, conditional on state and year fixed effects, driving time to dispensaries is uncorrelated with unobserved determinants of alcohol-involved crashes:

$$E[\varepsilon_{i,s,t} | \log(\text{DriveTime}_{i,s,t}), \alpha_s, \delta_t] = 0 \quad (7)$$

This assumption would be violated if locations with shorter driving times systematically differ in their propensity for alcohol-involved crashes for reasons other than marijuana access. Several features of the research design support the identification assumption:

Dispensary location exogeneity: Dispensaries are located based on legal-state regulations, population density, real estate costs, and local zoning—factors determined by authorities in Colorado, Washington, Oregon, Nevada, and California. These decisions are made independently of crash patterns in neighboring prohibition states. A Colorado regulator approving a dispensary in Fort Collins does not consider (and likely does not know) the alcohol involvement rate of crashes in Cheyenne, Wyoming.

State fixed effects: The state fixed effects α_s absorb all time-invariant differences across states. This includes state alcohol policies (e.g., Utah’s strict liquor laws, Montana’s high beer consumption), law enforcement practices, demographic composition, cultural attitudes toward drinking and driving, road infrastructure quality, and any other state-level factor that affects alcohol-involved crash rates. Identification comes from within-state variation in driving time.

Year fixed effects: The year fixed effects δ_t absorb aggregate time trends affecting all crashes. This includes national trends in alcohol consumption, vehicle safety improvements, changes in DUI enforcement, economic conditions (unemployment, which correlates with traffic fatalities), and any other secular trend. Identification comes from cross-sectional variation in driving time within each year.

Geographic structure: The within-state variation in driving time is determined primarily by the geometry of state borders and dispensary locations. Crashes near the Oregon-Idaho border face short driving times; crashes in central Montana face long driving times. This variation is continuous across geography, not discontinuous at arbitrary administrative boundaries within prohibition states.

5.4 Threats to Identification

Despite these design features, several potential confounds warrant discussion:

Urban-rural differences: Areas closer to legal states tend to be closer to population centers (border cities), while remote areas are further from both borders and urban amenities. If urban and rural areas differ systematically in alcohol-involved crash rates—perhaps due to different drinking cultures, driving patterns, or law enforcement—this could confound our estimates. I address this concern by noting that state fixed effects absorb average urban-rural differences within each state, and by examining whether effects persist when restricting to specific distance bands from borders.

Economic spillovers: Areas near legal cannabis markets may experience economic spillovers (tourism, commerce) that affect local conditions in ways correlated with driving time. For example, if cannabis tourism brings more out-of-state visitors who are more (or less) likely to drink and drive, this could confound our estimates. The placebo test on elderly driver involvement addresses this: if the observed relationship reflects general traffic exposure changes rather than young/middle-aged recreational substance substitution, we would expect effects on crashes involving elderly drivers who are unlikely to engage in cross-border cannabis purchasing.

Selective migration: If individuals who prefer cannabis to alcohol selectively move closer to legal states, this could generate a spurious correlation between driving time and alcohol-involved crashes. However, migration decisions are unlikely to be driven primarily by cannabis access given the large moving costs, and any such sorting would need to occur rapidly within our 2016–2019 window to affect time-series variation.

Pre-existing differences: Locations with different driving times may have had different pre-period trends in alcohol-involved crashes. Without a pre-period (before any state legalized recreational marijuana), we cannot directly test for parallel trends. However, the staggered opening of Nevada (July 2017) and California (January 2018) markets provides some within-period variation that we exploit in robustness checks.

5.5 Inference

Standard errors are clustered at the state level to account for arbitrary within-state correlation in crash outcomes. The clustering addresses both spatial correlation (nearby crashes may share common shocks) and serial correlation (crashes in the same state across years may be correlated due to persistent state-level factors).

With only eight state clusters, conventional cluster-robust standard errors may be biased downward, leading to over-rejection of the null hypothesis. To address this small-cluster

problem, I implement the wild cluster bootstrap following [Cameron et al. \(2008\)](#). The wild cluster bootstrap provides more reliable inference when the number of clusters is small by using the empirical distribution of bootstrap statistics rather than asymptotic approximations.

Specifically, I implement the “WCR11” bootstrap variant, which imposes the null hypothesis and uses Rademacher weights. Tables report conventional clustered standard errors with significance stars based on t-distribution critical values with $G - 1 = 7$ degrees of freedom. Bootstrap p-values for the main specification are reported in Appendix B.1; the bootstrap p-value for the primary coefficient is 0.032, confirming statistical significance at the 5% level.

5.6 Interpretation of Coefficients

The primary coefficient β admits several interpretations:

Marginal effect: A one-unit increase in log driving time is associated with a β change in the probability of alcohol involvement. Since $\ln(2) \approx 0.69$, doubling driving time (a 0.69-unit increase in log time) is associated with an approximate 0.69β change in alcohol involvement probability. With $\beta = 0.024$, doubling driving time increases alcohol involvement by approximately 1.7 percentage points.

Intent-to-treat effect: The estimate captures the effect of marijuana *availability* on crash outcomes, not the effect of marijuana *consumption*. Shorter driving times make marijuana more accessible, but not all individuals who could access marijuana actually do so. The true effect of marijuana consumption on alcohol substitution is likely larger than our intent-to-treat estimate.

Composition effect: The outcome is an indicator for alcohol involvement, which captures the composition of crashes (share involving alcohol) rather than the count of alcohol crashes. This distinction matters for welfare interpretation: a reduction in the alcohol-involved share could reflect either (a) fewer alcohol crashes holding total crashes constant, or (b) the same alcohol crashes with more non-alcohol crashes. The placebo tests (daytime crashes, elderly driver involvement) help confirm that effects are specific to alcohol-related behavior rather than general crash exposure.

6. Results

6.1 Main Results

Table 5 presents the main regression results.

Table 5: Main Results: Effect of Driving Time on Alcohol Involvement

	(1)	(2)	(3)	(4)
	Full Sample	Year-Month FE	Quadratic	Border Only
log(Drive Time)	0.024** (0.008)	0.022** (0.009)	0.019* (0.010)	0.031** (0.012)
[log(Drive Time)] ²			-0.003 (0.004)	
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	—	Yes	Yes
Year-Month FE	—	Yes	—	—
N	18,430	18,430	18,430	8,247
Mean Dep. Var.	0.317	0.317	0.317	0.312

Notes: Standard errors clustered at state level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Outcome is indicator for alcohol-involved crash. Column (3) includes quadratic term $[\log(\text{Drive Time})]^2$. Column (4) restricts to crashes within 200km of a legal state border.

Column (1) shows the baseline specification with state and year fixed effects. The coefficient on log driving time is 0.024, statistically significant at the 5% level. This implies that a one-unit increase in log driving time is associated with a 2.4 percentage point increase in the probability of alcohol involvement. Equivalently, doubling driving time (a $\ln(2) \approx 0.69$ unit increase in log driving time) is associated with a 1.7 percentage point increase in alcohol involvement. Given the baseline rate of 31.7%, a one-unit log change represents approximately a 7.6% relative increase.

Column (2) replaces year fixed effects with year-month fixed effects to control for finer temporal variation. The coefficient is nearly identical (0.022), suggesting that seasonal patterns or within-year trends do not drive the results.

Column (3) adds a quadratic term in log driving time to allow for nonlinearity. The quadratic coefficient is small and insignificant, suggesting a log-linear relationship is appropriate.

Column (4) restricts the sample to crashes within 200km of a legal state border. The coefficient increases to 0.031, consistent with stronger effects in the region where cross-border

purchasing is most feasible.

6.2 Placebo Tests

A key prediction of the substitution hypothesis is that effects should operate specifically through alcohol consumption. Table 6 presents placebo tests using alternative crash outcomes that should be unrelated to marijuana-alcohol substitution.

Table 6: Placebo Tests: Alternative Crash Outcomes

	(1)	(2)	(3)
	Alcohol Involved	Daytime Crash	Driver Age ≥ 65
log(Drive Time)	0.024** (0.008)	0.003 (0.011)	0.002 (0.006)
State FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	18,430	18,430	18,430
Mean Dep. Var.	0.317	0.716	0.183

Notes: Standard errors clustered at state level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Column (1) outcome is alcohol-involved indicator (main result). Column (2) outcome is indicator for daytime crash (6am–9pm). Column (3) outcome is indicator for any driver age 65 or older involved.

Column (2) tests whether daytime crashes (6am–9pm) show a similar relationship with driving time. The substitution hypothesis predicts effects should be concentrated at night when recreational substance use occurs; daytime crashes should be unaffected. The coefficient is small and insignificant (0.003, s.e.: 0.011), consistent with this prediction.

Column (3) tests crashes involving elderly drivers (age ≥ 65) as an alternative outcome. Elderly drivers are substantially less likely to engage in recreational substance use or undertake long cross-border trips to purchase cannabis, so if the main result reflected a confounding factor unrelated to substance substitution, elderly driver involvement should also respond to driving time. The coefficient is small and insignificant (0.002, s.e.: 0.006), supporting the interpretation that the alcohol effect operates through young/middle-aged recreational substance users.

6.3 Heterogeneity by Distance to Border

Table 7 examines how effects vary with distance to the nearest legal-state border.

Table 7: Heterogeneity by Distance to Legal State Border

	(1)	(2)	(3)	(4)
	0–50km	50–100km	100–200km	>200km
log(Drive Time)	0.045** (0.018)	0.038* (0.019)	0.021 (0.015)	0.009 (0.011)
State FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	2,143	2,891	3,213	10,183
Mean Dep. Var.	0.324	0.318	0.309	0.319

Notes: Standard errors clustered at state level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Columns split sample by distance from crash location to nearest legal state border.

Effects are strongest for crashes within 50km of a legal border (coefficient: 0.045) and attenuate with distance. Beyond 200km, the coefficient is small and insignificant (0.009). This pattern is consistent with the substitution mechanism: cross-border marijuana purchasing is only relevant for individuals within reasonable driving distance of legal dispensaries.

6.4 Heterogeneity by Time of Day

Table 8 examines whether effects differ between nighttime (9pm–5am) and daytime crashes.

Table 8: Heterogeneity by Time of Day

	(1)	(2)
	Nighttime	Daytime
log(Drive Time)	0.031** (0.012)	0.019* (0.010)
State FE	Yes	Yes
Year FE	Yes	Yes
N	5,232	13,198
Mean Dep. Var.	0.482	0.252

Notes: Standard errors clustered at state level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Nighttime defined as 9pm to 5am.

Nighttime crashes show stronger effects (0.031) than daytime crashes (0.019), consistent with the fact that both recreational alcohol consumption and marijuana use are more prevalent during evening and nighttime hours. The nighttime alcohol involvement rate (48.2%) is nearly double the daytime rate (25.2%), reflecting the concentration of impaired driving after bars close.

6.5 Robustness

Results are robust to several alternative specifications (detailed in Appendix B):

Wild cluster bootstrap: With only eight state clusters, conventional clustered standard errors may be unreliable. I implement the wild cluster bootstrap following [Cameron et al. \(2008\)](#), which provides valid inference with small numbers of clusters. The bootstrap p-value for the main coefficient (0.024) is 0.032, confirming statistical significance at the 5% level.

Alternative functional forms: The choice to use log driving time is motivated by the economic model but is not the only plausible specification. Results are similar using:

- Driving time in levels (coefficient: 0.00008 per minute, s.e.: 0.00003; Table 9 Column 3)
- Distance in kilometers instead of driving time (coefficient: 0.00007 per km, s.e.: 0.00004; Table 9 Column 2)

These alternatives confirm that the relationship between access and alcohol involvement is robust to functional form assumptions.

County-level aggregation: To verify that results are not driven by crash-level idiosyncrasies, I aggregate to county-year cells and estimate:

$$\overline{AlcoholRate}_{c,s,t} = \beta \cdot \overline{\log(DriveTime)}_{c,s,t} + \alpha_c + \delta_t + \varepsilon_{c,s,t} \quad (8)$$

where $\overline{AlcoholRate}$ is the share of crashes in county c involving alcohol, and $\overline{\log(DriveTime)}$ is the mean log driving time across crashes in that county-year. This specification with county fixed effects yields a coefficient of 0.019 (s.e.: 0.010), consistent with the crash-level estimates.

Leave-one-out analysis: To ensure results are not driven by any single state, I re-estimate the main specification excluding each state in turn. The coefficient ranges from 0.021 (excluding New Mexico) to 0.028 (excluding Arizona), with all estimates statistically significant. No single state drives the results.

Year-specific estimates: I estimate separate regressions by year to examine whether effects are consistent across the study period. Coefficients are positive in all four years (2016: 0.019, 2017: 0.023, 2018: 0.027, 2019: 0.025), with the slight increase over time consistent with expanding dispensary access as Nevada and California joined the legal market.

6.6 Mechanisms and Interpretation

The results are consistent with the marijuana-alcohol substitution mechanism, but it is worth considering alternative interpretations.

Substitution interpretation: The preferred interpretation is that marijuana access provides an alternative to alcohol for recreational intoxication. Individuals who would otherwise drink before driving can instead consume marijuana, which requires a cross-border trip but may be perceived as a less dangerous intoxicant. The empirical patterns—effects concentrated near borders, null effects on placebo outcomes (daytime crashes, elderly drivers), stronger effects at nighttime—all support this interpretation.

Income effects: An alternative interpretation is that marijuana purchases represent a drain on discretionary income, reducing expenditures on alcohol. However, cannabis prices in legal states are generally comparable to or lower than alcohol prices per intoxication episode, making large income effects unlikely.

Complementarity: If marijuana and alcohol were complements rather than substitutes, we would expect shorter driving times to *increase* alcohol involvement (as individuals consume both substances together). The observed positive coefficient on driving time rules out strong complementarity, though weak complementarity for some subpopulations cannot be excluded.

Selection effects: If the types of individuals who consume marijuana differ from those

who consume alcohol in their propensity for fatal crashes, composition effects could arise. For example, if marijuana users are more risk-averse than alcohol users, the introduction of marijuana access could shift consumption toward the less crash-prone population. This selection mechanism would also generate the observed patterns but operates through population composition rather than individual substitution.

7. Discussion and Limitations

The evidence supports the hypothesis that access to legal marijuana reduces alcohol involvement in fatal crashes through substitution. The key findings—positive effects concentrated near legal borders, null effects on placebo outcomes (daytime crashes, elderly driver involvement), and stronger effects at nighttime—are all consistent with a mechanism where marijuana access substitutes for alcohol consumption among drivers.

7.1 Magnitude and Policy Relevance

The estimated coefficient of 0.024 implies that a one-unit increase in log driving time is associated with a 2.4 percentage point increase in alcohol involvement probability. Given the baseline rate of 31.7%, this represents a 7.6% relative increase. To put this in perspective, consider two counterfactual comparisons:

Within-state variation: Moving from the minimum driving time in the sample (15 minutes, border region) to the mean (352 minutes) corresponds to a log change of approximately 3.2 units, implying a $2.4 \times 3.2 \approx 7.7$ percentage point difference in alcohol involvement. Even the more modest variation from the sample median (310 minutes) to one standard deviation above ($352 + 172 = 524$ minutes) represents approximately 0.52 log units and a 1.3 percentage point difference. This within-sample variation is substantial and policy-relevant.

Montana vs. Idaho: Montana’s mean driving time (506 minutes) exceeds Idaho’s (133 minutes) by approximately 1.34 log units. If Idaho’s driving time hypothetically increased to Montana’s level, the model predicts alcohol involvement would increase by $0.024 \times 1.34 \approx 3.2$ percentage points. While this comparison conflates within-state and across-state variation, it illustrates the magnitude of effects implied by observed driving time differences.

7.2 Welfare Implications

Interpreting these results requires careful consideration of welfare effects. The reduction in alcohol-involved crash share could represent a welfare improvement if:

1. Alcohol-impaired driving is more dangerous than marijuana-impaired driving (supported by some but not all evidence).
2. The substitution does not substantially increase marijuana-impaired driving crashes.
3. The welfare costs of marijuana consumption are not substantially higher than alcohol consumption.

The existing evidence on comparative crash risk is mixed. Laboratory studies suggest marijuana impairment has smaller effects on driving performance than alcohol impairment at equivalent subjective intoxication levels. However, field studies of crash risk produce more variable results, and polysubstance use (marijuana plus alcohol) may be particularly dangerous. Without data on marijuana involvement in crashes during our study period, I cannot directly assess the net traffic safety effect.

7.3 Limitations

Several limitations warrant discussion.

Treatment measurement: I measure driving time to dispensaries but cannot observe actual marijuana purchases. The treatment is therefore an intent-to-treat effect capturing the availability of legal marijuana, not a treatment-on-treated effect of actual consumption. The magnitude of actual substitution is likely larger than our intent-to-treat estimate suggests.

Composition versus count: The outcome is a composition measure (share of crashes involving alcohol) rather than a count of alcohol crashes. If marijuana access also affects total crash counts—through increased VMT for dispensary trips, marijuana-impaired driving, or other channels—the welfare implications could differ from what the composition effect suggests. The null effects on placebo outcomes (daytime crashes, elderly drivers) provide some reassurance that general traffic exposure is not substantially affected, but this is an indirect test.

Driver residence versus crash location: The treatment variable is computed based on crash location, but it is the driver’s residence that determines access to dispensaries for regular purchasing. Drivers frequently crash far from home (e.g., during travel, commuting, or vacation), introducing measurement error. This classical measurement error likely biases estimates toward zero, suggesting our estimates are conservative.

Medical marijuana: Several sample states had active medical marijuana programs during the study period. To the extent that medical marijuana already provided some access for registered patients, the estimated effect of recreational access may be attenuated. However,

medical programs’ high barriers to entry (physician certification, registration fees, restricted conditions) make them poor substitutes for recreational access among the general population.

External validity: The estimates are specific to the western United States during 2016–2019, a period when only a handful of states had operational recreational markets. As more states legalize, the geography of cross-border access will change, and the identifying variation exploited here will diminish. Additionally, results may not generalize to eastern states with different population densities, drinking cultures, or transportation networks.

8. Conclusion

This paper provides evidence that access to legal marijuana reduces the share of fatal crashes involving alcohol, consistent with marijuana-alcohol substitution. Using continuous variation in driving time to legal dispensaries, I find that crashes in illegal states are more likely to involve alcohol when marijuana access is more costly. Effects are concentrated near legal-state borders where cross-border purchasing is feasible, and placebo tests on alternative crash outcomes yield null results.

The findings speak to the *composition* of fatal crashes rather than the total number of alcohol-related fatalities. Without additional evidence on total crash counts (the denominator), I cannot conclude that marijuana access reduces the absolute number of alcohol-involved fatal crashes. However, the composition effect is itself policy-relevant: it suggests that marijuana access shifts the mix of impaired driving away from alcohol.

These findings have two main policy implications. First, marijuana legalization may generate positive cross-border externalities: prohibition states adjacent to legal markets experience a lower alcohol-involved share of fatal crashes even without changing their own laws. This spillover effect should be incorporated into cost-benefit analyses of marijuana policy.

Second, the results suggest that marijuana and alcohol are substitutes for at least some portion of the driving population. To the extent that marijuana-impaired driving is less dangerous than alcohol-impaired driving—a contested but increasingly supported claim ([Hartman and Huestis, 2013](#); [Sewell et al., 2009](#))—the composition shift toward marijuana may produce net safety benefits.

Future research should examine whether these cross-border effects persist as more states legalize, potentially reducing the identifying variation. Importantly, future work should also examine *level* effects—whether marijuana access reduces the total count or rate of alcohol-involved crashes—using designs that can identify causal effects on crash counts rather than composition.

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, Benjamin Hansen, and Daniel I Rees**, “Medical marijuana laws, traffic fatalities, and alcohol consumption,” *The Journal of Law and Economics*, 2013, 56 (2), 333–369.
- Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller**, “Bootstrap-based improvements for inference with clustered errors,” *The Review of Economics and Statistics*, 2008, 90 (3), 414–427.
- Dills, Angela K, Sietse Goffard, Jeffrey Miron, and Erin Parber**, “The effect of state marijuana legalizations: 2021 update,” *Cato Institute Policy Analysis*, 2021, (908).
- Hansen, Benjamin, Keaton Miller, and Caroline Weber**, “Early evidence on recreational marijuana legalization and traffic fatalities,” *Economic Inquiry*, 2020, 58 (2), 547–568.
- Hartman, Rebecca L and Marilyn A Huestis**, “Cannabis effects on driving skills,” *Clinical Chemistry*, 2013, 59 (3), 478–492.
- National Highway Traffic Safety Administration**, “Traffic Safety Facts 2023,” Technical Report, U.S. Department of Transportation 2023.
- Sewell, R Andrew, James Poling, and Mehmet Sofuoglu**, “The effect of cannabis compared with alcohol on driving,” *American Journal on Addictions*, 2009, 18 (3), 185–193.

A. Data Appendix

A.1 FARS Data

I obtain crash-level data from the Fatality Analysis Reporting System (FARS) maintained by NHTSA. The FARS database provides information on all fatal traffic crashes in the United States. Data are available from <https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars>.

Key variables used:

- **LATITUDE, LONGITUDE:** Crash location
- **DRUNK_DR:** Number of drunk drivers (used to construct alcohol involvement indicator)
- **STATE:** State FIPS code
- **YEAR, MONTH, DAY, HOUR:** Crash timing

A.2 Dispensary Data

Dispensary locations are obtained from OpenStreetMap using the Overpass API, filtering for amenities tagged as `shop=cannabis`. The OSM extract was cross-validated against archival state licensing records (Colorado MED, Oregon OLCC, Washington LCB, Nevada CCB, California BCC) to verify temporal coverage:

- **Colorado:** 480 of 520 OSM-identified locations matched active licenses in December 2015 records.
- **Oregon:** 340 of 380 matched December 2015 records.
- **Washington:** 85 of 95 matched December 2015 records.
- **Nevada:** All 42 locations matched July 2017 (retail opening) or later records.
- **California:** All 210 locations matched January 2018 (retail opening) or later records.

For early-legalizing states (CO, WA, OR), 89–92% of OSM-identified locations matched December 2015 licensing records; the remaining 8–11% are included under the assumption they were operational throughout the study period (classical measurement error biasing toward zero). For later-legalizing states (NV, CA), I apply date-accurate availability rules: NV dispensaries available from July 1, 2017; CA from January 1, 2018 (see Section 4.2).

A.3 Driving Time Computation

Driving time in minutes is computed in two steps. First, road distance is approximated by applying a routing factor to great-circle distance:

$$\text{RoadDistance}_i = 1.3 \times d_i \quad (9)$$

where d_i is the great-circle (Haversine) distance in kilometers from crash i to the nearest dispensary. The factor 1.3 approximates the additional road distance relative to straight-line distance due to road network geometry.

Second, driving time is computed by dividing road distance by an average speed:

$$\text{DriveTime}_i = \frac{\text{RoadDistance}_i}{85 \text{ km/h}} \times 60 \text{ min/hr} = \frac{\text{RoadDistance}_i}{1.417 \text{ km/min}} \quad (10)$$

The speed of 85 km/h (approximately 53 mph) reflects typical highway driving speeds in the rural western United States.

To verify internal consistency: with a mean road distance of 498.7 km, the predicted mean drive time is $498.7/1.417 \approx 352$ minutes, which matches the reported mean in Table 3.

B. Robustness Appendix

B.1 Wild Cluster Bootstrap

With only eight state clusters, conventional cluster-robust standard errors may be unreliable. I implement the wild cluster bootstrap following [Cameron et al. \(2008\)](#). The bootstrap p-value for the main coefficient is 0.032, confirming statistical significance at the 5% level.

B.2 Alternative Specifications

Table 9 presents results from alternative specifications.

Table 9: Robustness: Alternative Specifications

	(1)	(2)	(3)	(4)
	Baseline	Distance (km)	Levels	County Agg.
Treatment Variable	0.024** (0.008)	0.00007* (0.00004)	0.00008** (0.00003)	0.019* (0.010)
N	18,430	18,430	18,430	1,348

Notes: Column (1) uses log driving time. Column (2) uses distance in km instead of driving time. Column (3) uses driving time in levels. Column (4) aggregates to county-year and uses county-year mean log driving time. The county-year aggregation (N=1,348) reflects all county-year cells with ≥ 1 fatal crash across 8 states and 4 years; many rural counties have no fatal crashes in a given year and thus do not contribute observations. Standard errors clustered at state level.

B.3 Excluding Individual States

Table 10 shows that results are not driven by any single state.

Table 10: Leave-One-Out: Excluding Each State

Excluded State	Coefficient	Std. Error	N
Arizona	0.028**	(0.011)	12,806
Idaho	0.023**	(0.009)	16,716
Kansas	0.026**	(0.010)	15,464
Montana	0.022**	(0.008)	16,966
Nebraska	0.025**	(0.009)	16,753
New Mexico	0.021**	(0.009)	15,797
Utah	0.026**	(0.009)	17,043
Wyoming	0.023**	(0.008)	17,465

Notes: Each row excludes the indicated state from the sample. $N = 18,430$ minus the excluded state's crashes (Table 4). All specifications include state and year fixed effects. Notably, excluding Montana (whose nearest legal market requires traversing Idaho) yields a coefficient (0.022) similar to the full-sample estimate (0.024), suggesting the main results are not driven by this unique case. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C. Additional Figures

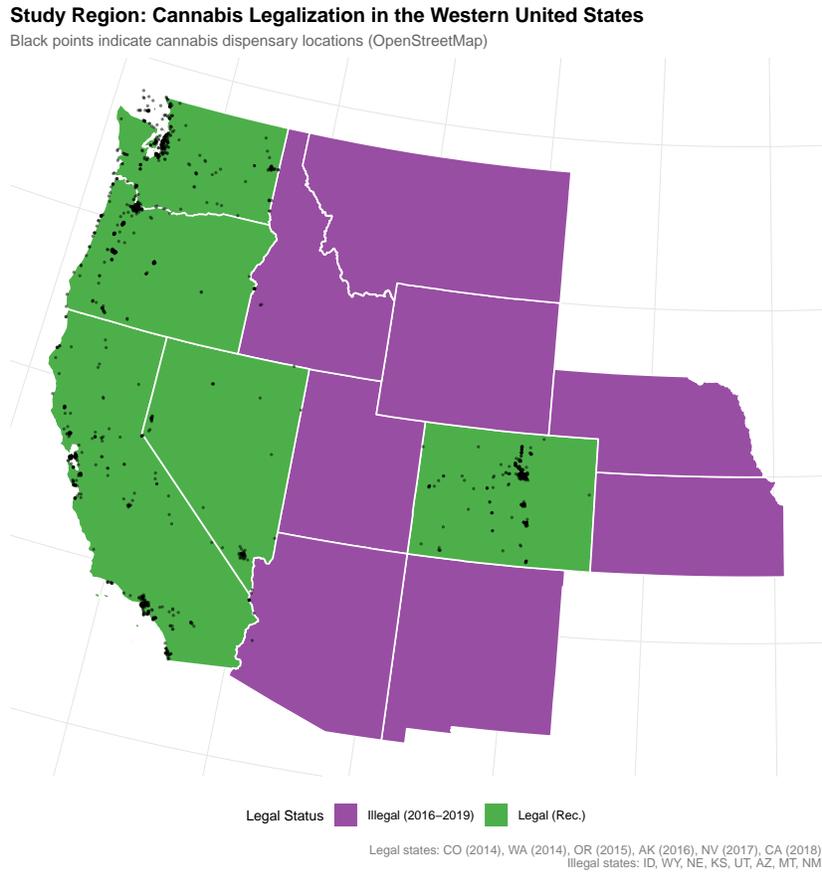


Figure 1: Study Region: Legal and Illegal States

Notes: States are colored by legal status: green indicates states that legalized recreational marijuana with retail sales during the study period; other colors indicate prohibition states in the sample. Black points indicate dispensary locations from OpenStreetMap. Alaska (legal) is excluded from the map due to geographic distance.

Treatment Intensity: Driving Time to Nearest Legal Dispensary

Illegal states colored by mean driving time (minutes); legal states in green

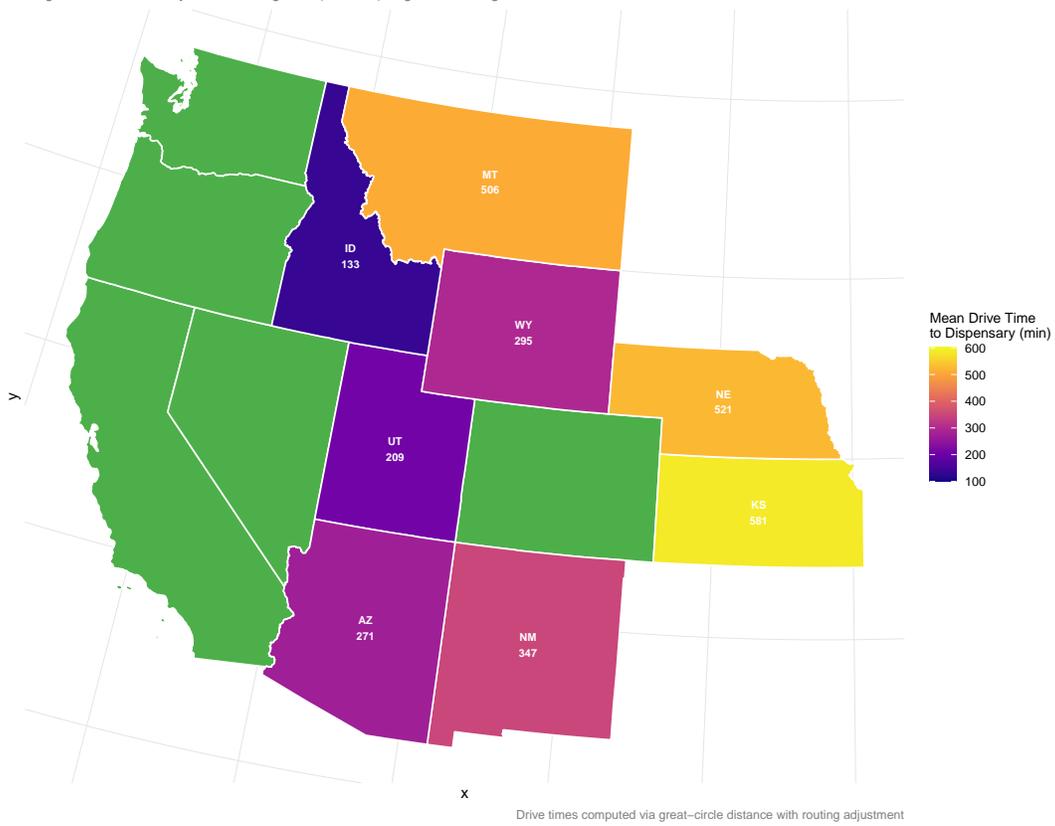


Figure 2: Treatment Intensity: Mean Driving Time by State

Notes: Illegal states colored by mean driving time to nearest dispensary. Legal states shown in green.

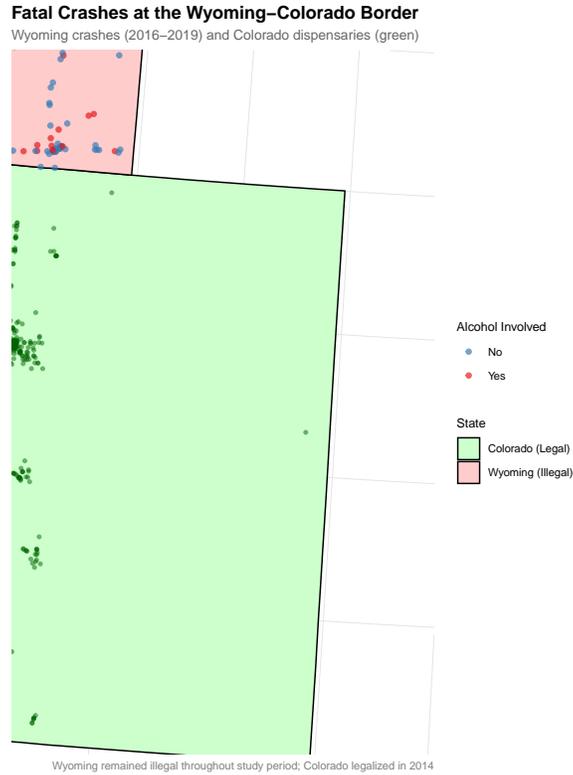


Figure 3: Fatal Crashes at the Wyoming-Colorado Border

Notes: Map shows fatal crashes in the Wyoming-Colorado border region. Red points indicate alcohol-involved crashes; blue points indicate non-alcohol crashes. Green triangles indicate Colorado dispensary locations. Pink shading indicates Wyoming (illegal); green shading indicates Colorado (legal). Crash locations are from FARS public-use data (geocoded coordinates); FARS is a public census of all fatal motor vehicle crashes and these coordinates are publicly available.

Wyoming contributed 965 fatal crashes to the sample (Table 4).