



A Natural Experiment to Test the Effect of Sanction Certainty and Celerity on Substance-Impaired Driving: North Dakota's 24/7 Sobriety Program

Greg Midgette^{1,2} · Beau Kilmer³ · Nancy Nicosia⁴ · Paul Heaton⁵

Published online: 4 May 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Objectives Evaluate the deterrent effect of a program that increases the certainty and celerity of sanction for arrestees ordered to abstain from alcohol and other drugs on substance-impaired driving arrests.

Methods We examine participant compliance with orders to abstain from alcohol and other drug use via breathalyzer, body-worn continuous alcohol monitoring devices, transdermal drug patches, and urinalyses. We then evaluate the impact of the 24/7 Sobriety program on substance-impaired driving arrests. Using variation across counties in the timing of program implementation in North Dakota as a natural experiment, we use difference-in-differences fixed effects Poisson regressions to measure the program's effect on county-level arrests for substance-impaired driving.

Results Over half of participants ordered to abstain from substance use complete 24/7 Sobriety without a detected substance use event. At the county level, the program is associated with a 9% reduction in substance-impaired driving arrests after accounting for the impact of oil exploration in the Bakken region, law enforcement intensity, alcohol availability, whether the state's large universities were in session, and socio-demographic characteristics. We find that the Bakken oil boom is associated with a 22% increase in substance-impaired driving arrests.

Conclusions The results suggest frequent monitoring combined with increased sanction celerity deters substance use-involved crime. While the results are generally consistent with an earlier study of 24/7 Sobriety in another state, differences in the study outcome measures implementation choices across states make direct comparisons difficult. More can be learned by conducting randomized controlled trials that vary time on program, testing technology, and/or level of sanction.

Keywords Deterrence · Decision-making · Alcohol · Community corrections · Substance use · Impaired driving

✉ Greg Midgette
gem@umd.edu

Extended author information available on the last page of the article

Introduction

Approximately 20 million Americans drive under the influence of alcohol in a given year, another six million report driving under the influence of alcohol and illegal drugs at the same time, and three million report driving under the influence of only illegal drugs (Lipari et al. 2016). The annual societal costs of alcohol-involved traffic crashes in the U.S. is on the order of \$125 billion (Zaloshnja et al. 2013). The National Highway Traffic Safety Administration reported that alcohol-involved crashes killed 10,874 people in 2017 (National Center for Statistics and Analysis 2019); less is known about fatalities from driving under the influence of prescription or illegal drugs, but the incidence of such deaths is thought to be substantial (Hedlund 2017).

There is a rich research literature examining efforts to target drunk driving (e.g., Ross 1973; Kenkel 1993; Ruhm 1996; Carpenter and Dobkin 2010) and the desire to reduce driving under the influence of other substances has become increasingly salient as more states consider alternatives to cannabis prohibition (Caulkins et al. 2016). Recent data suggest that roughly 25% of those convicted of driving while intoxicated nationally have a future intoxicated driving offense, with the recidivism rate rising to 70% in some states (Warren-Kigenyi and Coleman 2014).

For substance-involved individuals subject to community supervision, judges and community corrections officials may order abstinence from alcohol and other drug use and monitor their behavior with testing. This evaluation provides empirical evidence on North Dakota's 24/7 Sobriety program, which prioritizes certainty and celerity over severity to deter substance use and related crime. 24/7 Sobriety orders abstinence and combines frequent monitoring with a rapidly applied sanction for non-compliance; typically a program violation results in a night in jail that is not included in participants' criminal records. This study also contributes to the broader literature on community corrections approaches to reducing substance-use related harms. Importantly, North Dakota represents the first effort to adapt the innovative and successful South Dakota 24/7 Sobriety program to a new setting and, therefore, this paper also provides additional evidence regarding the generalizability of the *swift and certain* sanctioning approach.

The 24/7 Sobriety approach is notable for its differences from traditional community supervision in several key respects which comport with theoretical and empirical findings from criminology and behavioral economics. Relative to supervision as usual, it substantially increases the probability that violations will be detected, increases the probability of punishment conditional on detection to near certainty, and reduces the delay in administering punishment, but lowers the severity of punishment. Theory suggests that this model may have important advantages relative to the traditional approach, but there is limited and sometimes conflicting empirical evidence regarding the effectiveness of this model.

Exploiting variation across counties in the timing of program adoption within North Dakota, we estimate difference-in-differences models to identify the effect of program availability on county-level driving under the influence of alcohol or other drugs (DUI) arrests. We find a statistically and substantively significant reduction in DUI arrests associated with the introduction of the new supervision and sanctioning regime. Based on our preferred specification, DUI arrests decline by approximately 9% after 24/7 Sobriety becomes operational, a finding that is consistent across a number of samples and specifications. We assess the robustness of our findings with several sensitivity

analyses and validate the results with a permutation test. We also present event study analyses demonstrating that those counties that implemented 24/7 Sobriety early had similar pre-existing DUI arrest patterns as those that implemented later.

Background

Deterring Alcohol and Other Drug Use in Community Supervision

The certainty, celerity, and severity of sanctions to deter illicit behaviors were features of the seminal works of Beccaria (1764) and Bentham (1789). However, typical practice in community supervision settings tends to be inconsistent with deterrence theory and subsequent empirical evidence. For criminal justice-involved persons with histories of substance use, traditional community supervision practices introduce both delay and uncertainty into the negative incentives of jail time or fines for non-compliance. Violations are unlikely to be detected because of infrequent testing, thereby introducing uncertainty. An officer must then refer the person under supervision back to the court for a hearing, which introduces delay. Then at the hearing, the judge must decide whether to revoke conditional release, adding to the uncertainty.

To compensate for the uncertainty, sanctions are often severe when they are carried out. Revocation of release in the pre-trial context means detaining a (not yet convicted) person until the trial date, which may be weeks or months away. For those who were convicted and on probation, revocation of release means imposing the required custodial sentence, which often involves a substantial amount of jail time: for example, 12 states required at least 30 days in jail for a repeat DUI offense and an additional 32 states required at least a week in jail (Comoreanu 2017).

In the United States, the criminal justice protocol for repeat-DUI typically allows arrestees to remain in the community, but requires them to conform to a set of conditions established by judges. Monitoring is performed by community supervision officers. Requirements of community supervision can be both affirmative and restrictive. A participant may face a portfolio of conditions that, for example, combine orders to meet regularly with a probation officer, attend 12-step meetings, abstain and participate in alcohol or other drug testing, avoid environments such as bars, and limit use of vehicles. In theory, the goal of the supervision conditions is to enable the individual to engage in behavioral change to avoid future infractions. Because many conditions require behaviors that individuals would not engage in of their own volition, traditional community supervision relies largely on the threat of future incarceration to induce behavioral change. The incentive-based approach used in the criminal justice system contrasts with the disease model of addiction favored by many in the public health community, which conceptualizes addiction as a chronic medical condition that requires clinical intervention to correct.

Under an incentive-based paradigm exemplified by the classic (Becker 1968) model of rational offending, theory would predict that the perceived threat of incarceration, measured by certainty, celerity, and severity, would induce greater compliance with the conditions of community supervision, despite the disutility that persons under supervision might experience from abstinence and other requirements. Among the three levers, perceived sanction certainty appears to be most salient (Nagin 2013; Chalfin and McCrary 2017).

Intensive supervision programs (ISPs) were developed as an alternative to the routine model designed to address some of its limitations through more frequent contact with a

supervising officer, random drug testing, work or community service requirements that vary across programs, and sometimes substance use treatment. ISPs evaluated in experimental settings tended to be onerous and punitive compared to traditional supervision and resulted in more technical violations, but were found to have no effect on preventing crime (Petersilia and Turner 1990, 1993). A cost effectiveness analysis of prison-diversion ISPs in Minnesota found they are as costly per person as prison (Deschenes et al. 1995). This suggests close monitoring, frequent penalties, and coerced enrollment in substance use treatment and pro-social activities are an inefficient use of scarce supervision resources.

In an analysis of the British Road Safety Act of 1967 and prior deterrence literature, Ross (1973: 67) notes, “The literature suggests that it is necessary and sufficient for deterrence that credibility—the subjective perception of certainty—increase. This perception is very likely to be influenced by the objective increase in certainty of punishment.” The 1967 Act increased perceived certainty of penalties and in the short-run appeared to effectively deter drinking and driving without increasing penalties. However, the law’s impact waned as the legal system’s ambivalence and moderation toward its provisions became broadly apparent.

Subsequent empirical research continues to reinforce the idea that sanction certainty is central to policies’ deterrent potential, and that the deterrent value of more punitive sanctions is mixed (Nagin 2013; Chalfin and McCrary 2017).¹ Randomized controlled trials in community supervision demonstrate that individuals who face the threat of a short spell of incarceration are more likely pay outstanding fines (Weisburd et al. 2008) and to comply with rules while on probation (Hawken and Kleiman 2009). However, individuals with prior exposure to criminal sanction tend to perceive less risk of sanction (Paternoster et al. 1985), thus making the task of preventing recidivism among repeat-offenders through criminal justice deterrence alone more difficult.

The value of celerity is less well understood, but findings from criminology and behavioral economics with respect to substance use and related criminal justice outcomes suggest hyperbolic time discounting among justice-involved individuals (Loughran et al. 2012; Loughran 2019). People who use drugs also exhibit hyperbolic discount rates and time inconsistencies in decision making (Bickel et al. 2011). Hyperbolic valuation of time appears to be independent of detection certainty (Loughran et al. 2012), and is especially prominent among intoxicated drivers (Sloan et al. 2014) and dependent drug users (Bickel and Marsch 2001). Thus, despite a paucity of specific evidence, the stimulus provided by programs that deliver a high celerity stimulus appears less likely to be eroded by individuals’ cognitive processes.

24/7 Sobriety

In 2008, North Dakota implemented a program called 24/7 Sobriety designed to reduce alcohol and other drug consumption and associated adverse consequences among individuals under community supervision for substance-involved offenses. In addition to making abstinence from alcohol or illegal drugs a condition of bond or probation, participants were ordered to either (1) show up at the county jail *every* morning and *every* evening and blow into a breathalyzer, (2) wear an alcohol monitoring bracelet that tests their sweat every

¹ However, in the specific context of driving under the influence of alcohol, Hansen (2015) finds that graduated sanctions at increasing blood alcohol content thresholds (BAC) deter drunk driving recidivism.

30 minutes for alcohol, and/or (3) submit to urinalysis or drug patch testing for illegal drugs. Those testing positive for alcohol or illegal drugs, missing a test, or tampering with the device were subject to an immediate stay in jail, typically a night or two. A positive alcohol test under breathalyzer-based supervision is a BAC greater than .02; under alcohol bracelet monitoring the threshold is a sustained positive alcohol reading that has been confirmed by external expert review (SCRAM Systems 2019).

24/7 Sobriety was originally launched in South Dakota to reduce incarceration, especially in state prisons. Circa 2003, the state's attorney general recommended a pilot program that would make abstinence a condition of bond for repeat-DUI arrestees and require them to come to the county jail and blow into a breathalyzer once in the morning and again in the evening *every day*. If there was any alcohol at all in their system, they would immediately go to jail, but only for a night or two. Over time, the program expanded in terms of counties of operation, types of offenses, and alcohol testing technologies. Unlike drug and DUI courts, 24/7 Sobriety did not require participants to attend treatment; the program only mandated abstinence (Long 2009).²

Emerging evidence on 24/7 Sobriety in South Dakota suggests that the program reduced problem drinking. From 2005 to 2017, more than 30,000 South Dakotans participated in the original 24/7 Sobriety. More than 99% of the breathalyzers tests were taken and passed (including no-shows in the denominator) and more than 99% of the days on the bracelet have neither confirmed alcohol use nor an attempt to tamper with the device (South Dakota Attorney General 2019). While the absence of heavy drinking was the most proximate outcome, emerging evidence also suggests that the program was associated with reductions in alcohol-involved consequences. Using county-level variation in timing of 24/7 Sobriety implementation in South Dakota, Kilmer et al. (2013) documented a 12% reduction in repeat-DUI arrests and a 9% reduction in arrests for domestic violence at the county level. An individual-level analysis, which used the availability of 24/7 Sobriety in the county as an instrumental variable for program participation, found that 24/7 Sobriety leads to large and lasting reductions in the probability of rearrest among repeat DUI arrestees: nearly a 50% reduction in the probability of rearrest or probation revocation at 12 months, and potentially evidence of sustained reductions up to 36 months (Kilmer and Midgette 2020). Looking beyond DUI and other traditional criminal justice outcomes, Nicosia, Kilmer, and Heaton (2016) found that 24/7 Sobriety was associated with a reduction in total mortality at the county level. The program has also been a fiscal success, with fees from participants more than covering the costs of the program from the state's perspective (Midgette 2014).

North Dakota's pilot 24/7 Sobriety program began in 2008 in the South Central and North East Central Juridical Districts comprising 14 of the state's 53 counties. Although roughly 75% of the state's population lived in counties that could require 24/7 Sobriety, assignment to the program grew slowly, in part due to judicial discretion (Kubas et al. 2016). In April 2009, House Bill 1306 took the program statewide and allowed enrollment of any individual arrested for "offenses in which alcohol or controlled substances are involved" (North Dakota Attorney General 2019: 9). To generate support for 24/7, counties were allowed to tailor the program. For example, breathalyzer-based testing windows varied across counties as did the eligibility criteria for the monitoring technology (i.e.,

² While innovative, the program was not without controversy. The program required participants to pay \$1 per breathalyzer test or \$6/day for alcohol monitoring bracelets (plus fees for applying and removing the bracelet), thereby raising concerns about financial burden for participants. The program developers responded, however, that participants were generally spending more than that on alcohol each day.

in-person breathalyzer or body-worn continuous alcohol monitoring [CAM] bracelets) to suit both staffing constraints and the needs of the population. Further, some counties integrated drug monitoring via transdermal sweat patch or urinalysis while others did not.

In mid-2013, North Dakota House Bill 1302 (HB1302) mandated participation terms for 24/7 Sobriety among DUI arrestees. Anyone arrested for a DUI would lose their license for a set period; those who wished to regain driving privileges were required to enter 24/7 Sobriety. First-time offenders were now subject to the program for a six-month term, repeat offenders were subject for a 1-year term, offenders arrested with BAC over .18 or with two or more prior DUI offenses were subject for a two-year term, and those with two or more priors and BAC over .18 were subject for a three-year term. These terms were longer and, tautologically, less varied than prior assignment spells. Thus, HB1302 led to both longer spells and higher rates of participation among DUI offenders.

There were some important differences in how 24/7 Sobriety was implemented in North Dakota vis-à-vis South Dakota. First, whereas any trace of alcohol comprised a violation in South Dakota, North Dakota allowed up to the equivalent of a .02 BAC on a breathalyzer test. Thus, the North Dakota program targeted heavy drinking in contrast to South Dakota's approach targeting any drinking. The .02 BAC failure threshold is the empirical threshold for risk of alcohol-related crash risk (Nochajski and Stasiewicz 2006), but may be less effective at preventing all drinking than South Dakota's 0.00 BAC threshold. For participants who might "game" the program by limiting their alcohol consumption just enough to pass their next test, it's harder to hit the .02 target than 0, especially given that cognitive ability is inversely related to ethanol consumed (Dawson and Reid 1997). More generally, this small policy difference might carry implications for the frequency of violations and alcohol-related consequences.

Second, monitoring via a continuous alcohol monitoring (CAM) bracelet was more common in North Dakota than South Dakota (62% of monitoring days vs 30%).^{3,4} Under CAM, participants incurred greater monitoring fees, but its remote testing allowed for fewer in-person visits vis-à-vis breathalyzer. Third, drug testing was more commonplace in North Dakota, though combined testing for alcohol and other drugs was rare (2%).

To date, there has yet to be a rigorous empirical evaluation of the causal impact of 24/7 Sobriety on DUI crime in North Dakota; this paper seeks to fill that gap. Three prior studies conducted by the NDSU Upper Great Plains Transportation Institute examined how North Dakota's 24/7 Sobriety influenced dangerous driving behaviors. Using 24/7 Sobriety records linked with drivers' license records on crashes and convictions, the authors documented statistically significant improvements in dangerous driving behaviors (e.g. DUI, crashes) using pre-post comparisons of means among individuals assigned to 24/7 Sobriety, but the improvements did not hold for high and moderate-risk offenders (Kubas et al. 2015). Regression results among program participants confirmed lower levels of dangerous driving after program participation, especially among those enrolled for longer terms under HB1302 (Kubas et al. 2017; Vachal et al. 2018). These studies offer comprehensive descriptive analyses, but

³ In North Dakota, participants monitored for alcohol use typically begin using breathalyzers on the day they are enrolled, but are eligible for CAM after completing between week and a month of compliance to the breathalyzer-based program.

⁴ According to interviews conducted with program administrators in the state, this tendency is a pragmatic response to mitigate the administrative burden on staff, to reduce commute burden for those participants in rural areas or those who travel often for work (e.g., periodic work in the Bakken oil fields in the state's northwest).

the studies do not attempt to measure the program's causal impact. The simple pre-post analyses are suggestive of the program's effect, but potential endogenous selection into 24/7, and the need for a comparison group to control for general trends in dangerous driving beyond 24/7 Sobriety participants, necessitate a quasi-experimental research design.

Swift, Certain, Fair Programs in Replication

Establishing whether the 24/7 Sobriety effects measured in South Dakota are observable in North Dakota is of considerable import due to ambiguous findings for other replications of the swift and certain sanctioning model. Perhaps the best known and most celebrated program of this type, Hawaii's Project HOPE, generated high rates of compliance and sizeable reductions in recidivism (Hawken and Kleiman 2009). Proponents have argued that the program's success is attributable to a superior incentive architecture that increases deterrence without increasing incarceration.

Authors analyzing the four-site randomized replication of HOPE on the mainland reported no benefits to HOPE in terms of reducing arrests, convictions, and incarceration, and concluded that "HOPE/SCF seems unlikely to offer better outcomes and lower costs for broad classes of moderate-to-high-risk probationers" (Lattimore et al. 2016: 1103). However, after reviewing the complete analysis of the replication (Lattimore et al. 2018) and examining outcomes by implementation fidelity, Humphreys and Kilmer (2020) concluded site differences suggest that some outcomes varied by implementation success. Further, although the main article in *Criminology and Public Policy* included the largely negative pre-registered outcomes, it did not include the pre-registered substance use outcome, which was significantly improved.

Hawken (2016) posits that the HOPE model may require adaptation to important local-level differences across jurisdictions, and further notes that results based on immediate measurement of the program's impacts without first adapting the program to jurisdiction-specific implementation challenges through a pilot may understate the true effectiveness of the approach. The success of HOPE in Hawaii may be also due to idiosyncratic factors, such as the characteristics of the judge and other administrators overseeing the program, rather than or in conjunction with a better incentive structure. Alternatively, incentives may matter, but HOPE and its replications do not actually offer sufficiently swift and certain punishment required to fully leverage the deterrent possibilities of the model; the sanction is swift, but not immediate and random drug testing yields a probability of detecting a violation less than one (Cook 2016).

24/7 Sobriety arguably offers greater swiftness of detection than even HOPE, providing a better test of the underlying deterrence model. Should the North Dakota program successfully reduce intoxicated driving, it would provide new evidence that the swift and certain sanctioning approach works due to its appeal to the basic decision processes of offenders.

Methods

Data

Program participation data were provided by the North Dakota Attorney General for all enrollees through December 2015 including the dates of program initiation, termination,

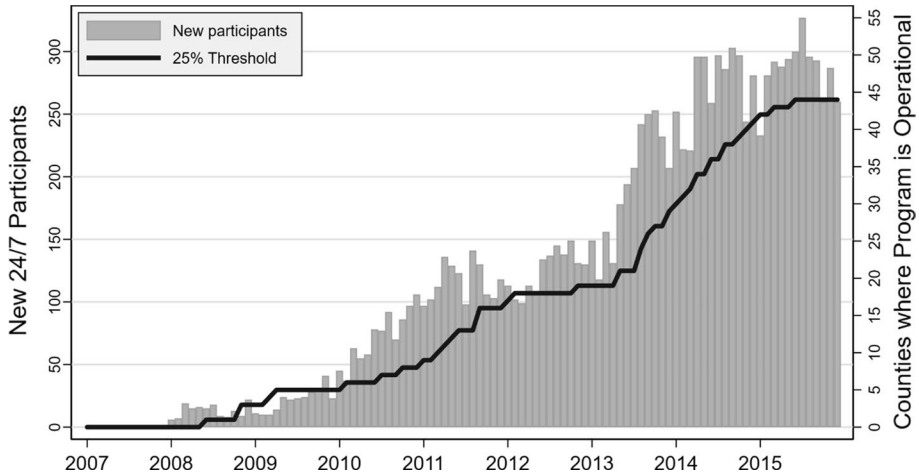


Fig. 1 Participation in North Dakota's 24/7 Sobriety Project, 2008–2015 (Note We define 24/7 as operational in each county once it equals or exceeds the threshold number of program participants per the number of DUI arrests in the county, where the latter is defined as the county's moving monthly average during the previous year to address any seasonality)

and all scheduled tests and results.⁵ Individuals may choose to be tested in counties other than those where they live. For our primary analysis we organize the data by county of enrollment, under the assumption that those who enroll outside of their home county do so because they are likely more often present in the county of enrollment due to work or other obligations, meaning that they pose a greater risk in that area.⁶ Following Kilmer et al. (2013), we defined 24/7 Sobriety as operational in each county once the number of county participants in 24/7 Sobriety for a given month equaled or exceeded a threshold share of the average number of DUI arrests in the county over the previous year. Figure 1 displays monthly new program participants and the count of counties reaching threshold share number of participants over time, varying that share between 10 and 40%. It also includes the county count based on the first participant in each county's 24/7 Sobriety program. By the end of 2015, 44 of the state's 53 counties had operational programs based on the 25% threshold, and individuals had been assigned to the program more than 13,000 times.⁷

Three-quarters of participants were male and three-quarters were between 21 and 45 years old, with a median age of 31 (Table 1). Fifty-eight percent of participants were enrolled in the program for a DUI offense, about one-quarter for a drug or paraphernalia possession arrest, and approximately 9% for a domestic violence or assault arrest.

Our primary outcome variable is the count of DUI arrests reported to the North Dakota Department of Transportation between 2004 and 2015. These administrative records do not

⁵ All participants listed as active as of the data censoring date were assumed to be participating as of the end of 2015, the study period.

⁶ The share of participants that participate in a county other than their residence is small. For example, individuals commute long distances to work in the Bakken oil fields; approximately 7% of those who participate in the Bakken region report residing elsewhere.

⁷ Administrative data provided by the State of North Dakota did not include identifying characteristics, so repeat enrollment under multiple participant identification numbers cannot be determined.

Table 1 24/7 Sobriety participant characteristics through December 2015

Participants	13,164
Male (%)	75.9
Median age	31
Offense leading to 24/7 Sobriety assignment (%)	
DUI	58.0
Drug possession	24.6
Assault, domestic violence, or abuse	9.4
Other	7.9
Participant testing method (% share of days)*	
Breathalyzer	30.5
CAM	61.6
Drug patch and urinalysis	27.8

Based on participant outcome information reported by North Dakota Attorney General for 13,496 participants; testing information for 332 participants was not included in administrative data. DUI includes charges for actual physical control (10.7%) and driving under suspicion (0.6%); Drug possession includes paraphernalia possession (7.6%)

*Testing media may be simultaneous so categories are not mutually exclusive

Table 2 Descriptive statistics (2004–2015; n = 7620)

	Mean	SD	Minimum	Median	Maximum
DUI arrests	9.41	19.67	0	2	153
24/7 Sobriety operational (1 = yes)	.228	.420	0	0	1
Percent white	91.86	16.21	13.54	97.01	99.47
Percent male 18–40	13.22	3.28	8.48	12.59	24.94
Sworn police per 10,000 residents	12.48	8.20	0	10.28	6.45
Bars per 10,000 residents	9.84	5.26	0	9.04	31.76
Package Stores per 10,000 residents	1.74	1.89	0	1.50	10.04
College (1 = in session)	.028	.166	0	0	1
Bakken labor expansion (1 = yes)	.080	.271	0	0	1
Exposure: population	12,833	25,069	643	4172	174,689

differentiate between alcohol and drug-related DUI arrests, and in the state, alcohol and drug-related DUI are governed by the same law (North Dakota Century Code 39-08-01). These data were reported monthly.^{8,9} The median number of DUIs per county-month is 2 with a mean of 9.41 over the period from 2004 to 2015 (Table 2).

The counties average ten bars and nearly two package stores per 10,000 residents. Nearly 92% of the state population is white, though there are several counties that overlap

⁸ According to state officials, these data were the most complete and accurate records of DUI arrests available and generally include larger counts of offenses than FBI Uniform Crime Reports.

⁹ Data for calendar year 2007 were not available for this analysis, so we interpolated missing values using a Poisson-based multiple imputation. The findings reported are robust to model specification and imputation method.

with Native American reservations. Roughly one-eighth of the state population are males 18–40 years of age. Additionally, we included an indicator for whether college was in session in Cass (North Dakota State University) and Grand Forks (University of North Dakota). Finally, we defined an indicator of Bakken oil field employment expansion to account for the influx of primarily young males to relatively high-paying, high-stress oil industry jobs in the five counties comprising the Bakken Formation in the State’s north-west corner during the oil boom over January 2009 to April 2015 (Federal Reserve Bank of Minneapolis 2019).¹⁰

Empirical Strategy

We first perform a descriptive analysis of the 24/7 Sobriety program at the test-event and individual levels. These data cover 8103 participants monitored using breathalyzers, 4015 using CAM, 3435 using transdermal drug patches, and 228 using in-person urinalysis screening. We report violation rates for each testing medium. Since participants may be assigned to breathalyzer or CAM for alcohol monitoring, may simultaneously be assigned drug patch or urinalysis, and may switch testing medium during participation, the counts of individuals are not mutually exclusive.

The main model estimates the relationship between 24/7 Sobriety and county-month DUI arrest counts using a difference-in-differences approach defined as:

$$Y_{it} = \alpha 24/7_{it} + \beta X_{it} + \gamma_i + \delta_t + \varepsilon_{it}, \quad (1)$$

where Y_{it} represents the count of DUI arrests in county i and month t . The coefficient α captures the effect of 24/7 Sobriety based on the indicator $24/7_{it}$, which is equal to 1 for all county-month combinations when 24/7 Sobriety was defined as being operational. As controls, we include a matrix of time-varying county-level variables (X_{it}) shown in Table 3: indicator variables for the Bakken oil expansion or whether the state’s large universities were in session, bars per capita, package stores per capita, sworn police officers per capita, log-transformed county population, percentage of population identifying as white, and percentage of population between 18 and 40 years of age identifying as male. The specification also includes county fixed effects (γ_i) to capture time-invariant unobservable factors as well as fixed effects for each month (δ_t) to control for seasonal and temporal factors common to all counties.

Our identification strategy assumes the program’s rollout across the state is exogenous to DUI counts. Based on interviews with program administrators in the state, this appears to be the case, as early implementation seems idiosyncratic, rather than based on particularly acute need for policies effective against DUI. Within the pilot counties, higher-risk arrestees may have been more likely to enter the program early in its growth, but potential statistical bias in our estimates is ambiguous, as the passage of HB1302 mandates all repeat-DUI arrestees to be assigned to 24/7 Sobriety by August 2013 (Kubas et al. 2017).

¹⁰ We define the Bakken Labor Expansion variable to equal one for Dunn, McKenzie, Stark, and Williams county between March 2010 and April 2015 and zero otherwise.

Table 3 24/7 Sobriety testing outcomes through December 2015

Test result characteristics	
Breathalyzer (test-level)	837,327 tests
	426,198 days
% Passed	95.6
% Positive for Alcohol	0.5
% Missed	1.9
% Excused	2.0
Breathalyzer (n = 8103 participants)	
% Never fail	53.5
% One failure	14.5
% Two failures	9.8
% Three failures	4.8
% Four or more failures	17.4
CAM (n = 4015 participants)	673,925 days
% Never fail	67.4
Drug Patch (n = 3435 participants)	30,392 tests
% Passed	92.0
% Positive for Drugs	7.5
% Excused	0.5
Urinalysis (n = 228 participants)	5185 tests
% Passed	97.2
% Positive for Drugs	1.5
% Excused	1.3

Participants may test using multiple media so testing media counts are not mutually exclusive

Figure 2 displays the roll-out of the program across counties over time based on the 25% threshold criterion. There is no apparent spatial pattern to roll-out.

While such selection would undermine an individual-level analysis, its effect on this community-level analysis is ambiguous. The effect on the estimates depends on whether certain type of individuals (e.g. high risk of re-offending) were more likely to be assigned early in the program and on heterogeneity in responsiveness to the program. Our approach guards against upward bias by defining 24/7 Sobriety operationalization as a minimum of 25% enrollment of DUI arrestees (as well as alternative thresholds in sensitivity analyses).

We estimate the main models using panel fixed effects Poisson regression to provide consistent estimates of the conditional mean function across a wide range of potential data generating processes (Wooldridge 2010). We report two sets of 95% confidence intervals for our main results based on alternative assumptions about potential spatial autocorrelation between counties. First, we consider cluster-robust standard errors proposed by Bertrand et al. (2004) at the county level to allow for valid statistical inference in case of within-county autocorrelation or failure of the Poisson equal mean–variance assumption over time. However, these standard errors may be inappropriate in the presence of spatial autocorrelation, e.g. potential displacement of drunk driving to neighboring counties in response to the enactment of 24/7 Sobriety. To confront potential spatial autocorrelation, Bester et al. (2011) demonstrate that clustering at a small number of large geographic units is asymptotically equivalent to accounting for arbitrary spatial autocorrelation as the

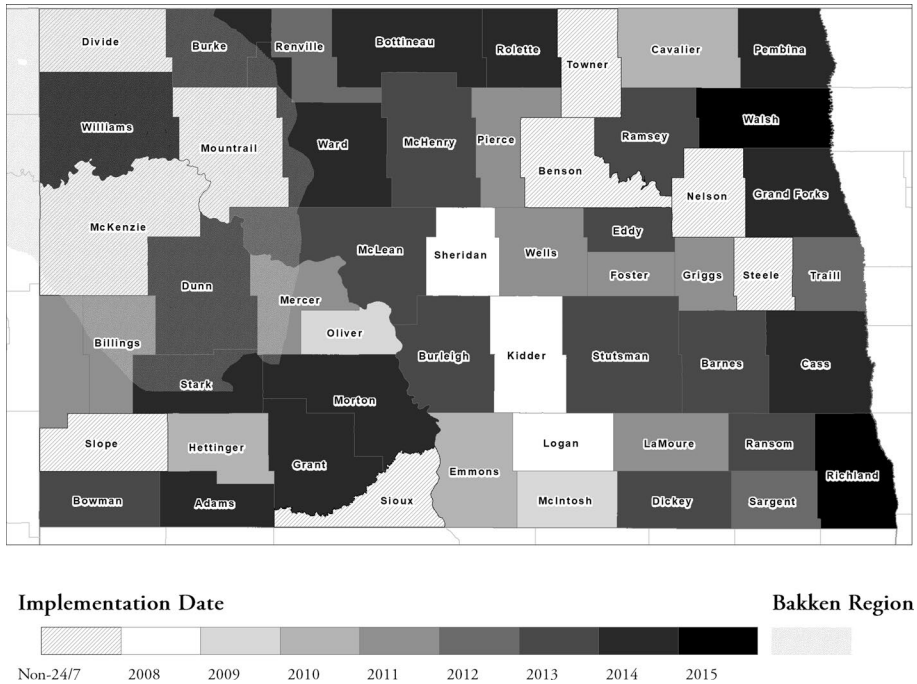


Fig. 2 Map of 24/7 Program roll-out across counties based on the 25% threshold criterion

number of sub-units grows large. Ridgeway et al. (2019) apply this method in the criminal justice context when testing the impact of Los Angeles Police Department gang injunctions in reporting districts by clustering at the larger bureau level. We in turn generate alternative standard errors by clustering at North Dakota’s eight judicial districts.

We separately report standard errors based on a permutation test to assess whether the effects reported are due to 24/7 Sobriety implementation rather than chance. This method randomizes treatment assignment such that all units are independent. We assign an alternative 24/7 Sobriety “implementation” date for each county in our sample by randomly switching program implementation dates among all counties in the state. We then re-estimate Eq. (1) to obtain a new “program effect” estimate over 1000 iterations. This approach essentially allows us to ascertain whether the particular dates of 24/7 Sobriety implementation in each county seem to be unusual relative to surrounding dates in terms of subsequent changes in alcohol-related crime. All analyses were performed in Stata/MP version 16.1 (StataCorp LP, College Station, TX).

Results

Participants were monitored for alcohol in North Dakota for just over 1.1 million days between 2008 and 2015 (see Table 3). Roughly 60% of those days were via CAM. Over 95% of breathalyzer tests were taken and passed; and, among violations, missed tests were four times as common as positive tests (consecutive skipped tests are counted separately). Over half of participants never tested positive via breathalyzer (53.5%) and two-thirds of

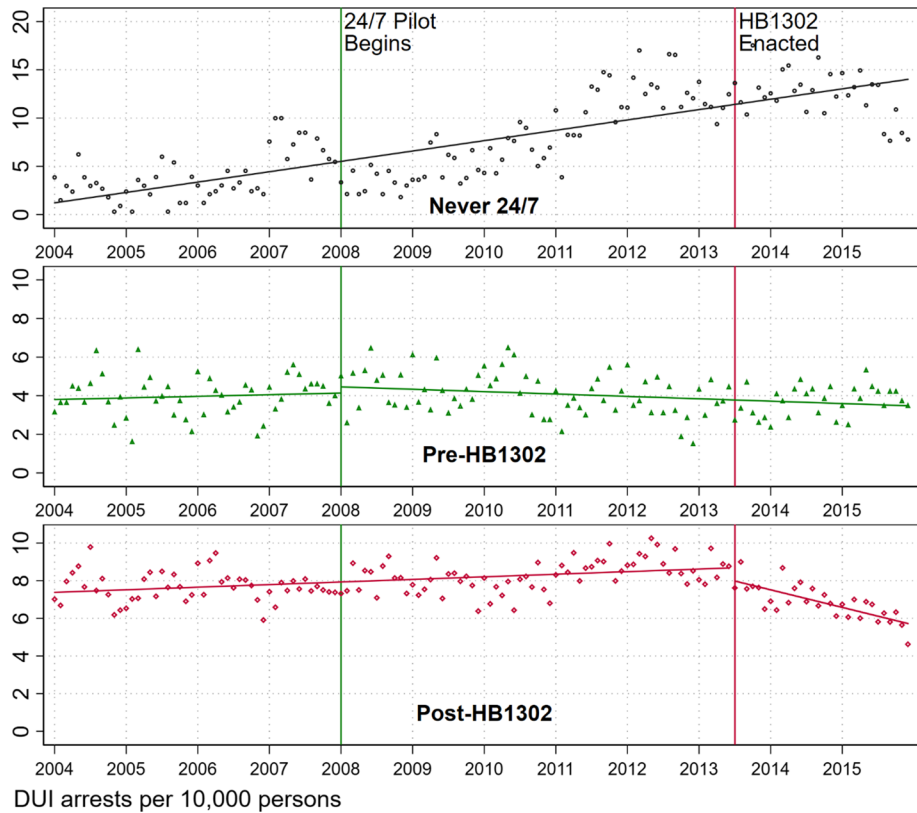


Fig. 3 Pre- and post-enactment DUI rates by 24/7 Sobriety start period based on 25% threshold

CAM participants (67.4%) completed the program without a confirmed violation.¹¹ Compared to South Dakota 24/7, the test-level violation rate is notably higher in North Dakota. However, at the individual-level, non-compliance rates are similar to those seen in South Dakota 24/7. This suggests that differences in participants or the program do not impact the extensive margin—the share of participants that violate the program—but do affect the intensive margin—the frequency of violations among violators. Among those monitored for illegal drugs other than alcohol, 92% of participants monitored using the drug patch are fully compliant and 97.2% of urinalysis screens were negative. This is a surprisingly high compliance rate, especially given the possibility that early drug tests may even detect pre-arrest and pre-enrollment drug use.

We test the pre-trend equivalence assumption desired for difference-in-differences models and the policy exogeneity assumption by splitting DUI per capita rates in three groups of counties; (1) those that never meet the 25% threshold for active programs in our study period; (2) those that met the threshold after the pilot began in January 2008 and before

¹¹ The cost of each testing medium to participants differs. The conditions under which participants are assigned to each testing mechanism may also vary in unobservable ways. Thus, we caution against assessments of relative effectiveness without a randomized controlled experimental design.

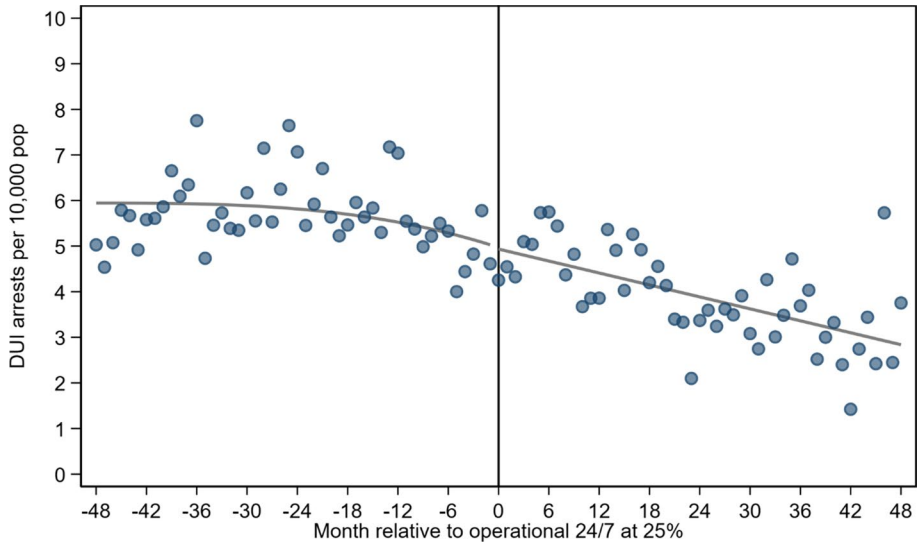


Fig. 4 Trend in the 48 months before and after county-level program activation

the enactment of HB-1203 in August 2013; and (3) those that met the threshold after August 2013. Figure 3 shows similar trends in the pre-24/7 Sobriety period between 2005 and 2008, as well as some evidence of reduction in per capita DUI after implementation for each group that eventually implements 24/7 Sobriety. Before the introduction of 24/7 Sobriety in 2008, there is no difference in per capita DUI arrest trends across groups. There is also no difference in levels between counties that never adopt 24/7 Sobriety and those that adopt prior to HB1302. Counties that implemented 24/7 Sobriety after HB1302 generally had higher DUI rates than the other groups. This suggests that early adoption was not related to escalating need for DUI intervention, an important assumption in our identification strategy.

We next plot the trend in DUI arrests per capita by centering the month when a county crosses the 25% enrollment threshold at zero, again testing the identifying assumption that the change in DUI arrest rates should be associated with enactment of the policy (Fig. 4). There is visual evidence of a reduction in rate coinciding with the date we define counties to be operational based on the 25% enrollment threshold. However, for many counties, rollout of 24/7 Sobriety progressed over a period of time spanning a number of months; on average, six months pass between the time that a county in North Dakota first assigns a participant to 24/7 Sobriety and when the county reaches the 25% threshold, so typically enrollment began several months prior to the county reaching the implementation threshold. We assess the sensitivity of our findings to this phenomenon with a set of robustness checks.

Effect of 24/7 Sobriety on DUI Arrests

In an unadjusted model with only the 24/7 Sobriety policy variable and panel fixed effects as explanatory variables (Table 4), the estimated incident rate ratio (IRR) for the policy indicator is .735. In this model, we find that the clustering based on county results

Table 4 Estimated impact of 24/7 Sobriety on DUI arrests based on 25% threshold

	DUI arrests	DUI arrests
24/7 Sobriety	0.735 [0.622, 0.869] [0.571, 0.947]	0.912 [0.856, 0.973] [0.875, 0.951]
Percent white		1.037 [0.995, 1.082] [1.005, 1.071]
Percent male 18–40		1.073 [1.045, 1.102] [1.047, 1.100]
Sworn police per capita		1.017 [0.996, 1.037] [1.001, 1.032]
Bars per capita		0.983 [0.960, 1.007] [0.955, 1.012]
Package stores per capita		0.973 [0.916, 1.034] [0.908, 1.043]
College		1.122 [1.086, 1.160] [1.084, 1.162]
Bakken labor expansion		1.216 [1.039, 1.422] [1.149, 1.286]
N	7620	7620

95% confidence intervals based on cluster-robust standard errors in brackets; top interval clustered at the county-level, bottom interval clustered at the Judicial District-level. All models include county and month fixed effects and log-population as exposure

in marginally narrower standard errors [95% CI 0.622, 0.869] than clustering at the judicial district level [0.571, 0.947]. In our preferred specification that controls for time-varying covariates, the IRR estimate for 24/7 Sobriety is 0.912, indicating the DUI rate decreases by approximately 9% when the program is active. Here county clustering results in a slightly wider 95% confidence interval [0.856, 0.973] than judicial district clustering [0.875, 0.951], suggesting spatial autocorrelation is not a salient concern for our inference. Based on the permutation test, 95% of simulated effect sizes were between 0.913 and 1.084, with mean 1. Our actual estimated effect of 0.912 is in the third percentile of simulated outcomes (see Fig. 5), suggesting our results are unlikely to be due to chance, and instead that a pattern of decreased DUI is tightly tied to the particular months the program was implemented in each county.

Among the covariates, the proportion of the population that is male between 18 and 40, the indicator for whether the state's large universities were in session, and Bakken-related market expansion were found to be positively associated with DUI arrests. Of particular note, even after conditioning on population, labor market expansion in the Bakken region

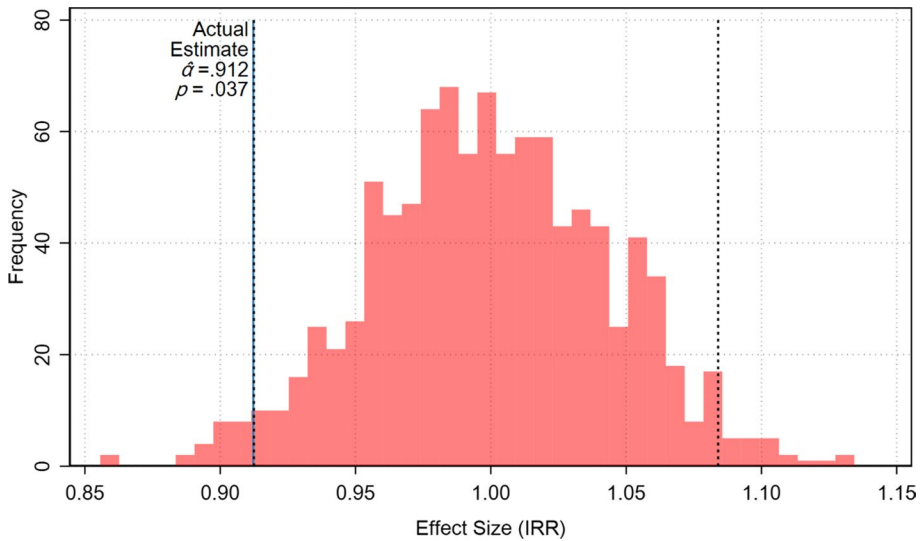


Fig. 5 Permutation test of the estimated impact of 24/7 Sobriety implementation on DUI arrests (empirically defined 95% confidence interval denoted by dotted lines)

is associated with a 22% increase in DUI rates in affected counties. The DUI arrest rate also increases by about 12% in the counties that are home to the state’s two largest universities when school is in session. Bars per capita has a statistically significant but small relationship with county DUI arrest rates.

Robustness Checks

To assess the robustness of the findings, we conduct seven sensitivity analyses: (1) utilizing a negative binomial model instead of Poisson, (2) removing counties with histories of non-reporting, (3) changing the operationalization of the implementation variable based on the participants’ county of residence (versus participation), (4) including county-specific time trends, (5) excluding counties that did not implement 24/7 Sobriety over the study period, and (6 and 7) alternating the threshold to define 24/7 Sobriety as operational between 10 and 40%. We present confidence intervals based county clustering, as they are uniformly more conservative than those produced by judicial district clustering. These model estimates are summarized in Table 5.

Negative Binomial

Poisson regression is often preferred when modeling count processes for the simplicity of assumptions: the conditional mean of the outcome of interest is equal to the conditional variance. Many small counties in North Dakota record zero DUI arrests each month. Due to this feature of the data, we find empirically that a Poisson model is likely still consistent, but inefficient (Pearson $\chi^2 = 12601.8$; $p < .001$). The negative binomial is a frequently used alternative count model that will be more efficient if the underlying data generating process is negative binomial rather than Poisson, though it will not solve potential bias or inconsistency in our estimators (Berk and MacDonald 2008). We find that under this flexible

Table 5 Sensitivity analyses

	(0) Base model	(1) Negative binomial model	(2) Large counties only	(3) Residence county	(4) County-specific trends	(5) Exclude never- takers	(6) 10% threshold	(7) 40% threshold
24/7 Sobriety	0.912 [0.856, 0.973]	0.855 [0.805, 0.909]	0.920 [0.864, 0.979]	0.959 [0.898, 1.025]	0.772 [0.741, 0.804]	0.917 [0.862, 0.976]	0.983 [0.926, 1.045]	1.009 [0.892, 1.142]
Percent white	1.037 [0.995, 1.082]	1.012 [1.006, 1.018]	1.035 [0.991, 1.080]	1.039 [0.995, 1.085]	1.023 [1.001, 1.046]	1.02 [0.986, 1.055]	1.04 [0.995, 1.087]	1.073 [1.030, 1.117]
Percent male, 18–40	1.073 [1.045, 1.102]	1.169 [1.155, 1.183]	1.074 [1.046, 1.103]	1.075 [1.046, 1.104]	1.065 [1.038, 1.092]	1.080 [1.049, 1.112]	1.076 [1.048, 1.105]	1.063 [1.022, 1.104]
Sworn police per capita	1.017 [0.996, 1.037]	0.995 [0.988, 1.002]	1.009 [0.988, 1.030]	1.017 [0.997, 1.037]	0.99 [0.980, 1.001]	1.018 [1.005, 1.031]	1.017 [0.997, 1.036]	1.014 [0.999, 1.029]
Bars per capita	0.983 [0.960, 1.007]	0.957 [0.947, 0.967]	0.981 [0.952, 1.011]	0.982 [0.959, 1.006]	0.977 [0.965, 0.990]	1.001 [0.980, 1.021]	0.982 [0.959, 1.006]	0.988 [0.962, 1.013]
Package stores per capita	0.973 [0.916, 1.034]	0.950 [0.929, 0.972]	0.962 [0.897, 1.032]	0.972 [0.915, 1.033]	0.962 [0.930, 0.995]	1.018 [0.964, 1.075]	0.972 [0.914, 1.034]	1.04 [0.971, 1.115]
College	1.122 [1.086, 1.160]	1.160 [1.090, 1.236]	1.122 [1.085, 1.160]	1.123 [1.087, 1.161]	1.04 [0.996, 1.086]	1.119 [1.083, 1.156]	1.125 [1.088, 1.163]	1.149 [1.101, 1.199]
Bakken labor expansion	1.216 [1.039, 1.422]	1.074 [0.998, 1.156]	1.233 [1.061, 1.433]	1.211 [1.042, 1.407]	1.312 [1.230, 1.399]	1.171 [1.086, 1.263]	1.206 [1.031, 1.411]	1.219 [1.131, 1.314]
N	7620	7620	5371	7620	7620	6336	7620	6180

95% confidence intervals based on cluster-robust standard errors in brackets. All models include county and month fixed effects and log-population exposure

alternative model, 24/7 Sobriety is associated with a 14.5% reduction ($IRR = .855$ [0.805, 0.909]) in DUI arrests with statistical precision such that its confidence interval overlaps with our main result found by Poisson, but not with a null effect. The difference across these estimates in part reflects different weighting structures—variance is modeled as a linear function of the mean under Poisson compared to a quadratic function under negative binomial (Ver Hoef and Boveng 2007).

Exclude Small Counties

Due to the low counts of monthly DUIs in small counties, small changes in levels (e.g., movements among single-digit arrest counts) are expressed as percentage changes. The estimates are not sensitive to bias of this type by virtue of the log-difference approach in regressions for count data such as the negative binomial, but that assumes all arrests are reported. Some smaller counties may not report arrests at all in some months, even when they occur. To assess whether such non-reporting, if it exists, might contaminate our results, we repeat the analysis on the subset of counties with populations over 2500 persons. This is generally equivalent to the subset of counties that report at least one DUI arrest per month on average. We find that the estimated effect is of similar magnitude and precision as our main result ($IRR = 0.920$ [0.864, 0.979]).

County of Residence Versus Enrollment

Earlier we discussed the potential importance of defining 24/7 Sobriety as operational based on the county of participation rather than residence. The concern revolves around counties with relatively few DUI arrests that would be more likely to qualify falsely as operational because some residents will neither have been arrested nor participated the program in the county of residence. Therefore, we would expect less precise estimates when defining program status based on county of residence vs participation. Our results are consistent with that hypothesis ($IRR = 0.959$ [0.898, 1.025]).

County-Specific Time Trends

The main model includes county and month fixed effects, but one might instead assume that counties have unique time trends with respect to DUI counts. When county-specific time trends are included, the estimated effect of 24/7 Sobriety increases considerably in magnitude ($IRR = .772$ [0.741, 0.804]) and remains precise.

Exclude Counties that Did Not Implement 24/7

Considering the pre-trends across treatment groups (displayed in Fig. 3), the counties that never take up 24/7 Sobriety in the study period appear to experience an increase in DUI rates in 2007, the year before the 24/7 Sobriety pilot began. To test whether this difference in pre-trend is driving our main result, we exclude the never-takers from the sample as a

separate analysis and find nearly identical results as our main finding (IRR = .917 [0.862, 0.976]).

This analysis also reveals an area for future study related to program heterogeneity. Among the counties that implement 24/7 Sobriety at some point, compared to the most populous county (Cass) the estimated IRR from the county-level fixed-effects vary between 0.168 and 3.873. This describes significant variation in DUI rates across counties, and suggests characteristics including particular program implementation choices, the composition of participant, and environmental factors may play an important role in efficacy. We do not have sufficient data to study these factors rigorously.

Alternating the Threshold to Define 24/7 Sobriety as Operational

As the 25% participation threshold used to define 24/7 Sobriety as operational is only based on previous studies (Kilmer et al. 2013; Nicosia et al. 2016) and not on any official definition, we test the sensitivity of that decision by redefining the threshold at the lower level of 10% and higher level of 40%. Neither estimate was statistically significant, which differs from the findings Kilmer et al. (2013) report in South Dakota.

The lower 10% threshold indicates that many potentially eligible DUI offenders are untreated in counties included as treated in the model. We are thus essentially measuring a response from a relatively small dose. By complementary logic, a higher threshold will falsely classify a larger proportion of counties with any share under 40% of eligible arrestees assigned to 24/7 Sobriety participants as untreated. In this latter situation, the effect of the program is biased toward a null finding if the program's impact could be detected at a lower threshold (i.e., the 25% threshold), and if several months pass between when a county crosses from one threshold to the next. Unlike the relatively rapid expansion in South Dakota, the program expanded out slowly in North Dakota. An average of 13.8 months passed between when a county crossed the 10% threshold and the 25% threshold, and subsequently another 9.1 months passed before crossing the 40% threshold.¹² Thus, we might expect a nonlinear relationship between estimated treatment effect and the threshold definition. The results are consistent with this assumption. The effect we estimate is thus more sensitive to our empirical specification on this dimension than Kilmer et al. (2013) found in their evaluation of 24/7 Sobriety in South Dakota.

Conclusions

The vast majority of 24/7 Sobriety participants in North Dakota submit and pass their tests and the majority complete the program without a detected violation. The 13,164 participants passed more than 95% of the alcohol tests that were administered and accumulated approximately one million days under supervision without a drinking event, and over 90% of all drug tests that were administered were passed. This implied reduction in heavy drinking and other drug use is striking given the high-risk nature of this vulnerable population. Looking beyond substance use, we also find quasi-experimental evidence of a reduction

¹² This leads to a small set of counties breaking the 40% threshold by the end of the study period, so we subset the data to counties that either never met the threshold or had at least twelve months of post-enactment data (see Fig. 1). This resulted in exclusion of 19% of the sample.

in DUI arrests following the operationalization of 24/7 Sobriety in the counties. The point estimate for this effect is somewhat smaller than the effect found by Kilmer et al. (2013), though not statistically differentiable. However, this analysis uses a different data source and slightly different outcome measure.

Given the lower rate of test compliance in North Dakota than South Dakota, a smaller but still non-trivial program impact is consistent with expectations through the lens of deterrence theory, though we caution that several components of the North Dakota program suggest it is not directly comparable. First, North Dakota's 24/7 Sobriety program has a large and distinct drug testing component which suggests it may serve a higher-risk drug-using population than its southern neighbor. Second, the data series from which we measure the effect include first-time DUI arrests and potentially higher rates of drug-involved DUI, which was not reported separately in North Dakota over the timeframe for this analysis.

In terms of broader generalizability, the status quo for criminal justice and public health programs to prevent impaired driving varies from state to state, as does the base rate of offending. The Dakotas may represent the upper range of 24/7 Sobriety's effect size. In particular, the large increase in high-risk pool of high-earning, young male drivers due to activity in the Bakken is unique to the region. We demonstrate that DUI arrests in the Bakken region spike during periods of high associated employment.

In this study, we find that the North Dakota program was less effective at deterring alcohol consumption than its neighbor to the south as measured by breathalyzer test violation rates, and may be marginally less effective at preventing intoxicated driving.¹³ While the compliance rates in both states are notably high, the difference in the share of taken and passed tests in South Dakota (99%) and North Dakota (95%) is non-trivial, especially when each violation results in a brief jail admission which is costly for both the participant and the jurisdiction. Given the similarities in the populations and policies across the Dakotas, the increase in test violations may be due, at least in part, to the change in the alcohol violation threshold from zero to 0.02 BAC. The higher threshold is more permissive, so any violation in South below 0.02 BAC would not result in a violation in North Dakota. However, the laxer standard may invite more risk-taking. This is an area for future research.

Regarding biases in participant decision-making, we may be able to better understand the mechanisms leading to the decision to consume alcohol under enforced prohibition—in this case through 24/7 Sobriety—and the decision to drive while intoxicated. Nagin and Pogarsky (2004) establish that high discounting was a factor in actions requiring relatively more forethought, while urge-driven behaviors are dominated by poor impulse control. Here, the decision to drive while intoxicated may be considered the former, and the decision to drink is the latter.

In a recent critique of the HOPE model, Cullen et al. (2016: 120) note, "The language of swift, certain, and fair is misleading or, in the least, open to debate." Their critique of HOPE may too apply to 24/7 Sobriety. Penalties under HOPE, and 24/7 Sobriety, are swifter than the status quo, but are not immediate. Criminal offenses and program violations that go undetected erode certainty, and jail time for a missed test or consumption of alcohol may not seem fair to many observers. We argue that the elements of 24/7 Sobriety

¹³ While our point estimate is smaller, the confidence intervals on our estimate suggest it is comparable to the effect Kilmer et al (2013) estimate for the South Dakota 24/7 Sobriety Program. However, there are differences in the evaluated outcome and the program itself that may make direct comparisons inappropriate. Prior analyses have not examined non-alcohol substance use in South Dakota.

that reduce intoxicated driving should be better understood and adapted to each specific setting given that setting's norms and objectives. Some choices may diminish the program's efficacy, but the tradeoff may be worthwhile. As the Dakotas demonstrate, 24/7 Sobriety is not a monolith, and both individual and community-level outcomes may vary depending on the choices made by each jurisdiction.

24/7 Sobriety shifts monitoring and sanction from intoxicated driving, which is difficult to detect and risks harm to others in the community, to alcohol consumption which is detected reliably among participants who have demonstrated a higher risk of intoxicated driving with lower stakes than a DUI for the participant. While experienced alcohol-related crime offenders typically perceive the risk of subsequent similar crime relatively well (Apel 2013), the 24/7 Sobriety paradigm may be appealing for the repeat alcohol-related crime offender group given recent findings regarding the non-probabilistic nature of offender risk perception, particularly when probability of detection diminishes (Pogarsky et al. 2017).

Mothers Against Drunk Driving has argued that 24/7 Sobriety might have limited ability to reduce harms by its design because the program theoretically allows participants to drink immediately after testing and does not directly prevent that participant from operating a motor vehicle (Vock 2015). We acknowledge this as an important risk of the 24/7 Sobriety program as typically run, but further suggest that we confront this risk rigorously and scientifically. Evidence regarding the effectiveness of ignition interlock-type devices suggests they are effective deterrents against DUI when they are installed on vehicles. The bulk of the evidence suggests they do not reduce drinking or have a lasting deterrent effect against DUI (Willis et al. 2004; Government Accountability Office 2014; Voas et al. 2016), but the merits of available alcohol testing media are beyond the scope of this analysis. Given available technology, a straightforward trial can be run comparing 24/7 Sobriety using any alcohol testing medium that allows rapid detection—including an ignition interlock device—as the testing medium against an alternative prevention program of a jurisdiction's choice. Cost and wireless data infrastructure are the prime considerations for such a pilot—the basic design of 24/7 Sobriety suggests it can complement or strengthen an interlock, breathalyzer, or alcohol monitoring bracelet-based program.

Additional work is needed to fully understand, and potentially improve, the effects and mechanisms of such programs. Heterogeneous program effects across jurisdictions and participant groups may also comprise an important consideration as different sociodemographic characteristics or measured risk factors may be associated with recidivism in programs like 24/7. Further, jurisdictions may additionally tailor the program to account for decision-making biases to reduce crime (Pickett 2018). For example, does consistent high-frequency testing under 24/7 Sobriety appeal to the availability heuristic in a way that random testing under HOPE does not?

A randomized control trial varying assignment to 24/7 Sobriety and alcohol testing medium within those assigned to 24/7 Sobriety would allow one to tailor the program to improve its effectiveness across groups. To date, there is no evidence on whether the testing mechanism (e.g. CAM versus breathalyzer-based testing) has a differential effect within 24/7 Sobriety. The aforementioned selection mechanism, which may also play a role in the assigned testing mechanism, undermines the ability to estimate the impact of the testing technology choice on program effectiveness. Individual-level analyses that track both participants and a suitable control group and that address potential selection to the program would strengthen the evidence for this program's effectiveness. Further, looking beyond 2015 would allow one to examine the dose–response relationship by examining whether the longer participation spells required by HB1032 improve outcomes. Finally, given the

implied reductions in heavy drinking episodes, future research may also explore effects on less proximate outcomes such as morbidity and mortality.

Acknowledgements Funding was provided by National Institute of Justice, National Institute on Alcohol Abuse and Alcoholism (US) (Grant Nos. 2015-R2-CX-0016, R21AA022439). We thank presentation attendees at the American Society of Criminology, International Health Economics Association, and University of Pennsylvania Criminology Colloquium for valuable comments and suggestions. The manuscript was also improved by the feedback we received from the editors and two anonymous reviewers. We would also like to thank the Office of the North Dakota Attorney General and the alcohol testing companies (Alcohol Monitoring Systems and Intoximeters) for graciously providing data for this analysis. The views presented here represent only those of the authors.

References


- Apel RJ (2013) Sanctions, perceptions, and crime: implications for criminal deterrence. *J Quant Criminol* 29:67–101. <https://doi.org/10.1007/s10940-012-9170-1>
- Beccaria CM (1764) *On crimes and punishments*, 1st edn. Hackett Pub Co, Indianapolis
- Becker G (1968) Crime and punishment: an economic approach. *J Polit Econ* 76:169–217
- Bentham J (1789) *An introduction to the principles of morals and legislation*. Athlone Press, London
- Berk R, MacDonald J (2008) Overdispersion and poisson regression. *J Quant Criminol* 24:269–284. <https://doi.org/10.1007/s10940-008-9048-4>
- Bertrand M, Dufo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates?*. *Q J Econ* 119:249–275. <https://doi.org/10.1162/003355304772839588>
- Bester CA, Conley TG, Hansen CB (2011) Inference with dependent data using cluster covariance estimators. *J Econ* 165:137–151. <https://doi.org/10.1016/j.jeconom.2011.01.007>
- Bickel WK, Marsch LA (2001) Toward a behavioral economic understanding of drug dependence: delay discounting processes. *Addiction* 96:73–86. <https://doi.org/10.1046/j.1360-0443.2001.961736.x>
- Bickel WK, Jarmolowicz DP, Mueller ET, Gatchalian KM (2011) The behavioral economics and neuroeconomics of reinforcer pathologies: implications for etiology and treatment of addiction. *Curr Psychiatry Rep* 13:406–415. <https://doi.org/10.1007/s11920-011-0215-1>
- Carpenter C, Dobkin C (2010) Alcohol regulation and crime. In: Cook P, Ludwig J, McCrary J (eds) *Controlling crime: strategies and tradeoffs*. University of Chicago Press, pp 291–329
- Caulkins JP, Kilmer B, Kleiman MAR et al (2016) *Considering marijuana legalization*, 2nd edn. Oxford University Press, New York
- Chalfin A, McCrary J (2017) Criminal deterrence: a review of the literature. *J Econ Lit* 55:5–48. <https://doi.org/10.1257/jel.20141147>
- Comoreanu A (2017) Strictest and most lenient states On DUI. WalletHub. <https://wallethub.com/edu/dui-penalties-by-state/13549/#detailed-findings>
- Cook PJ (2016) Behavioral science critique of HOPE. *Criminol Public Policy* 15:1155–1161. <https://doi.org/10.1111/1745-9133.12256>
- Cullen FT, Pratt TC, Turanovic JJ (2016) It's Hopeless. *Criminol. Public Policy* 15:1215–1227. <https://doi.org/10.1111/1745-9133.12260>
- Dawson D, Reid K (1997) Fatigue, alcohol and performance impairment. *Nature* 388:235. <https://doi.org/10.1038/40775>
- Deschenes EP, Turner S, Petersilia J (1995) A dual experiment in intensive community supervision: Minnesota's prison diversion and enhanced supervised release programs. *Prison J* 75:330–356. <https://doi.org/10.1177/0032855595075003005>
- Federal Reserve Bank of Minneapolis (2019) *The Bakken Oil Boom*. <https://www.minneapolisfed.org/region-and-community/bakken>. Accessed 28 Apr 2020
- Government Accountability Office (2014) *Traffic safety: alcohol ignition interlocks are effective while installed; less is known about how to increase installation rates*. Government Accountability Office, Washington, DC
- Hansen B (2015) Punishment and deterrence: evidence from drunk driving. *Am Econ Rev* 105:1581–1617
- Hawken A (2016) All Implementation Is Local. *Criminol Public Policy* 15:1229–1239. <https://doi.org/10.1111/1745-9133.12264>

- Hawken A, Kleiman M (2009) Managing drug involved probationers with swift and certain sanctions: evaluating Hawaii's HOPE: executive summary. National Criminal Justice Reference Services, Washington, DC
- Hedlund J (2017) Drug impaired driving: a guide for states. Governors Highway Safety Association, Washington, DC
- Humphreys K, Kilmer B (2020) Still HOPEful: reconsidering a "failed" replication of swift, certain, and fair approach to reducing substance use among individuals under criminal justice supervision. *Addiction*. <https://doi.org/10.1111/add.15049>
- Kenkel DS (1993) Drinking, driving, and deterrence: the effectiveness and social costs of alternative policies. *J Law Econ* 36:877–913
- Kilmer B, Midgette G (2020) Deterring crime: insights from an individual-level analysis of 24/7 sobriety. *J Policy Anal Manag*
- Kilmer B, Nicosia N, Heaton P, Midgette G (2013) Efficacy of frequent monitoring with swift, certain, and modest sanctions for violations: insights from South Dakota's 24/7 Sobriety Project. *Am J Public Health* 103:e37–e43
- Kubas A, Kayabas P, Vachal K (2015) Assessment of the 24/7 Sobriety Program in North Dakota: participant behavior during enrollment. Upper Great Plains Transportation Institute, Fargo
- Kubas A, Kayabas P, Vachal K (2016) The effects of legislatively-mandated sobriety on first-time and repeat DUI offenders in North Dakota. Upper Great Plains Transportation Institute, Fargo
- Kubas A, Kayabas P, Vachal K (2017) Does the 24/7 Sobriety program positively influence driver behaviors in North Dakota?. Upper Great Plains Transportation Institute, Fargo
- Lattimore PK, MacKenzie DL, Zajac G et al (2016) Outcome findings from the HOPE demonstration field experiment. *Criminol Public Policy* 15:1103–1141. <https://doi.org/10.1111/1745-9133.12248>
- Lattimore PK, Dawes D, MacKenzie DL et al (2018) Evaluation of the honest opportunity probation with enforcement demonstration field experiment (HOPE DEF). RTI International, Research Triangle Park
- Lipari RN, Hughes A, Bose J (2016) Driving under the influence of alcohol and illicit drugs (The CBHSQ Report). Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration
- Long L (2009) The 24/7 sobriety project. *Pub Law* 17:2
- Loughran TA (2019) Behavioral criminology and public policy. *Criminol Public Policy* 18:737–758
- Loughran TA, Paternoster R, Weiss D (2012) hyperbolic time discounting, offender time preferences and deterrence. *J Quant Criminol* 28:607–628. <https://doi.org/10.1007/s10940-011-9163-5>
- Midgette G (2014) Monitoring with swift, certain, and moderate sanctions to reduce alcohol-related crime: the South Dakota 24/7 Sobriety Program. Doctoral dissertation, Pardee RAND Graduate School. <https://doi.org/10.7249/RGSD339>
- Nagin DS (2013) Deterrence in the twenty-first century. *Crime Just* 42:199–263. <https://doi.org/10.1086/670398>
- Nagin DS, Pogarsky G (2004) Time and punishment: delayed consequences and criminal behavior. *J Quant Criminol* 20:295–317. <https://doi.org/10.1007/s10940-004-5866-1>
- National Center for Statistics and Analysis (2019) State alcoholimpaired-driving estimates: 2017 data (Traffic Safety Facts). National Highway Traffic Safety Administration, Washington, DC
- Nicosia N, Kilmer B, Heaton P (2016) Can a criminal justice alcohol abstention programme with swift, certain, and modest sanctions (24/7 Sobriety) reduce population mortality? A retrospective observational study. *Lancet Psychiatry* 3:226–232
- Nochajski TH, Stasiewicz PR (2006) Relapse to driving under the influence (DUI): a review. *Relapse Addict Behav* 26:179–195. <https://doi.org/10.1016/j.cpr.2005.11.006>
- North Dakota Attorney General (2019) North Dakota Century Code t54c12. <https://www.legis.nd.gov/cencode/t54c12.pdf>. Accessed 24 Apr 2020
- Paternoster R, Saltzman LE, Waldo GP, Chiricos TG (1985) Assessments of risk and behavioral experience: an exploratory study of change. *Criminology* 23:417–436. <https://doi.org/10.1111/j.1745-9125.1985.tb00348.x>
- Petersilia J, Turner S (1990) Comparing intensive and regular supervision for high-risk probationers: early results from an experiment in California. *Crime Delinquency* 36:87–111. <https://doi.org/10.1177/001128790036001007>
- Petersilia J, Turner S (1993) Intensive probation and Parole. *Crime Just* 17:281–335
- Pickett JT (2018) Using behavioral economics to advance deterrence research and improve crime policy: some illustrative experiments. *Crime Delinq* 64:1636–1659. <https://doi.org/10.1177/001128718763136>
- Pogarsky G, Roche SP, Pickett JT (2017) Heuristics and biases, rational choice, and sanction perceptions. *Criminology* 55:85–111. <https://doi.org/10.1111/1745-9125.12129>

- Ridgeway G, Grogger J, Moyer RA, MacDonald JM (2019) Effect of gang injunctions on crime: a study of Los Angeles from 1988–2014. *J Quant Criminol* 35:517–541. <https://doi.org/10.1007/s10940-018-9396-7>
- Ross HL (1973) Law, science, and accidents: the British Road Safety Act of 1967. *J Leg Stud* 2:1–78
- Ruhm CJ (1996) Alcohol policies and highway vehicle fatalities. *J Health Econ* 15:435–454. [https://doi.org/10.1016/S0167-6296\(96\)00490-0](https://doi.org/10.1016/S0167-6296(96)00490-0)
- SCRAM Systems (2019) Transdermal alcohol testing. <https://www.scramsystems.com/products/alcohol-monitoring/transdermal-testing/>. Accessed 28 Apr 2020
- Sloan FA, Eldred LM, Xu Y (2014) The behavioral economics of drunk driving. *J Health Econ* 35:64–81
- South Dakota Attorney General (2019) 24/7 Program Statistics
- Vachal K, Benson L, Kubas A (2018) Effects of regular alcohol monitoring on North Dakota impaired drivers. Upper Great Plains Transportation Institute, Fargo
- Ver Hoef JM, Boveng PL (2007) Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology* 88:2766–2772
- Voas RB, Tippetts AS, Bergen G et al (2016) Mandating treatment based on interlock performance: evidence for effectiveness. *Alcohol Clin Exp Res* 40:1953–1960. <https://doi.org/10.1111/acer.13149>
- Vock DC (2015) The best way to prevent drunk driving? That's Debatable. *Governing.com*. https://www.governing.com/templates/gov_print_article?id=351031791. Accessed 28 Apr 2020
- Warren-Kigenyi N, Coleman H (2014) DWI Recidivism in the United States: an examination of state-level driver data and the effect of look-back periods on recidivism prevalence. *Traffic Saf Facts - Res Note*
- Weisburd D, Einat T, Kowalski M (2008) The miracle of the cells: an experimental study of interventions to increase payment of court-ordered financial obligations. *Criminol Public Policy* 7:9–36. <https://doi.org/10.1111/j.1745-9133.2008.00487.x>
- Willis C, Lybrand S, Bellamy N (2004) Alcohol ignition interlock programmes for reducing drink driving recidivism. *Cochrane Database Syst Rev* 4:CD004168
- Wooldridge JM (2010) *Econometric analysis of cross section and panel data*. MIT Press, Cambridge, MA
- Zaloshnja E, Miller TR, Blincoe LJ (2013) Costs of alcohol-involved crashes, United States, 2010. *Ann Adv Autom Med Annu Sci Conf* 57:3–12

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Greg Midgette^{1,2}  · Beau Kilmer³ · Nancy Nicosia⁴ · Paul Heaton⁵

¹ Department of Criminology and Criminal Justice, University of Maryland, College Park, MD, USA

² RAND Corporation, Arlington, VA, USA

³ RAND Corporation, San Francisco, CA, USA

⁴ RAND Corporation, Boston, MA, USA

⁵ University of Pennsylvania Law School, Philadelphia, PA, USA