Lowering the Legal Limit: Evidence from Utah's 0.05 BAC Law

Sean Kiely, *NBER**

June 2026

Abstract

In December 2018, Utah became the first U.S. state to lower its legal blood alcohol concentration (BAC) limit from 0.08 to 0.05 grams per deciliter. Using synthetic difference-in-differences (SDID) on a balanced state-year panel covering all 50 states from 1994 to 2023, I estimate the causal effect of the law on traffic fatalities. The law reduces traffic fatalities per 100 million vehicle miles traveled by approximately 12 percent, marginally significant at the 10 percent level under permutation-based placebo inference, a result that holds up when COVID-affected years are excluded. Heterogeneity analysis finds a weakly statistically significant weekday reduction, consistent with deterrence concentrated among moderate weekday drinkers in the newly criminalized 0.05–0.08 BAC range, a pattern likely reflecting Utah's large Latter-day Saints population, though this interpretation remains exploratory. A welfare analysis of the policy implies an average of approximately 40 lives and $530 million in avoided fatality costs per year in Utah. However, as the main effect is only significant at the 10% level, no positive lower bound on welfare can be established.

Keywords: blood alcohol concentration; traffic safety policy; drunk driving; synthetic difference-in-differences; vehicle fatalities; deterrence

*Email address: kielys@nber.org. The findings and conclusions in this manuscript are mine and should not be construed to represent the views of the funding organizations or institutions. All errors are my own. I acknowledge support from the United States Department of Agriculture through a National Bureau of Economic Research administered fellowship (agreement 59-1000-2-0076) received while working on this project.

1 Introduction

Alcohol-impaired driving is a leading cause of preventable death in the United States, responsible for approximately 11,000 fatalities annually, roughly 30 percent of all traffic deaths. Drivers at or above the legal limit are estimated to be substantially more likely to cause a fatal crash than sober drivers (Levitt and Porter, 2001; National Center for Statistics and Analysis, 2026). The central legal instrument for deterring impaired driving is the per se blood alcohol concentration (BAC) limit, above which a driver is presumed legally impaired regardless of observed behavior. All 50 states adopted a 0.08 gram-per-deciliter standard by 2004 following federal highway funding incentives, and the evidence suggests this transition meaningfully reduced alcohol-related fatalities (Dee, 2001), though the magnitude and robustness of the effect remain debated (Eisenberg, 2003). Whether further reducing the limit to 0.05, the standard recommended by the World Health Organization, the National Transportation Safety Board, and the National Highway Traffic Safety Administration, would produce additional safety gains is the central empirical question addressed in this paper. In March 2017, Utah's legislature passed House Bill 155 to lower the per se BAC limit from 0.08 to 0.05 grams per deciliter. The law took effect on December 30, 2018, making Utah the first and only U.S. state to adopt a sub-0.08 per se standard. This paper provides a credible causal evaluation of the law, applying modern synthetic difference-in-differences methods to a 30-year panel of state-level fatality data.

The evidence base on per se BAC limit reductions draws primarily on international experience. Studies of Australia's 0.05 standard, introduced alongside a random breath testing program in the early 1980s, find sustained reductions in fatal crashes across multiple states (Henstridge et al., 1997). A comprehensive international review by Mann et al. (2001), covering eleven countries, finds consistent crash reductions following the introduction or lowering of per se limits, with estimated effects ranging from 5 to 20 percent. Fell and Voas (2006) reach similar conclusions in an updated synthesis focused specifically on the evidence for adopting the 0.05 standard. Fell and Scherer (2017) extend this analysis to the U.S. context, estimating the potential reduction in alcohol-involved fatalities from a 0.08-to-0.05 transition using U.S. crash data alongside international evidence. While these international findings motivate the 0.05 standard, they are difficult to extrapolate directly to the U.S. context given differences in enforcement intensity, driving culture, and baseline alcohol consumption patterns.

Within the United States, the evidence on BAC limits comes primarily from the transition to the 0.08 standard. Dee (2001) estimates that states adopting the 0.08 BAC standard experienced approximately a 7 percent reduction in traffic fatalities, identifying the effect through the staggered rollout of 0.08 laws across states during the 1980s and 1990s. Eisenberg (2003) extends this analysis by controlling for concurrent policy changes (zero tolerance laws, administrative license revocation, and seat belt mandates) and finds the 0.08 estimate attenuates substantially. Kenkel (1993) uses individual-level survey data to estimate the demand for drunk driving, finding that per se laws generate deterrence concentrated among drivers near the legal threshold rather than among heavy drinkers. Evans et al. (1991) provide aggregate-level evidence of deterrence from per se BAC laws, finding significant reductions in alcohol-related fatalities from DWI legislation across U.S. states using pooled cross-section time-series data. Wagenaar et al. (2007) analyze BAC limit changes across 28 states from 1976 to 2002 and estimate that the 0.08 limit prevented approximately 360 deaths annually, further projecting that a reduction to 0.05 could prevent an additional 538 fatalities per year nationwide. Utah's 2019 adoption provides a rare opportunity to estimate the effect of a sub 0.08 per se policy directly within the U.S. institutional setting.

Two studies currently exist that evaluate Utah's 0.05 law. The official NHTSA evaluation by Thomas et al. (2022) found that Utah's per-VMT fatal crash rate fell by 19.8 percent and its per-VMT fatality rate fell by 18.3 percent in 2019 relative to a 2016 baseline, but relied on ARIMA interrupted time series models applied to Utah alone, used comparison states only as descriptive context rather than formal controls, examined only the first post-treatment year, and could not formally test for pre-existing trend differences between Utah and any comparison group. Portillo et al. (2024) apply a difference-in-differences design using counties in neighboring states (Idaho, Nevada, Colorado, and Wyoming) as controls, finding reductions in property damage accidents and a marginally significant nighttime fatality effect, though total fatality estimates are statistically insignificant. Using neighboring states as the control group introduces a potential SUTVA concern: if Utah's law generated behavioral spillovers into bordering states through shared media coverage or cross-border travel, the control counties would be partially treated, biasing the estimated fatality effect toward zero and making a true reduction harder to detect. Portillo et al. (2024) do not examine this spillover channel directly.

I apply the synthetic difference-in-differences (SDID) estimator of Arkhangelsky et al. (2021) to a balanced panel of 50 states from 1994 to 2023 using a national donor pool and evaluating a five-year post-treatment window. SDID constructs a data-driven synthetic control, a weighted average of control states whose pre-treatment trajectory closely matches Utah's, and estimates the treatment effect as a double-differenced weighted comparison, combining the pre-treatment trend matching of synthetic control methods (Abadie and Gardeazabal, 2003; Abadie et al., 2010) with the robustness of difference-in-differences to common time shocks. I use permutation-based placebo inference to construct exact finite-sample p-values without large-sample approximations. I also address COVID-19 confounding by estimating the model under multiple post-treatment sample restrictions that progressively exclude the pandemic years.

I find that Utah's 0.05 BAC law reduces traffic fatalities per 100 million vehicle miles traveled by approximately 12 percent (log point estimate: -0.130, $p = 0.066$). The estimate is virtually unchanged across covariate specifications that control for a broad set of traffic safety policies and demographic characteristics. Robustness checks that restrict to the only pre-pandemic year (2019), exclude 2020, or exclude 2020–2021, all yield estimates in the narrow range of -0.125 to -0.150, each marginally significant at the 10 percent level. Analysis of alcohol-specific sub-outcomes finds a negative but statistically insignificant estimate for BAC-positive fatalities, a null I attribute to low statistical power given Utah's unusually low baseline rate of alcohol involvement rather than to absence of an alcohol channel. The no-alcohol sub-outcome cannot serve as a clean falsification test because NHTSA's multiple imputation procedure assigns estimated BAC values to drivers never directly tested, contaminating the sub-category in both directions, and I present it as a data-quality document rather than an identification exercise. The negative no-alcohol point estimate is directionally consistent with imputation reclassification of deterred 0.05–0.08 drivers, whose fatalities would be assigned BAC $= 0$ by the imputation model if the drivers were never tested.

Heterogeneity analysis finds a marginally significant reduction in weekday fatalities, while the weekend and nighttime estimates are statistically indistinguishable from zero. The nighttime sub-outcome additionally fails its own placebo-in-time test, indicating the nighttime synthetic control does not meet the identification conditions for causal interpretation. I treat the heterogeneity results as exploratory. They suggest which behavioral margin the law may have reached in this particular setting but do not constitute independently identified causal estimates of day-of-week or time-of-day effects. A placebo-in-time test assigns a fake treatment to Utah in 2010 and finds no spurious effect ($p = 0.430$). Translated into welfare terms at the U.S. Department of Transportation's 2023 value of a statistical life ($13.2 million), the point estimate implies an average of approximately 40 lives saved valued at $530 million annually in Utah (U.S. Department of Transportation, 2023). Because the 95 percent confidence interval on the ATT is consistent with zero, no positive lower bound on national welfare benefits can be derived from this estimate. National projections derived from a confidence interval consistent with zero are therefore not reported here.

Relative to prior evaluations, I make three contributions. First, I replace the geographic-proximity control group with a data-driven synthetic control whose composition is determined entirely by pre-treatment outcome trajectories, directly addressing the concern that neighboring states may not constitute a valid counterfactual for a demographically distinctive state like Utah. Second, I extend the evaluation window from one year to five (2019–2023), providing evidence on whether the initial effect persists rather than fades. Third, to my knowledge I provide the first formal welfare quantification of the policy, translating the point estimate into lives saved and economic value. Thomas et al. (2022) evaluate safety impacts comprehensively but do not convert their findings into economic terms. The heterogeneity results additionally speak to which behavioral margin the law reached.

The remainder of the paper proceeds as follows: Section 2 describes the data, Section 3 presents the SDID estimator and identification strategy, Section 4 reports the main results, heterogeneity analysis, robustness checks, and a welfare analysis, and Section 5 discusses limitations and policy implications for other states considering similar legislation.

2 Data

Utah's 0.05 per se BAC law (House Bill 155, signed March 2017) took effect December 30, 2018, becoming the first state to adopt a sub-0.08 per se standard. As only two calendar days of

2018 fell under the new limit, I use 2019 as the first treatment year without loss of precision. To evaluate this policy change, I construct a state-year panel from several administrative sources.

The primary data source is the Fatality Analysis Reporting System (FARS), an annual census of fatal motor vehicle traffic crashes in the United States compiled by the National Highway Traffic Safety Administration (NHTSA). FARS has been operational since 1975 and covers every crash on a public trafficway that results in at least one death within 30 days of the crash. It records crash circumstances, vehicle characteristics, and the reported or tested BAC of involved drivers. For drivers not formally tested at the scene, I use NHTSA's officially imputed BAC values, which NHTSA derives from a multiple imputation procedure using crash and driver characteristics. I use FARS microdata from 1994 to 2023, aggregating to annual state-level counts of total traffic fatalities, fatalities in crashes involving at least one driver with a positive BAC, and fatalities in crashes with no alcohol involvement. The 1994 start year reflects the scope of the initial data collection, and for the window sensitivity robustness analysis I additionally collect FARS fatality counts and Federal Highway Administration (FHWA) state VMT data for 1990–1993, finding that extending the panel to 1990 or 1992 yields virtually identical ATTs (see Appendix Figure 5).

The primary outcome is the natural log of total traffic fatalities per 100 million vehicle miles traveled. Normalizing by vehicle miles traveled (FHWA, 2026) controls for differences in driving exposure across states and over time (using raw counts rather than rates would otherwise conflate changes in per-mile crash risk with changes in total driving) and is a standard normalization in the traffic safety literature. I take the log so the SDID estimate is interpretable as a proportional effect. The dataset is a balanced state-year panel covering 50 states from 1994 to 2023, yielding 1,500 observations. The pre-treatment period spans 1994–2018 (25 years), providing a long baseline for the synthetic difference-in-differences estimator to identify state and time weights.

Table 1 reports pre-treatment means (1994–2018) for Utah, the SDID-weighted synthetic control, and the unweighted donor pool average. The comparison is dominated by one main outlier: Utah's LDS adherent share is more than 35 times the unweighted donor pool average and more than 40 times the SDID-weighted synthetic control. For this reason, I exclude the variable from the preferred specification. This demographic gap shows up directly in the outcome itself. Consistent with the LDS population's low rates of alcohol consumption, Utah's pre-treatment log BAC-positive fatality rate (−1.416) sits well below both the synthetic control (−0.816) and the unweighted pool (−0.712), roughly half as many alcohol-involved fatalities per vehicle mile as the average donor state. Geography sets Utah apart as well, though in a way that is less concerning for identification. Population density is 31 persons per square mile against 193 in the unweighted pool, and even the SDID-weighted synthetic control, chosen to match outcome trajectories rather than covariates, carries a weighted density of 180.8, roughly six times Utah's. Because the outcome is normalized per VMT, this density gap does not directly confound the estimate. Rural highway driving at higher speeds carries higher crash severity per incident, while urban driving generates more total collisions. The net result is that per-VMT fatality rates are broadly comparable across the density range, and alcohol-involved fatalities per VMT show no systematic relationship with population density in the pre-treatment donor pool. More directly, the pre-treatment synthetic control fit ($R^2 \approx 0.87$, RMSPE $= 0.132$ log points) over 25 years confirms that the density-diverse donor pool nevertheless tracks Utah's fatality trajectory closely, absorbing any density-correlated composition differences into the outcome-matching weights.

A third imbalance is harder to dismiss: Utah's real beer tax ($0.605/gallon) runs 65 percent above the synthetic control ($0.366/gallon), and since beer taxes affect alcohol consumption directly, I include the real beer tax as a time-varying covariate. However, the results are robust to including this covariate, confirming that differential beer tax trends are not driving the result. The share of licensed drivers aged 15–24 is higher in Utah as well (17.8 percent versus 14.0 percent in the unweighted pool), another outlier tied to the LDS demographic profile. On the dimensions that matter most for confounding (income, unemployment, educational attainment), Utah tracks the donor pool closely, which helps rule out differential economic trends as an alternative explanation. The 2019–2023 post-treatment window covers five years of policy exposure, though COVID-19 substantially disrupted driving and fatality patterns in 2020–2021, which I address directly in the robustness analysis by varying the included post-treatment years.

State-year covariates fall into three groups. The most direct confounding threat is concurrent legislative changes, so I include seven traffic safety policy variables: mandatory ignition interlock laws for first-time DUI offenders, the 0.08 per se BAC law, open container restrictions, recreational cannabis legalization, the real state beer tax (which reduces alcohol-involved fatalities through price effects Saffer and Grossman, 1987), the real state gasoline tax, and texting-while-driving laws.[1] Demographic and economic variables address two further confounding pathways: 1) cross-state differences in crash risk exposure such as population density, male population share, white population share, and educational attainment and 2) secular trends in economic conditions correlated with both driving and alcohol consumption, including log median household income, unemployment rate, and Democratic presidential vote share. I include weather controls (mean annual temperature and total annual precipitation) because climate affects driving conditions and fatality rates independently of policy. I include two additional variables in a separate specification, the LDS adherent share per 1,000 population and the share of licensed drivers aged 15–24, because they are extreme outliers for Utah relative to any donor state.

3 Empirical Strategy

Given the single-treated-unit, single-adoption-date of the law, I use the synthetic difference-in-differences (SDID) estimator of Arkhangelsky et al. (2021), which extends the synthetic control framework to accommodate additive common time effects. SDID finds two sets of non-negative weights: 1) *unit weights* $\hat{\omega}_i \geq 0$ (with $\sum_{i \neq \mathrm{UT}} \hat{\omega}_i = 1$) that minimize the pre-treatment discrepancy between Utah's outcome trajectory and a weighted average of control states, analogous to the synthetic control method, and 2) *time weights* $\hat{\lambda}_t \geq 0$ ($\sum_{t<2019} \hat{\lambda}_t = 1$) that emphasize pre-treatment years whose outcome patterns most closely resemble those of the post-treatment period, thereby improving temporal comparability. The ATT estimate solves the

[1] Policy data are from the National Conference of State Legislatures (National Conference of State Legislatures, 2024, 2026), the National Institute on Alcohol Abuse and Alcoholism (National Institute on Alcohol Abuse and Alcoholism, 2024), The Urban Institute & Brookings Institution Tax Policy Center (Tax Policy Center, 2026), the U.S. Department of Transportation Federal Highway Administration (FHWA, 2026), and the Insurance Institute for Highway Safety (Insurance Institute for Highway Safety, 2026). All dollar values are in 2023 dollars.

weighted least squares problem

$$\hat{\tau}^{\text{SDID}} = \arg \min_{\tau, \alpha_i \, \beta_t \, \gamma} \sum_{i=1}^{N} \sum_{t=1}^{T} \hat{\omega}_i \, \hat{\lambda}_t \left(Y_{it} - \alpha_i - \beta_t - X_{it}' \gamma - D_{it} \, \tau \right)^2, \tag{1}$$

where α_i and β_t are unit and time fixed effects absorbing permanent state-level differences and common time trends, X_{it} is a vector of time-varying covariates, and $D_{it} = \mathbf{1}[i = \text{Utah}, \, t \geq 2019]$ is the treatment indicator. The preferred specification excludes covariates, but I add additional covariate sets to ensure robustness across these time-varying variables. Arkhangelsky et al. (2021) show that $\hat{\tau}^{\text{SDID}}$ is consistent and asymptotically normal under a low-rank factor model for potential outcomes, and that it dominates both difference-in-differences (uniform weights) and the synthetic control (no time reweighting) when the data-generating process contains both additive unit effects and latent factor structure.

The outcome Y_{it} is the natural log of total traffic fatalities per 100 million vehicle miles traveled in state i and year t. I take the log so that the estimate $\hat{\tau}^{\text{SDID}}$ is interpretable as the proportional change in the fatality rate attributable to the law. The panel covers 50 states from 1994 to 2023, with 1994–2018 as the pre-treatment window and 2019–2023 as the post-treatment period. The 25-year pre-treatment period is beneficial for estimating both the unit weights, which require a long pre-treatment history to identify a synthetic control that tracks Utah closely, and the time weights, which require many pre-treatment periods to identify the most informative ones. In the preferred no-covariates specification, the time weights concentrate on four years: 2018 (37.9 percent), 1994 (25.5 percent), 2017 (25.4 percent), and 2015 (11.2 percent), with the remaining 21 pre-treatment years receiving zero weight. The heavy weight on the years immediately preceding treatment is a characteristic feature of SDID and reflects the fact that the late pre-treatment period provides the most informative comparison for identifying the post-treatment counterfactual.[2]

I use permutation-based placebo inference throughout to construct p-values without large-sample approximations. The procedure assigns the treatment indicator to each of the 49 control states in turn, re-estimates the SDID ATT for each artificial treated unit, and builds an empirical reference distribution of placebo ATTs under the null of no treatment effect. The reported

[2] I implement SDID using the `sdid` command in Stata (Clarke et al., 2024).

p-value is the normal approximation to the permutation distribution, $p = 2\Phi(-|\hat{\tau}/\hat{\sigma}_{\text{placebo}}|)$, where $\hat{\sigma}_{\text{placebo}}$ is the standard deviation of the 49 placebo ATTs.[3] With 49 control states, the minimum achievable exact p-value is $1/50 = 0.02$. Permutation inference is used for all regression estimates reported in the paper.[4]

Figure 1 shows the unit weights for the donor pool states under the no covariates specification. Weight concentrates on Minnesota (17.2 percent), Michigan (11.0 percent), Mississippi (9.0 percent), Iowa (8.1 percent), Florida (7.6 percent), and Nevada (7.4 percent), states whose pre-treatment fatality trajectories track Utah's most closely. This concentration on a small, interpretable subset of the donor pool means the counterfactual is not diluted by states whose fatality paths diverged from Utah's well before 2019. While these weights apply to the no covariates specification, the heterogeneity and robustness checks use different outcomes and therefore generate different unit weights, which I do not separately report.

4 Results

Figure 2 plots Utah's log fatality rate alongside that of "Synthetic Utah" for the full 1994–2023 period. The two series track each other closely throughout the pre-treatment period: the pre-treatment root mean squared prediction error (RMSPE) between Utah and synthetic Utah over 1994–2018 is 0.132 log points, accounting for approximately 87 percent of Utah's pre-treatment time-series variance. Following the law's effective date in December 2018, the series diverge. Utah's fatality rate falls relative to the synthetic control and does not recover. Figure 3 translates this divergence into the year-by-year gap (Utah minus synthetic control) with 95 percent bootstrap confidence bands. In the pre-treatment period (1994–2018), the gaps fluctuate around zero with no systematic trend; Table 5 provides the formal pre-trend test, and I describe it below. The 2018 coefficient is negative, a potential pre-trend concern I address directly in Section 4.5. From 2019 onward, the gap is persistently negative. The

[3]The exact permutation p-value counts the share of the 49 placebo ATTs at least as large in magnitude as the true ATT: $p_{\text{exact}} = \#\{|\hat{\tau}_j^{\text{placebo}}| \geq |\hat{\tau}^{\text{SDID}}|\}/49$. For the main result, 3 of 49 placebos exceed the Utah ATT in magnitude, giving $p_{\text{exact}} = 3/49 = 0.061$ versus 0.066 under the normal approximation. The `vce(placebo)` option in the `sdid` command reports the normal approximation throughout (Clarke et al., 2024); it closely tracks the exact permutation quantile in all specifications.

[4]The event study figure uses cluster-block bootstrap confidence bands as a visual diagnostic of pre-trend stability; all formal inference rests on the permutation-based p-values in the tables.

COVID years 2020–2021 (orange shading) show elevated variance but do not overturn the overall pattern.

Table 2 reports estimates of the average treatment effect over all post-treatment years (2019–2023) across four covariate specifications. The preferred specification (Column 1) estimates equation (1) without covariates. Under this specification, the no-covariates estimate is -0.130 ($p = 0.066$, with 3 of 49 placebo ATTs exceeding this magnitude), implying a reduction in traffic fatalities per 100 million VMT of 12.2%. Column (2) adds the seven traffic-safety policy variables. Column (3) adds the full set of demographic, economic, and weather controls while Column (4) includes the LDS adherent share and the young driver share. The estimates in columns (1)–(3) span fewer than 0.01 log points, confirming that the headline figure is insensitive to the choice of control variables and that pre-treatment trajectory matching delivers the result without covariate adjustment.

I exclude the LDS adherent share and the young driver share from the preferred specifications because their inclusion distorts the synthetic control weights. Utah's LDS adherent share (68.7 percent in the pre-treatment period, versus approximately 25 percent in Idaho, the next highest state) exceeds every donor state by a large margin, creating a situation in which the covariate projection distorts the synthetic control weights rather than improving the counterfactual. This mechanical mismatch inflates the estimated treatment effect substantially, making the all-covariates estimate unreliable for causal interpretation. Table 9 makes this distortion concrete. In the preferred (no-covariates) specification, weight concentrates on states with fatality trajectories that closely track Utah's: Minnesota (17.2 percent), Michigan (11.0 percent), Mississippi (9.0 percent), Iowa (8.1 percent), Florida (7.6 percent), and Nevada (7.4 percent). When the LDS adherent share is added, the optimizer cannot match Utah's 68.7 percent LDS share against any state in the donor pool and compensates by shifting weight to satisfy other covariate constraints: Florida's weight doubles to 16.8 percent, Vermont jumps from 0.7 to 8.3 percent, and six states (Colorado, Tennessee, Nebraska, North Carolina, Ohio, and Alaska) lose all weight.

The resulting synthetic control is constructed from a set of donor states chosen to satisfy a covariate constraint that cannot be satisfied rather than to match Utah's outcome trajectory, explaining why the all-covariates ATT of -0.254 should not be interpreted causally. More generally, Abadie (2021) shows that covariate-adjusted synthetic control specifications are preferred over trajectory-matching alone only when they improve pre-treatment outcome fit. The near-identical estimates in columns (1)–(3) of Table 2 confirm that the policy and demographic covariates add nothing beyond what trajectory-matching achieves, validating the no-covariates specification as the appropriate baseline. All remaining results use the no-covariates specification.

I present the placebo-in-time test in Table 5, restricting the sample to 1994–2018 and reassigning the treatment indicator to Utah beginning in 2010. I choose 2010 because it falls after all U.S. states had adopted the 0.08 per se standard (fully by 2004) and roughly in the middle of the remaining stable-policy window (2004–2018), providing balanced pre- and post-placebo periods. The placebo ATT is −0.052 ($p = 0.430$), statistically indistinguishable from zero. This null result confirms that the SDID estimator does not spuriously attribute a trend break to Utah in a period without a policy intervention, and serves as the formal pre-trend test for the identification strategy. Combined with the pre-treatment RMSPE of 0.132 log points quantifying the synthetic control's tracking accuracy, the placebo test supports the validity of the parallel-trends identification in the reweighted sample.

4.1 Alcohol-Involved Fatalities

I estimate the effect of the law on fatalities in crashes where at least one driver recorded a positive BAC in Table 3, the most direct measure of whether the law reduced alcohol-involved driving outcomes. The full post-period estimate is +0.035 (SE = 0.139), positive in sign but consistent with zero across a very wide confidence interval spanning roughly [−0.24, +0.31]; this estimate cannot distinguish between a large reduction and a large increase in alcohol-involved fatalities. The 2019-only estimate is −0.176 (SE = 0.124), directionally consistent with alcohol deterrence, though also not statistically significant.

I interpret both estimates as reflecting low statistical power rather than absence of an alcohol channel effect. Utah's large LDS population means that the baseline count of alcohol-involved fatalities is substantially below what would be expected for a state of similar size. Standard errors for the BAC-positive outcome are approximately twice as large as those for the total fatality outcome, reflecting the noisier signal in the smaller sub-sample. The positive sign on the full-period estimate is consistent with COVID confounding. National alcohol-involved fatality rates rose sharply in 2020–2021 (National Center for Statistics and Analysis, 2026), and Utah's small denominator sub-group is particularly sensitive to these national shocks, pushing the full-period average rightward relative to the cleaner 2019-only benchmark. To detect the main 12 percent effect in this sub-outcome under permutation inference at the 10 percent level would require approximately doubling the sample size or the effect size, so the noisy positive full-period estimate carries limited evidential weight against the alcohol mechanism.

4.2 BAC Imputation Contamination and the No-Alcohol Outcome

Table 4 reports SDID estimates for fatalities in crashes where no driver recorded any alcohol involvement, for the full post-treatment period (column 1) and for 2019 alone (column 2). This outcome cannot serve as a clean falsification test, and I do not present it as one. The fundamental obstacle is NHTSA's multiple-imputation procedure: approximately half of all FARS fatalities lack directly-tested BAC values, and NHTSA imputes these using a model based on crash characteristics and driver demographics (Subramanian, 2002). The resulting BAC $= 0$ category conflates drivers confirmed sober by direct measurement with drivers who were never tested and assigned zero by the imputation model. If the imputation model systematically misclassifies drivers with BAC in the newly-prohibited 0.05–0.08 range as BAC $= 0$ (plausible because this range is near the threshold that governs officer testing decisions), then a deterrent effect on 0.05–0.08 drinkers would appear mechanically in the "no-alcohol" outcome, contaminating the sub-category distinction in both directions: the no-alcohol null could reflect either absence of an effect or systematic contamination. Restricting to directly-breathalyzed drivers would not resolve the problem because tested drivers are a non-random subsample selected precisely on characteristics correlated with alcohol involvement, introducing selection bias.

I note that this same contamination concern applies to the *positive*-BAC outcome (Table 3): if deterred 0.05–0.08 drivers shift to the imputed-zero category, the measured reduction in BAC-positive fatalities would be attenuated. This is a data-quality limitation inherent to NHTSA FARS data that this analysis cannot resolve. The main identification result, a 12 percent reduction in total traffic fatalities, does not rely on BAC sub-classification and is entirely unaffected by imputation measurement error. Two mechanisms could produce a negative no-alcohol estimate that is consistent with, rather than independent of, the law's alcohol channel. First, NHTSA's imputation model assigns BAC $= 0$ to untested drivers using crash characteristics and driver demographics as predictors; if deterred 0.05–0.08 drinkers are less likely to trigger the officer cues that lead to testing (because they are more alert and less visibly impaired than 0.08+ drivers), their fatalities would be reclassified from the BAC-positive to the BAC-zero category by the imputation algorithm, generating a mechanical negative in the no-alcohol outcome. Second, if the law's public-awareness campaign reduced general risky driving behavior among the broad driving population, not exclusively among drinkers, a modest cross-category reduction is plausible. I cannot quantify either mechanism with available data, and I present the no-alcohol table as data-limitation documentation rather than an identification exercise. The main finding on total fatalities is unaffected by whichever of these channels generates the no-alcohol pattern.

4.3 Heterogeneity by Timing

Table 6 decomposes the effect by day of week and time of day, using the same no-covariates specification and full post-treatment sample. The weekday fatality rate shows a statistically significant reduction of -0.141 ($p = 0.093$), while the weekend estimate is -0.036 and statistically indistinguishable from zero. Both the daytime (6:00 a.m.–6:00 p.m.) and nighttime estimates are noisily estimated, with neither reaching conventional significance thresholds. I also report the pre-treatment RMSPE and a placebo-in-time p-value (fake 2010 treatment on 1994–2018 data) for each sub-outcome. The diagnostics are satisfactory for weekday and daytime, both of which show close pre-treatment tracking and no spurious pre-2019 trend. The weekend synthetic control has weaker pre-treatment fit, though the null placebo nonetheless does not reject. The nighttime specification is problematic with RMSPE $= 0.374$, the highest of all four sub-outcomes, and the placebo-in-time test rejects the null ($p = 0.015$), indicating a spurious pre-trend in the nighttime series that the SDID estimator cannot fully account for. This failure should be read as a warning about the design's reliability for time-of-day sub-outcomes, not merely as a reason to interpret the nighttime point estimate with additional caution.

The weekday concentration is itself unexpected. While the conventional expectation is that an impaired-driving law would reduce fatalities most on weekends and at night, when heavy drinking typically occurs, the opposite pattern appears in Utah. I offer one exploratory interpretation. Utah's large LDS population substantially suppresses night and weekend drinking, shifting the behavioral margin of the 0.05 law onto moderate weekday drinkers whose consumption with a meal or after work brings them into the newly criminalized 0.05–0.08 range. This hypothesis is post-hoc and unidentified in the data. A formal inference test comparing weekday and weekend ATTs would require sub-group placebo distributions that are not directly comparable, and the point estimates alone are insufficient to establish a statistically significant differential. The main heterogeneity results, rather than directly identifying a behavioral mechanism, serve primarily as suggestive evidence about which margin the law may have reached in this particular institutional setting. States with higher baseline rates of nighttime alcohol-involved driving should not expect the weekday pattern to replicate.

4.4 Welfare Analysis

Table 8 translates the SDID point estimate into welfare terms using the U.S. Department of Transportation's 2023 value of a statistical life of $13.2 million (U.S. Department of Transportation, 2023). The counterfactual fatality count for each post-treatment year is the observed count scaled by $e^{0.130} \approx 1.139$, with lives saved equal to the difference. The point estimate implies approximately 40 lives and $530 million in avoided fatality costs per year in Utah. Because the confidence interval on the ATT is consistent with zero, the lower bound on welfare is zero, and these figures are point estimates conditional on the main finding and should not be interpreted as forecasts with a tight range. I note an internal tension in the full-period welfare figures: the robustness section treats 2020 and 2021 as anomalous COVID-affected years, yet the welfare table applies a constant ATT to those years. The pre-COVID 2019 figure (34 lives, $454 million) is therefore the most internally consistent welfare benchmark; the 2022 and 2023 figures (after the acute COVID period) also apply the constant-effect assumption to a more normal traffic environment. Readers should accordingly weight the 2019, 2022, and 2023 point estimates more heavily than the 2020–2021 figures when interpreting Table 8.

A positive lower bound on national welfare benefits cannot be derived from this estimate, and I do not extrapolate the Utah figures to national fatality counts. For context, Wagenaar et al. (2007) projected 538 annual fatalities prevented nationally from 0.05 adoption by extrapolating from 0.08-limit coefficients estimated over 1976–2002. The Utah point estimate implies a larger per-percentage-point effect, likely because the present analysis identifies directly from the 0.05 law itself rather than projecting from a different policy transition. The figures in Table 8 are informative for Utah-specific policy evaluation but should not be extrapolated to other states without accounting for demographic differences in baseline alcohol involvement.

4.5 Robustness Checks

Table 7 examines sensitivity to the inclusion of COVID-affected years as well as anticipation effects of the law in 2018. The first four columns consider four post-treatment samples designed to account for the confounding effects of the pandemic: all years (2019–2023), 2019 only (the only post-treatment year prior to the pandemic), all post-treatment years excluding 2020, and all post-treatment years excluding both 2020 and 2021. Point estimates range from −0.125 (excluding 2020) to −0.150 (excluding 2020–2021), with the 2019-only estimate at −0.147 ($p = 0.074$). All four estimates are statistically significant at the 10 percent level. The 2019-only estimate is particularly informative because it isolates the law's first year of operation, entirely free of COVID-related confounding, and yields an effect consistent in magnitude with the full-period estimate. The stability of results across post-treatment windows that sequentially exclude the most anomalous COVID years rules out the hypothesis that the main finding is driven by pandemic-era changes in driving behavior or fatality reporting.

The law passed the Utah legislature in March 2017 and received substantial public attention during the intervening period, so behavioral responses beginning before the December 30, 2018 effective date are plausible. Column (5) addresses the 2018 pre-trend concern directly by recoding treatment as beginning in 2018, shrinking the pre-treatment window to 1994–2017. The resulting ATT is −0.128 ($p = 0.113$). The point estimate is virtually identical to the main estimate of −0.130, ruling out the 2018 movement as the mechanical source of the result while not resolving whether the 2018 coefficient reflects genuine behavioral anticipation or a pre-existing trend. I note that the direction of potential bias differs across these two interpretations. If the 2018 movement reflects anticipation of the law (behavioral responses before enforcement), the main estimate captures some anticipation effect but remains a valid causal estimate of the cumulative effect. If it reflects a pre-existing downward trend unrelated to the law, the main estimate would be biased upward in magnitude, overstating the true effect. Appendix Table 13 reports placebo-in-time tests at 2016 and 2017, each using the full 1994–2018 pre-treatment sample with the fake treatment date reassigned within that window (the same approach as the 2010 placebo). The 2016 placebo ATT is -0.084 ($p \approx 0.33$) and the 2017 placebo ATT is -0.016 ($p \approx 0.79$), neither statistically significant. The near-zero 2017 estimate is particularly reassuring, since 2017 is the year HB 155 passed the Utah legislature and any announcement-driven anticipation would most plausibly begin there. The 2016 estimate, while negative, falls well within the range of placebo sampling variation and is statistically indistinguishable from zero. It does not constitute evidence of a pre-existing trend. However, $p = 0.113$ for the 2018 recode fails to meet the 10 percent threshold used throughout, and I acknowledge that the main finding's marginal significance is sensitive to this specification choice.

Column (6) extends the anticipation exercise to the law's legislative passage date. Recoding treatment as beginning in 2017 tests whether behavioral responses began at announcement rather than enforcement, with the pre-treatment window contracting to 1994–2016 (23 years). The passage-date ATT is $\hat{\tau} = -0.125$ ($p = 0.141$), statistically indistinguishable from zero. The null at the 2017 passage date, combined with the significant 2019 implementation estimate, is consistent with enforcement-based deterrence rather than announcement effects. Drivers modified behavior when the law carried legal consequences, not when it was signed.

To verify that the main estimate is not driven by any single donor state, I re-estimate the preferred no-covariates specification 49 times, each time dropping one control state from the donor pool. Figure 4 reports the resulting distribution of ATT estimates. Table 11 reports the full set of ATT estimates and p-values for all 49 runs. The estimates range from -0.145 (Minnesota, the highest-weight state) to -0.124 (Illinois, a near-zero-weight state whose removal has minimal effect on the synthetic control) log points, with a mean of -0.130, virtually identical to the full-sample baseline. All 49 leave-one-out estimates share the same sign as the baseline and the range spans only 0.021 log points. Forty-eight of the 49 specifications yield p-values below 0.10, with the sole exception being New Hampshire ($\hat{\tau} = -0.126$, $p = 0.109$), a marginal departure that does not alter the substantive conclusion. The main finding does not hinge on the inclusion of any particular control state.

Appendix Table 12 reports the statistical power of the preferred SDID design as a function of the hypothetical true effect, computed analytically using the placebo-based standard error $\hat{\sigma}_{\text{placebo}} \approx 0.071$ log points. For a one-tailed test at $\alpha = 0.10$, the design achieves 80 percent power for effects of 13.9 percent (0.150 log points) or larger; for a two-tailed test at $\alpha = 0.10$, the corresponding minimum detectable effect rises to approximately 16 percent (0.177 log points). The reported permutation p-values are two-tailed by construction (based on the magnitude of the placebo ATT relative to the true ATT), so the two-tailed threshold is the appropriate comparator: the observed ATT of 12.2 percent falls below the two-tailed 80 percent power threshold, consistent with the marginal $p = 0.066$. For context, effects in the 5–10 percent range found in some international studies of sub-0.08 BAC limits would fall well below the detectable range of this design. Effects at or above 20 percent would be detected with high reliability at $\alpha = 0.10$. The 25-year pre-treatment window is the primary source of precision. Designs with shorter pre-treatment histories face substantially wider confidence intervals and correspondingly higher minimum detectable effects.

The sensitivity of significance to the pre-treatment window length reflects how SDID uses pre-treatment data. Unlike standard difference-in-differences, which treats all pre-treatment years symmetrically, SDID jointly optimizes both synthetic control unit weights and pre-treatment time weights to minimize the discrepancy between the treated unit's trajectory and the reweighted donor pool. This joint optimization identifies the component of the pre-treatment path that is common across treated and control units, isolating it from unit-specific level differences. The precision of this decomposition improves with longer pre-treatment panels, additional years sharpen the time weights and reduce sampling variance in the synthetic control. A short pre-treatment window yields imprecise weights and wider confidence intervals. This identification argument justifies extending the sample back to 1994. To show differences in the pre-treatment window on the treatment effect I vary start date of the window every two years from 1990 to 2014. Appendix Figure 5 shows SDID estimates across all window lengths from 2014 through 1990. The significant result in the preferred specification requires extending the sample back to at least 1998. Windows starting in 1992 ($p = 0.074$) and 1990 ($p = 0.081$) produce virtually identical ATTs (−0.127 and −0.125, respectively), confirming the result is not specific to the 1994 start year.

A further identification concern is the stable unit treatment value assumption (SUTVA). If Utah's law altered behavior in neighboring states, through cross-border driving, shared media coverage, etc., the donor pool may no longer represent a valid untreated counterfactual. Utah borders five states: Idaho, Wyoming, Colorado, Arizona, and Nevada. Of these, only Nevada receives material weight in the preferred synthetic control (7.4 percent), and none of the other four appears among the states gaining or losing weight in Table 9. Spillover through cross-border drinking and driving is the most plausible channel for Nevada specifically, given shared commuting and tourism corridors near the Utah-Nevada border. If Nevada's own fatality trend were partly affected by Utah's law, the contamination would bias $\hat{\tau}^{\text{SDID}}$ toward zero rather than away from it, since a control state that partially adopts Utah's outcome improvement would narrow, not widen, the estimated gap. This would make the reported estimate conservative under this particular threat to identification. Appendix Table 14 addresses this directly by re-estimating the preferred specification after dropping all five bordering states from the donor pool. The ATT is -0.126 ($p = 0.080$), virtually identical to the baseline -0.130, confirming that the estimate is not sensitive to potential spillover contamination through bordering states.

5 Conclusion

Applying synthetic difference-in-differences to a 30-year state-year panel of FARS fatality data, I find that Utah's 2019 reduction of the per se BAC limit from 0.08 to 0.05 reduced traffic fatalities per 100 million vehicle miles traveled by approximately 12 percent. The estimate is stable across covariate specifications and robust to COVID-19 sample exclusions and sample window choices, supporting causal interpretation of the main finding. The estimate is consistent in sign and magnitude with the international literature examining similar policy changes.

Heterogeneity analysis finds a marginally significant reduction in weekday fatalities, while weekend and nighttime estimates are statistically insignificant. The nighttime sub-outcome also fails its own placebo-in-time test, indicating that the nighttime synthetic control does not satisfy the identification conditions required for causal interpretation, and I treat this as a diagnostic warning rather than a credible null result. The heterogeneity evidence is best understood as exploratory. While the conventional expectation is that impaired-driving laws reduce fatalities most on nights and weekends, when heavy drinking typically occurs, this pattern is not borne out in Utah. I postulate that Utah's large Latter-day Saints population substantially suppresses weekend and nighttime alcohol consumption, likely shifting the deterrent effect onto moderate weekday drinkers in the newly criminalized 0.05–0.08 BAC range, but this interpretation is post-hoc and unidentified in the data. States with higher baseline rates of night and weekend alcohol involvement should not assume the weekday pattern will replicate, and the heterogeneity results should not be treated as independently established causal findings.

Three limitations condition the conclusions drawn here. First, the single-treated-unit design constrains statistical precision. As there are only 49 permutation placebos, the design can reliably identify only relatively moderate to large effects, and the main estimate falls near the boundary of what this design can detect. Second, the analysis of alcohol-specific fatality sub-outcomes is underpowered because Utah's baseline count of alcohol-involved crashes is unusually low, limiting direct evidence on the behavioral channel. Third, having only a single post-treatment year before the pandemic means the long-run estimate carries more uncertainty than would be possible with a longer follow-up window.

For state and federal policymakers evaluating a 0.05 per se standard, this study provides the first credible causal estimate from a U.S. institutional context. The welfare implications are economically meaningful: the point estimate implies approximately 40 lives saved and $530 million in avoided fatality costs per year in Utah. Any bias from spillovers to neighboring states or undetected anticipation effects would work toward zero rather than away from it, suggesting the estimate, if anything, understates the true policy effect. The 95 percent confidence interval is consistent with zero, so these welfare figures carry substantial uncertainty and should not be read as precise forecasts.

Interpreting this estimate for other states requires care. Utah is a best-case setting for a law targeting moderate drinkers. The LDS population substantially compresses the baseline rate of alcohol-involved driving, concentrating the at-risk population in the newly criminalized 0.05–0.08 BAC range. For a median U.S. state with higher baseline alcohol involvement and a larger share of fatalities occurring at night and on weekends, the behavioral margin would differ. If the moderate-drinker population is proportionally larger in higher-consumption states, a national 0.05 standard could produce larger effects than the Utah estimate; if that population is less responsive because night and weekend drinking is entrenched, effects could be smaller. The welfare calculation in Wagenaar et al. (2007), which projects 538 annual prevented fatalities nationally from 0.05 adoption, was based on extrapolation from 0.08-limit coefficients and not directly from a 0.05 law; the Utah estimate is the closest available direct evidence but almost certainly not representative of the national average. Policymakers should treat the Utah figure as an informative upper bound for demographically similar low-consumption states and as an uncertain benchmark for higher-consumption states, while recognizing that the direction of the effect, a reduction in fatalities, is consistent with the broader international evidence supporting 0.05 as an effective standard.

Figures and Tables

Table 1: Pre-Treatment Summary Statistics: Utah, Synthetic Utah, and Donor Pool

	Utah (1994–2018)	Synthetic Utah (SDID-weighted)	Donor Pool (Mean)	Donor Pool (SD)
Primary outcomes (log fatalities per 100M VMT)				
Total fatalities	0.204	0.315	0.364	0.360
BAC-positive fatalities	−1.416	−0.816	−0.712	0.345
Traffic safety policy variables (proportion of years active)				
Mandatory interlock (first-time DUI)	0.40	0.12	0.18	0.38
0.08 BAC per se law	0.96	0.69	0.76	0.42
Open container law	1.00	0.83	0.77	0.42
Recreational cannabis legal	0.00	0.03	0.03	0.17
Texting-while-driving ban	0.40	0.39	0.34	0.48
Beer tax ($/gallon, 2023$)	0.605	0.366	0.433	0.402
Gas tax (cents/gallon, 2023$)	37.5	32.3	33.4	9.5
Demographic and economic controls				
Population density (per sq. mi.)	31.0	180.8	193.0	256.4
Log median household income	11.263	11.183	11.159	0.156
Unemployment rate (%)	4.24	5.50	5.40	1.86
Mean annual temperature (°F)	49.4	51.1	52.6	8.7
Total annual precipitation (inches)	13.6	36.5	38.9	14.8
Share male	0.502	0.493	0.493	0.008
Share white	0.940	0.821	0.824	0.121
Share with bachelor's degree or higher	0.284	0.261	0.261	0.055
Democratic presidential vote share	0.284	0.477	0.464	0.094
Outlier covariates (excluded from preferred specification)				
LDS adherent share	0.687	0.016	0.019	0.039
Share of drivers aged 15–24	0.178	0.138	0.140	0.009

Notes: Pre-treatment means over 1994–2018. Synthetic Utah uses SDID unit weights $\hat{\omega}_i$ from the preferred no-covariates specification. Donor pool is the simple unweighted average of all 49 control states. Policy indicator variables are expressed as the proportion of pre-treatment years in which the policy was active. The Donor Pool SD column reports the cross-state standard deviation across the 49 donor states in the pre-treatment period. Utah's LDS adherent share (68.7 percent) and its substantially lower BAC-positive fatality rate (−1.416 vs. −0.712 in the unweighted pool) illustrate the demographic features that motivate excluding these variables from the preferred specification.

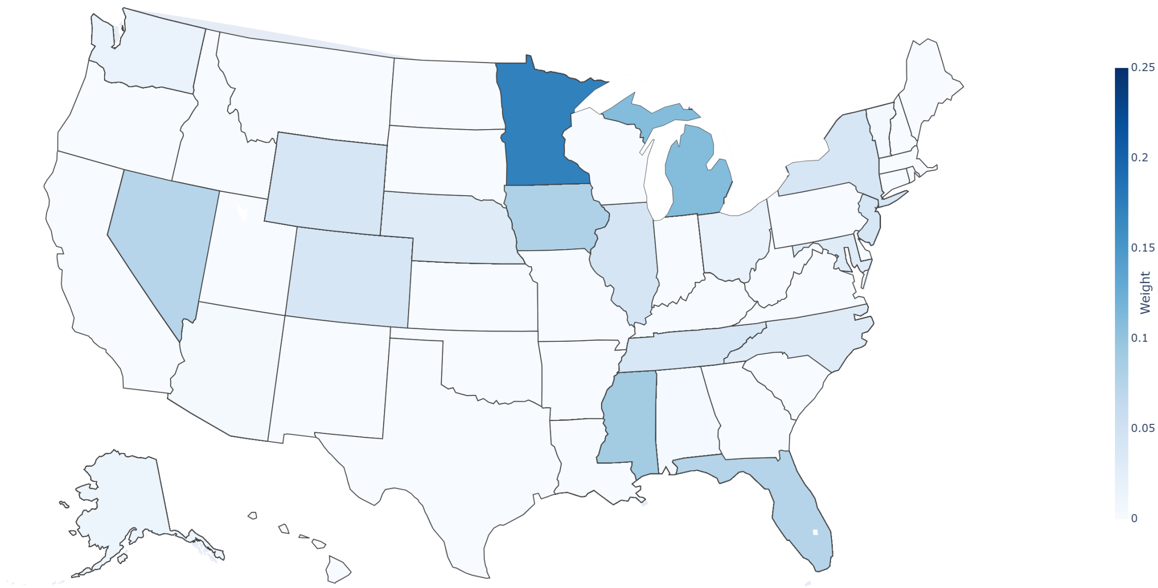


Figure 1: SDID Donor Pool Weights: Preferred Specification

Notes: Shading shows the SDID unit weight $\hat{\omega}_i$ assigned to each control state in the preferred specification (dependent variable log total traffic fatalities per 100 million VMT, no covariates, 1994–2023). Darker blue indicates a higher weight. SDID is sparse and concentrates weight on a small subset of the 49-state donor pool. Utah (the treated state) is shown in white.

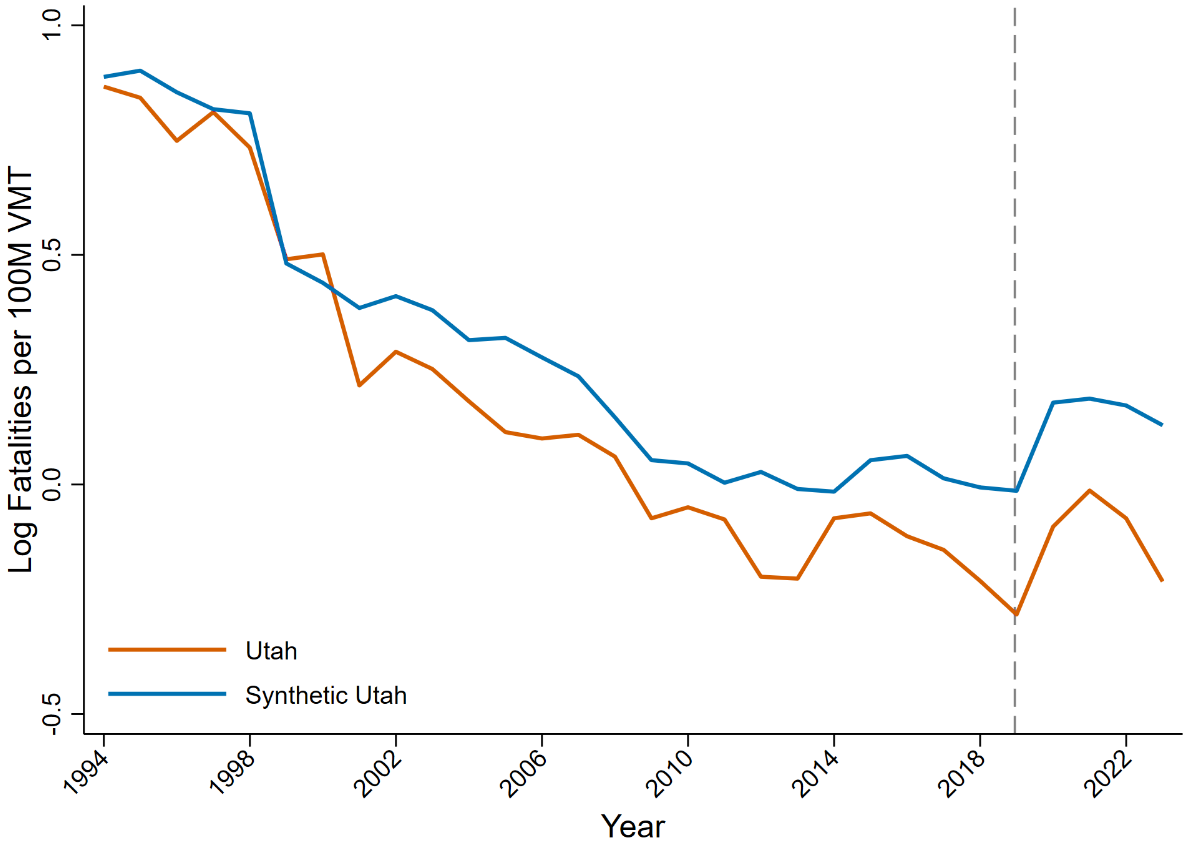


Figure 2: Parallel Trends: Utah vs. Synthetic Utah, Log Traffic Fatalities per 100 Million VMT

Notes: The orange line shows Utah's actual log traffic fatality rate (per 100 million VMT) from 1994 to 2023. The blue line shows the SDID synthetic control: a weighted average of the 49 donor states using unit weights $\hat{\omega}_i$ (no covariates). The vertical dashed gray line marks the effective treatment year, 2019, when Utah's 0.05 BAC law took effect.

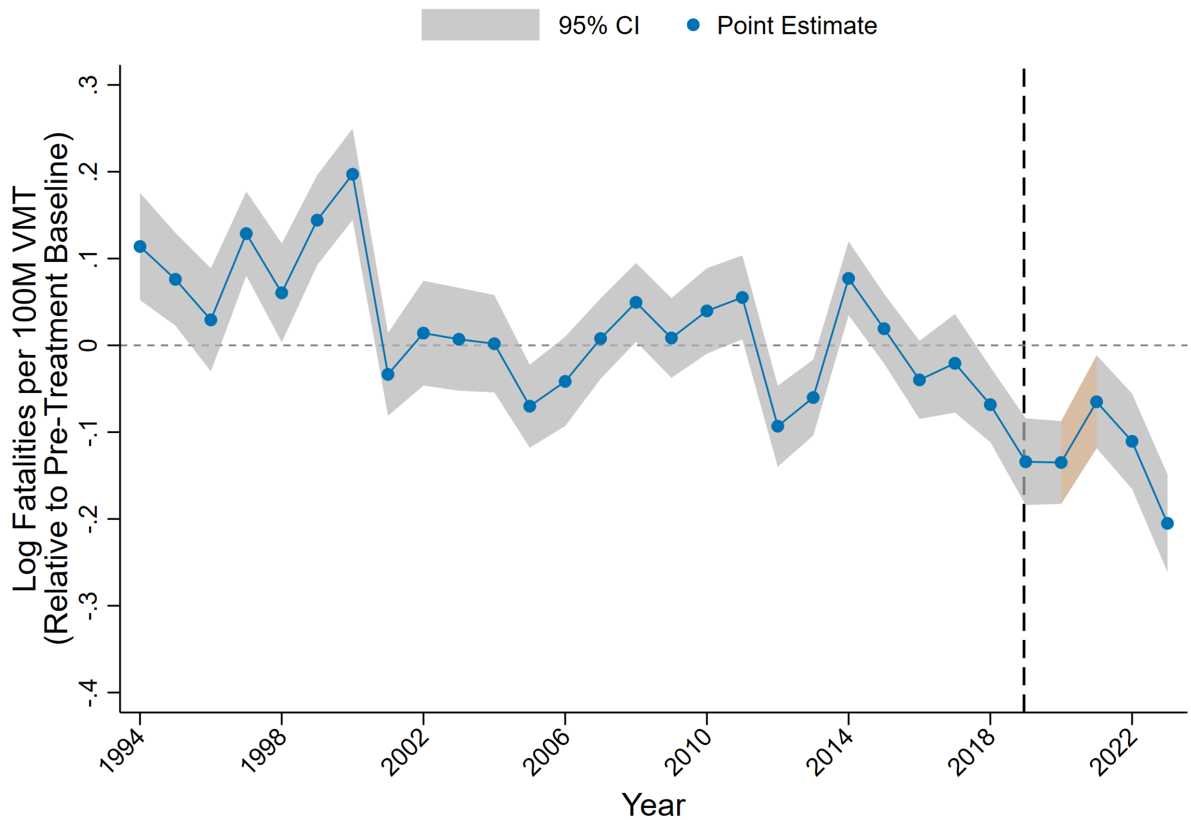


Figure 3: SDID Event Study: Utah Minus Synthetic Utah, Log Traffic Fatalities per 100 Million VMT

Notes: Each point plots the SDID-weighted difference (Utah minus synthetic control) in log total traffic fatalities per 100 million VMT, centered on the pre-treatment weighted mean. Unit and time weights are estimated without covariates. Dashed lines are 95 percent bootstrap confidence bands ($B = 500$ cluster-block bootstrap replicates, sampling states with replacement). The vertical dashed line marks the effective treatment year, 2019. Orange shading covers 2020–2021 (COVID-affected years).

Table 2: Effect of Utah's 0.05 BAC Law on Log Traffic Fatalities per 100 Million VMT

	(1) No Covariates	(2) Policy Covariates	(3) Alt. Covariates	(4) All Covariates
BAC 0.05 Law	-0.130*	-0.120*	-0.123*	-0.254***
	(0.071)	(0.065)	(0.074)	(0.077)
N	1500	1500	1500	1500

Notes: Dependent variable is the log of total fatalities per 100 million vehicle miles traveled. SDID estimator with placebo-based permutation inference (49 control states). Standard errors in parentheses. Policy covariates includes state-level policy variables including mandatory ignition interlock laws, BAC 0.08, open container laws, recreational cannabis, distracted driving and texting, beer tax, and gasoline tax laws. Alt. Covariates add unemployment rate, male population share, white population share, population density, log of mean household income, democratic vote share in Presidential elections, average temperature, and total precipitation covariates. All covariates includes LDS adherent share and young driver share. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 3: Effect on Alcohol-Involved Fatalities per 100 Million VMT

	(1) All Years	(2) 2019 Only
BAC 0.05 Law	0.035 (0.139)	-0.176 (0.124)
N	1500	1300

Notes: Dependent variable is the log of fatalities in crashes with at least one driver BAC >= 0.01 per 100M VMT. No covariates specification. Column (1) estimates the full 2019–2023 post-treatment period; column (2) restricts to 2019 only to avoid COVID confounding. The full-period estimate is +0.035 (SE = 0.139), positive but consistent with zero across a 95 percent confidence interval of approximately [−0.24, +0.31]; it cannot distinguish a large reduction from a large increase in alcohol-involved fatalities. The positive sign is consistent with COVID-era confounding: national alcohol-involved fatality rates rose sharply in 2020–2021, and Utah's small denominator sub-group is particularly sensitive to those national shocks. The 2019-only estimate (−0.176, SE = 0.124) is directionally consistent with alcohol deterrence and avoids this confound. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: BAC Imputation Contamination: Effect on Fatalities with No Alcohol Involvement

	(1) All Years	(2) 2019 Only
BAC 0.05 Law	-0.146	-0.146
	(0.101)	(0.116)
N	1500	1300

Notes: Dependent variable is the log of fatalities in crashes with no driver alcohol involvement per 100M VMT. No covariates specification. Standard errors in parentheses. NHTSA imputes BAC for 50% of FARS fatalities; BAC = 0 includes untested drivers assigned zero by the imputation model. This limits the test to directional evidence only. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Placebo-in-Time Test: Fake Treatment in 2010

	(1)
Placebo Law (2010)	-0.052
	(0.066)
N	1250

Notes: Dependent variable is the log of total fatalities per 100M VMT. No covariate specification. Standard errors in parentheses. Sample restricted to 1994–2018 to exclude actual post-treatment period with placebo treatment reassigned to Utah in 2010.2010 is chosen as the placebo treatment year as it falls after all U.S. states had adopted the 0.08 per se standard (fully by 2004) and roughly in the middle of the remaining stable-policy window (2004–2018), providing balanced pre- and post-placebo periods. The null result ($p \approx 0.43$, 49 control-state placebos) confirms the SDID estimator generates no spurious trend break in Utah absent a policy intervention; this test serves as the formal pre-trend test for the identification strategy. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Heterogeneity: Effect by Day of Week and Time of Day

	(1) Weekday	(2) Weekend	(3) Daytime	(4) Nighttime
BAC 0.05 Law	-0.141*	-0.036	-0.131	-0.084
	(0.084)	(0.097)	(0.084)	(0.127)
Pre-treatment RMSPE (log pts.)	0.118	0.301	0.074	0.374
Placebo p-value (2010 fake)	0.397	0.555	0.621	0.015

Notes: Dependent variable is the log of fatalities per 100M VMT for the indicated sub-period. No covariates specification. Standard errors in parentheses. Daytime defined as 6:00 am – 6:00 pm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: COVID Robustness and Anticipation Checks: Sample Window and Treatment Date Variations

	(1) All Years	(2) 2019 Only	(3) Excl. 2020	(4) Excl. 2020–21	(5) Treat from 2018	(6) Treat from 2017
BAC 0.05 Law	-0.130*	-0.147*	-0.125*	-0.150*	-0.128	-0.125
	(0.071)	(0.083)	(0.074)	(0.080)	(0.081)	(0.085)
N	1500	1300	1450	1400	1500	1500

Notes: Dependent variable is the log of total fatalities per 100M VMT. No covariates specification. Standard errors in parentheses. Columns (1)–(4) vary the post-treatment window to assess sensitivity to COVID-affected driving patterns. Columns (5)–(6) recode the treatment date to test for anticipation: column (5) moves the start to 2018 (the law officially went into effect on December 30th, 2018), column (6) to the March 2017 legislative passage date. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

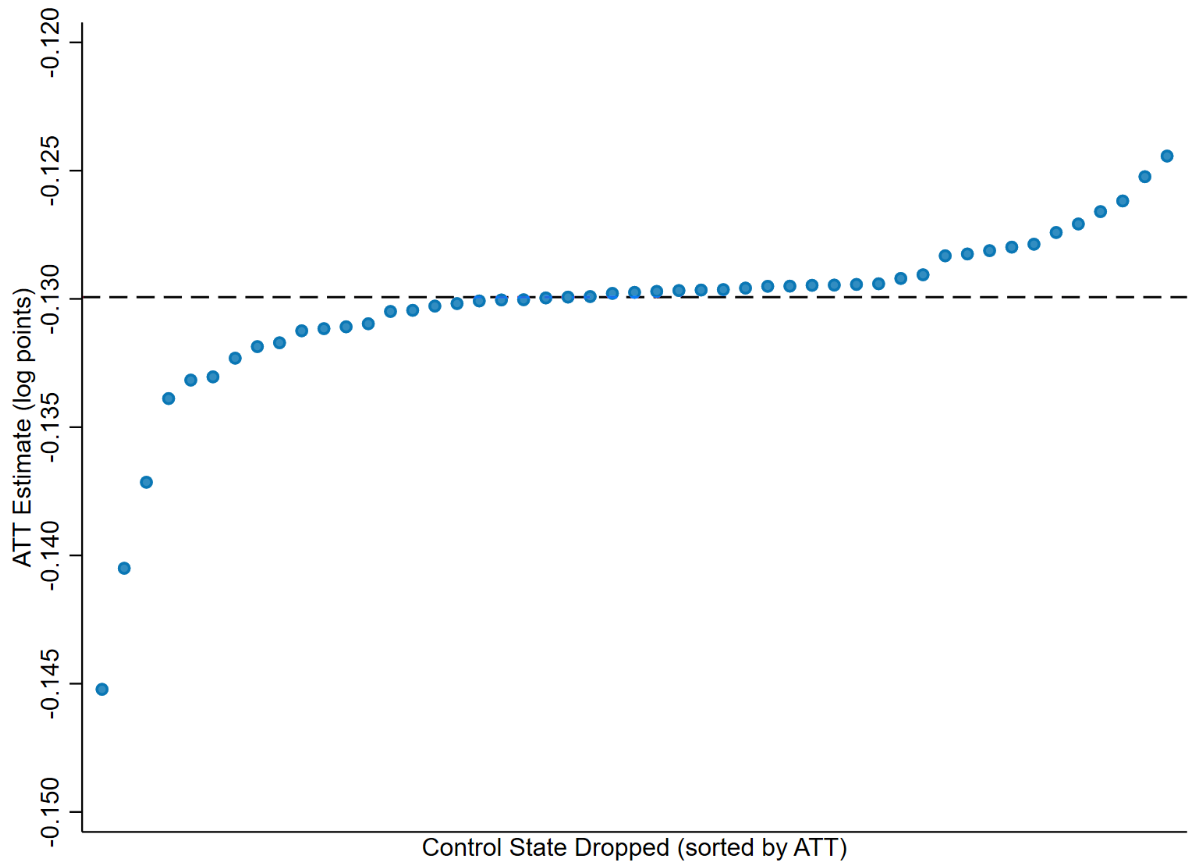


Figure 4: Leave-One-Out Sensitivity: ATT Estimates Dropping Each Control State

Notes: Each point shows the SDID ATT estimate for log total traffic fatalities per 100 million VMT when one of the 49 control states is removed from the donor pool. Estimates are sorted by magnitude. The dashed black line marks the baseline estimate using the full 49-state donor pool (−0.130). No-covariates specification, full 1994–2023 sample.

Table 8: Welfare Calculation: Estimated Lives Saved and Benefits, Utah 2019–2023

Year	Observed Fatalities	Counterfactual Fatalities	Lives Saved	Annual Welfare (million 2023 $)
2019	248	282	34	$454
2020	276	314	38	$506
2021	332	378	46	$608
2022	319	363	44	$585
2023	280	319	39	$514
Annual average	291	331	40	$533
Cumulative total	1,455	1,657	202	$2,666

Notes: Counterfactual fatalities = observed × $e^{|\hat{\tau}|}$, where $\hat{\tau} = -0.130$ is the SDID ATT estimate (no covariates, full 1994–2023 sample). Lives saved = counterfactual − observed. Annual welfare = lives saved × $13.2M, using the U.S. Department of Transportation's 2023 value of a statistical life (U.S. Department of Transportation, 2023). Confidence interval on lives saved: the 95% CI on the ATT under permutation inference spans approximately [−0.257, 0.000] log points (one-sided lower bound; permutation distributions are discrete with 49 mass points), implying a range of 0 to approximately 90 lives saved per year (0 to $1.2 billion annually). Because the CI is consistent with zero, the lower bound on welfare is zero; the figures in the table are point estimates conditional on the main finding and should not be interpreted as forecasts with a tight range. All dollar figures in 2023 dollars.

Appendix

Table 9: SDID Unit Weights: Preferred Specification vs. All-Covariates Specification

State	Preferred ($\hat{\omega}_i$)	All Covariates ($\hat{\omega}_i$)
Minnesota	0.172	0.189
Michigan	0.110	0.148
Mississippi	0.090	0.037
Iowa	0.081	0.075
Florida	0.076	0.168
Nevada	0.074	0.082
Illinois	0.043	0.012
New Jersey	0.042	0.044
Wyoming	0.042	0.018
Colorado	0.040	0.000
New York	0.040	0.069
Tennessee	0.039	0.000
Maryland	0.030	0.038
Nebraska	0.030	0.000
North Carolina	0.030	0.000
Ohio	0.017	0.000
Washington	0.016	0.035
Alaska	0.012	0.000
Vermont	0.007	0.083
All other states	≈ 0	≈ 0

Notes: SDID unit weights $\hat{\omega}_i$ for states with weight above 0.01 in at least one specification; rows sum to approximately 1 within each column. Preferred specification uses the log of total fatalities per 100 million VMT as the dependent variable with no covariates between 1994–2023. As no state comes close to Utah's 68.7 percent LDS share included in column (2), the covariate projection cannot construct a valid match and instead shifts weight to satisfy other covariate constraints, inflating the implied ATT to −0.254 ($p = 0.001$).

Table 10: SDID Time Weights: Preferred (1994–2023) vs. Extended (1990–2023) Window

Year	Main (1994–2023) $\hat{\lambda}_t$ (%)	Extended (1990–2023) $\hat{\lambda}_t$ (%)
1990	0.0	7.0
1994	25.5	20.9
2015	11.2	11.1
2017	25.4	26.2
2018	37.9	34.8
All other years	0.0	0.0
Total	100.0	100.0

Notes: SDID time weights $\hat{\lambda}_t$ for pre-treatment years receiving weight ≥ 0 percent in either specification. Weights sum to 100 percent over all pre-treatment years within each specification. The main specification uses the 1994–2023 panel (25 pre-treatment years) while the extended specification adds 1990–1993.

Table 11: Leave-One-Out Sensitivity: ATT Estimates and p-Values for All 49 Donor States

State Dropped	ATT Estimate	p-Value
Alabama	-0.131*	0.083
Alaska	-0.129*	0.060
Arizona	-0.130*	0.062
Arkansas	-0.130*	0.062
California	-0.131*	0.076
Colorado	-0.130*	0.064
Connecticut	-0.132*	0.056
Delaware	-0.128*	0.059
Florida	-0.132*	0.063
Georgia	-0.130*	0.069
Hawaii	-0.127*	0.064
Idaho	-0.130*	0.057
Illinois	-0.124*	0.095
Indiana	-0.130*	0.074
Iowa	-0.137*	0.068
Kansas	-0.130*	0.085
Kentucky	-0.130*	0.085
Louisiana	-0.130*	0.086
Maine	-0.130*	0.085
Maryland	-0.130*	0.086
Massachusetts	-0.129*	0.087
Michigan	-0.128*	0.092
Minnesota	-0.145*	0.056
Mississippi	-0.133*	0.082
Missouri	-0.130*	0.084
Montana	-0.130*	0.085
Nebraska	-0.133*	0.081
Nevada	-0.129*	0.096
New Hampshire	-0.126	0.109
New Jersey	-0.130*	0.093
New Mexico	-0.134*	0.061
New York	-0.128*	0.079
North Carolina	-0.130*	0.077
North Dakota	-0.128*	0.069
Ohio	-0.128*	0.072
Oklahoma	-0.132*	0.063
Oregon	-0.140**	0.045
Pennsylvania	-0.129*	0.069
Rhode Island	-0.130*	0.070
South Carolina	-0.130*	0.072
South Dakota	-0.129*	0.071
Tennessee	-0.125*	0.076
Texas	-0.130*	0.068
Vermont	-0.129*	0.070
Virginia	-0.131*	0.061
Washington	-0.127*	0.060
West Virginia	-0.130*	0.090
Wisconsin	-0.131*	0.088
Wyoming	-0.127*	0.098

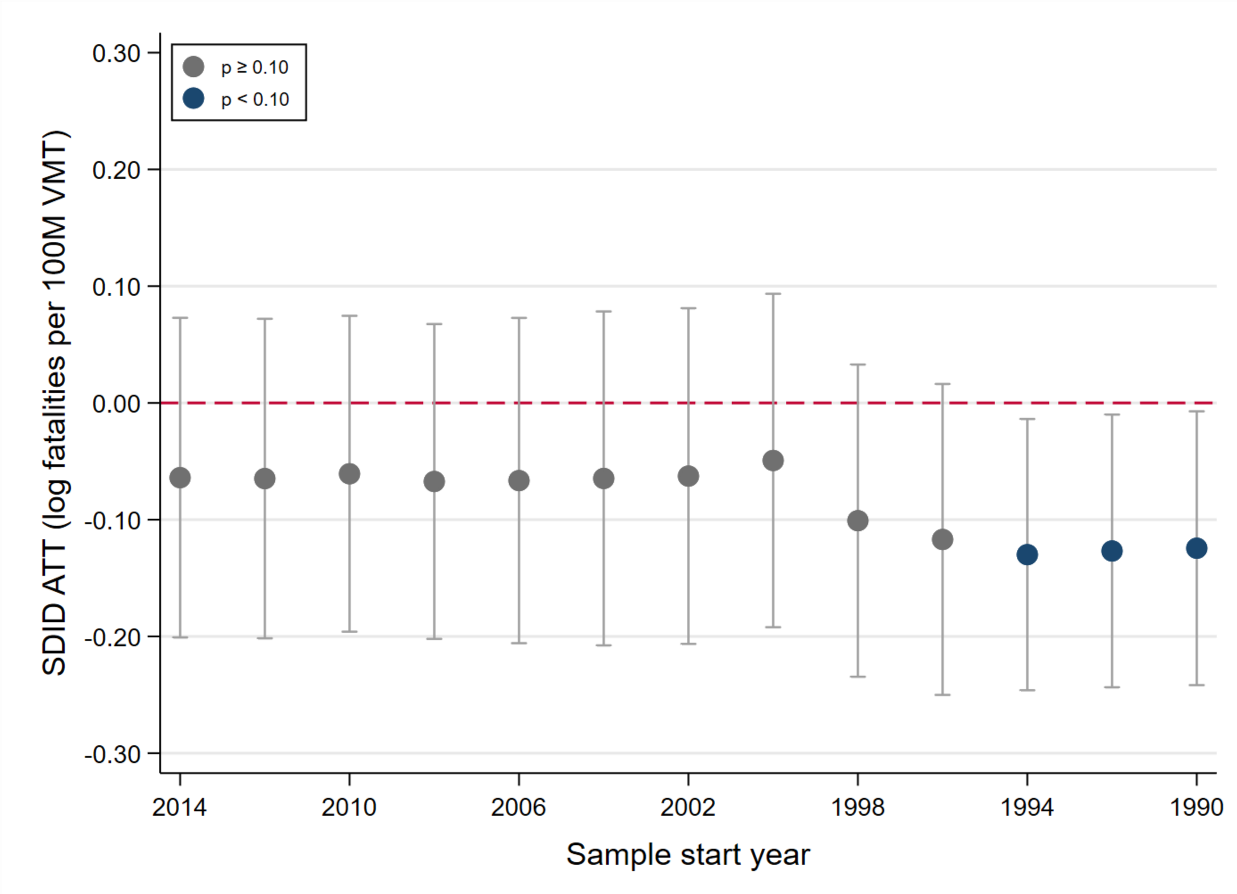
Notes: Each row reports the SDID ATT estimate and p-value when the listed state is removed from the 49-state donor pool. The remaining 48 states serve as donor pool, and each LOO run uses 48 permutation placebos. For computational tractability across 49 LOO iterations, p-values use the normal approximation $z = \hat{\tau}/\hat{\sigma}_{\text{placebo}}$. With 48 placebos per LOO run the normal approximation closely tracks the exact permutation p-value. States sorted alphabetically. Baseline estimate using the full donor pool: -0.130 ($p = 0.066$). * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 12: Statistical Power of the SDID Design: Minimum Detectable Effect Analysis

True Effect (%)	Log-Point Effect	Power α=0.10	Power α=0.05
0	0.000	10.0%	5.0%
5	0.051	28.9%	17.9%
8	0.083	46.0%	32.1%
10	0.105	58.3%	43.9%
12.2*	0.130	71.3%	57.8%
15	0.163	84.6%	74.4%
17	0.186	91.3%	84.0%
20	0.223	97.0%	93.5%
25	0.288	99.7%	99.2%
80% power threshold: α=0.10: 13.9% (0.150 log pts.); α=0.05: 16.1% (0.176 log pts.)			

Notes: Power is computed analytically using the placebo-based standard error $\hat{\sigma}_{\text{placebo}}$ from the preferred specification (no covariates, 1994–2023 sample). One-tailed test at the stated significance level (α): the law is expected to reduce fatalities, so the critical region is $\hat{\tau}/\hat{\sigma} < -z_\alpha$. Power$(\delta,\alpha) = \Phi(-z_\alpha + |\delta|/\hat{\sigma}_{\text{placebo}})$ where $z_\alpha = \Phi^{-1}(1-\alpha)$. True effects are expressed as the percent reduction in traffic fatalities per 100 million VMT and the equivalent log-point effect. The asterisk (*) marks the effect size corresponding to the observed ATT of −0.130 log points. MDE: minimum detectable effect at 80% power.

Figure 5: SDID Estimates by Sample Start Year



Notes: Each point is a separate SDID estimation using all years from the indicated start year through 2023; the post-treatment period (2019–2023) is held fixed across all windows. Bars show 90 percent confidence intervals based on placebo permutation inference (50 replications). A filled navy circle indicates $p < 0.10$ (two-tailed); a grey circle indicates $p \geq 0.10$. Dependent variable is the log of total traffic fatalities per 100 million VMT. No-covariates specification.The horizontal dashed line marks zero. Moving rightward along the x-axis extends the pre-treatment window further back in time. FARS fatality and FHWA VMT data for 1990–1993 were collected separately from the main analysis dataset.

Table 13: Placebo-in-Time Tests: 2016 and 2017 Fake Treatments

	(1) Placebo: 2016	(2) Placebo: 2017
Placebo Law	-0.084 (0.087)	-0.016 (0.061)
N	1250	1250

Notes: Dependent variable is the log of total fatalities per 100M VMT. No covariates specification. Standard errors in parentheses. Both columns restrict to the full pre-treatment sample (1994–2018) and reassign Utah's treatment date within that window. Column (1) assigns a fake treatment beginning in 2016 with a pre-period from 1994–2015 and the post-period from 2016–2018. Column (2) assigns fake treatment beginning in 2017 with a pre-period between 1994–2016 and the post-period is 2017–2018. A non-significant ATT rules out a spurious pre-existing downward trend in Utah's fatality rate at that date. These tests address the concern that the negative 2018 event-study coefficient reflects a pre-trend rather than anticipation of the law. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 14: SUTVA Robustness: Excluding Utah's Bordering States from the Donor Pool

	(1) Baseline	(2) Excl. bordering states
BAC 0.05 Law	-0.130*	-0.126*
	(0.071)	(0.072)
Donor states	49	44

Notes: Dependent variable is the log of total traffic fatalities per 100M VMT, full 1994–2023 sample.. No covariates specification. Standard errors in parentheses. Column (1) reproduces the preferred specification using all 49 donor states. Column (2) drops the five states bordering Utah (Arizona, Colorado, Idaho, Nevada, and Wyoming) from the donor pool, leaving 44 control states. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

ABADIE, A. (2021): "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects," *Journal of Economic Literature*, 59, 391–425.

ABADIE, A., A. DIAMOND, AND J. HAINMUELLER (2010): "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 105, 493–505.

ABADIE, A. AND J. GARDEAZABAL (2003): "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, 93, 113–132.

ARKHANGELSKY, D., S. ATHEY, D. A. HIRSHBERG, G. W. IMBENS, AND S. WAGER (2021): "Synthetic Difference-in-Differences," *American Economic Review*, 111, 4088–4118.

CLARKE, D., D. PAILAÑIR, S. ATHEY, AND G. W. IMBENS (2024): "On Synthetic Difference-in-Differences and Related Estimation Methods in Stata," *The Stata Journal*, 24, 557–598.

DEE, T. S. (2001): "Does Setting Limits Save Lives? The Case of 0.08 BAC Laws," *Journal of Policy Analysis and Management*, 20, 111–128.

EISENBERG, D. (2003): "Evaluating the Effectiveness of Policies Related to Drunk Driving," *Journal of Policy Analysis and Management*, 22, 249–274.

EVANS, W. N., D. NEVILLE, AND J. D. GRAHAM (1991): "General Deterrence of Drunk Driving: Evaluation of Recent American Policies," *Risk Analysis*, 11, 279–289.

FELL, J. C. AND M. SCHERER (2017): "Estimation of the Potential Effectiveness of Lowering the Blood Alcohol Concentration (BAC) Limit for Driving From 0.08 to 0.05 Grams per Deciliter in the United States," *Alcoholism: Clinical and Experimental Research*, 41, 2128–2139.

FELL, J. C. AND R. B. VOAS (2006): "The Effectiveness of Reducing Illegal Blood Alcohol Concentration (BAC) Limits for Driving: Evidence for Lowering the Limit to .05 BAC," *Journal of Safety Research*, 37, 233–264.

FHWA (2026): "Highway Statistics Series," https://www.fhwa.dot.gov/policyinformation/statistics.cfm, accessed 2026-07-21.

Henstridge, J., R. Homel, and P. Mackay (1997): "The Long-Term Effects of Random Breath Testing in Four Australian States: A Time-Series Analysis," Tech. Rep. DOT HS 808 841, National Highway Traffic Safety Administration, Washington, DC.

Insurance Institute for Highway Safety (2026): "Electronic Device Laws," https://www.iihs.org/research-areas/distracted-driving/electronic-device-laws, accessed 2026-07-21.

Kenkel, D. S. (1993): "Drinking, Driving, and Deterrence: The Effectiveness and Social Costs of Alternative Policies," *Journal of Law and Economics*, 36, 877–913.

Levitt, S. D. and J. Porter (2001): "How Dangerous Are Drinking Drivers?" *Journal of Political Economy*, 109, 1198–1237.

Mann, R. E., S. Macdonald, G. Stoduto, S. Bondy, B. Jonah, and A. Shaikh (2001): "The effects of introducing or lowering legal per se blood alcohol limits for driving: An international review," *Accident Analysis & Prevention*, 33, 569–583.

National Center for Statistics and Analysis (2026): "Alcohol-Impaired Driving: 2024 Data," Traffic Safety Facts DOT HS 813 815, National Highway Traffic Safety Administration, Washington, DC.

National Conference of State Legislatures (2024): "State Ignition Interlock Laws," National Conference of State Legislatures. https://www.ncsl.org/transportation/state-ignition-interlock-laws.

——— (2026): "State Medical Cannabis Laws," National Conference of State Legislatures. https://www.ncsl.org/health/state-medical-cannabis-laws.

National Institute on Alcohol Abuse and Alcoholism (2024): "Open Containers of Alcohol in Motor Vehicles," National Institutes of Health. https://alcoholpolicy.niaaa.nih.gov/.

PORTILLO, J. E., W. SUGIARTO, AND K. WILLARDSEN (2024): "Drink...then drive away: The effects of lowering the blood alcohol concentration in Utah," *Health Economics*, 33, 1869–1894.

SAFFER, H. AND M. GROSSMAN (1987): "Beer Taxes, the Legal Drinking Age, and Youth Motor Vehicle Fatalities," *Journal of Legal Studies*, 16, 351–374.

SUBRAMANIAN, R. (2002): "Transitioning to Multiple Imputation — A New Method to Impute Missing Blood Alcohol Concentration (BAC) Values in FARS," Tech. Rep. DOT HS 809 403, National Highway Traffic Safety Administration, National Center for Statistics and Analysis, Washington, DC, revised October 2002.

TAX POLICY CENTER (2026): "State Alcohol Excise Tax Rates," Urban Institute & Brookings Institution. https://taxpolicycenter.org/statistics/state-alcohol-excise-tax-rates.

THOMAS, F. D., R. BLOMBERG, J. R. DARRAH, L. A. GRAHAM, T. SOUTHCOTT, R. DENNERT, E. TAYLOR, R. TREFFERS, S. TIPPETTS, S. MCKNIGHT, AND A. BERNING (2022): "Evaluation of Utah's .05 BAC Per Se Law," Tech. Rep. DOT HS 813 233, National Highway Traffic Safety Administration, Washington, DC.

U.S. DEPARTMENT OF TRANSPORTATION (2023): "Revised Departmental Guidance on Valuation of a Statistical Life in Economic Analysis," Tech. rep., Office of the Secretary, U.S. Department of Transportation, https://www.transportation.gov/office-policy/transportation-policy/revised-departmental-guidance-on-valuation-of-a-statistical-life-in-economic-ana

WAGENAAR, A. C., M. M. MALDONADO-MOLINA, L. MA, A. L. TOBLER, AND K. A. KOMRO (2007): "Effects of Legal BAC Limits on Fatal Crash Involvement: Analyses of 28 States from 1976 through 2002," *Journal of Safety Research*, 38, 493–499.