

The Impact of Ambiguity-induced Error in Offender Decision-making: Evidence from the Field

Greg Midgette¹ , Thomas A. Loughran², and Sarah Tahamont¹

Journal of Research in Crime and Delinquency

2021, Vol. 58(6) 635–665

© The Author(s) 2021

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/00224278211000088

journals.sagepub.com/home/jrc



Abstract

Objectives: To invoke behavioral economics theories of ambiguity in the context of offender decision-making, and to test the impact of ambiguity in punishment certainty on offender decisions. *Methods:* We leverage a quasi-experimental condition among a sample of drunk driving arrestees that are tested for alcohol use and subject to mandatory brief incarceration for a violation. The treatment condition relaxes a zero-tolerance alcohol rule, thereby introducing design-based ambiguity surrounding the certainty of punishment. We use Mahalanobis matching and propensity score weighting methods to estimate the impact of ambiguity on violations. We then interrogate this finding with complementary sensitivity analyses.

¹ University of Maryland, College Park, MD, USA

² Pennsylvania State University, University Park, PA, USA

Corresponding Author:

Greg Midgette, University of Maryland, College Park, 2220 Samuel J. LeFrak Hall, 7251 Preinkert Drive, College Park, MD 20742, USA.

Email: gem@umd.edu

Results: When facing the ambiguity condition participants are 27–28 percentage points (84–93 percent) more likely to violate program conditions after 30 days of supervision. We demonstrate that a statistical difference in violations due to ambiguity is still detectable at 90 and 180 days of supervision. These results are robust to alternative specifications and falsification tests. **Conclusions:** This study is the first to examine the impact of ambiguity on criminal justice program compliance using a quasi-experiment from the field. We further demonstrate the unintended costs to persons under supervision and jurisdictions of laxity in program design, which are applicable across criminal justice domains.

Keywords

rational choice, offender decision-making, community corrections, substance abuse, deterrence

Drawing on both theoretical tradition (see, e.g., Beccaria 1764; Nagin 2013; Paternoster 2010) and empirical evidence (Durlauf and Nagin 2011; Kleiman 2009; Nagin, Solow, and Lum 2015) scholars have long advocated for policies designed to deter crime by increasing the certainty of detection rather than increasing the severity of punishment.¹ The empirical literature that demonstrates a link between perceived risk of detection and subsequent criminal behavior (Loughran et al. 2016; Matsueda, Kreager, and Huizinga 2006; Paternoster et al. 1985; Thomas, Loughran, and Hamilton 2020) spans multiple decades. This line of research is primarily rooted in traditional rational choice theory (Becker 1968), notably applied to explain crime displacement through a criminal opportunities framework (Cornish and Clarke 1986) and situational crime prevention (Clarke 1995).

Classical rational choice is predicated on the concept of expected utility theory (Rubinstein et al. 2007), which theorizes individual decisions to be the output of a rational calculus weighing the expected costs and benefits of potential actions. However, much of what we know about the influence of sanction certainty on offender decision-making comports with the behavioral science perspective that departures from rational behavior observed among individuals often manifest in predictable ways (Loughran 2019; Pogarsky, Roche, and Pickett 2018). In particular, scholars have demonstrated that misjudgments about probabilities, specifically in terms of how humans judge and act on them, are numerous (Camerer 1998; Kahneman and Tversky 1979). Cook (2016:1159) argued this point explicitly in

consideration of the mixed evidence generated by Hawaii's Opportunity Probation with Enforcement (HOPE) and subsequent Bureau of Justice Assistance-funded replications (Hawken & Kleiman 2009; Lattimore et al. 2016), which are programs built to deter recidivism in probation through certain sanctions: "a relatively high (but far from certain) punishment . . . will be the more effective deterrent when assessed by the traditional expected utility framing but that that conclusion is less obvious under the findings of modern behavioral science."

This study focuses on how justice-involved individuals process information about—and more importantly—*act on* perceived uncertainty through the behavioral economics concept of *ambiguity*, or the degree of confidence one holds regarding an unknown probability such as the chance of detection or punishment. In this study, we integrate theoretical insights from criminology and behavioral economics to develop and test novel predictions about the key role of ambiguity in offender decision-making. The study of ambiguity in perceptions of risk in criminal deterrence is relatively nascent, but its importance has been demonstrated in the offender decision-making process (Loughran et al. 2011; Pickett, Loughran, and Bushway 2016; Pogarsky et al. 2018). We contribute to the understanding of ambiguity and certainty in the study of offender decision-making by estimating an ambiguity parameter that to this point has only been described in theory and approximated empirically in contrived settings. Leveraging novel data from a "real-world" community supervision program, we test the impact of introducing ambiguity around the certainty of detection into a decision process in which an individual faces a real threat of incarceration for program violations.

Using data from two states, we directly consider the role of ambiguity in sanction certainty in the context of 24/7 Sobriety, a program that has demonstrated important reductions in drunk driving for community supervision participants (National Institute of Justice 2015). Specifically, we leverage a single difference in the implementation of 24/7 Sobriety that leads to consistent but ambiguous sanction certainty in one state (Montana) compared to another (South Dakota), thereby creating a credible counterfactual comparison. The 24/7 program in Montana mirrors South Dakota, except that they differ on one important dimension: the breathalyzer-determined blood-alcohol concentration (BAC) level that denotes a program violation is zero in South Dakota, and 0.02 mg/l in Montana. Central to our quasi-experimental design, these violation thresholds are explicitly provided to participants. However, the choice to drink (or not) is made with ambiguity surrounding the probability of violation in Montana

that does not exist in South Dakota. Evidence of drinking may not result in a violation in Montana, but it certainly will in South Dakota. These conditions provide the opportunity to make reasonable comparisons between a program that incorporates strict sanction certainty and one that, in an attempt to be more flexible, operates with more ambiguity, thus allowing us to test the theoretical prediction that individuals will be ambiguity-seeking in the presence of a near-certain loss.

Deterrence and Ambiguity

Deterrence is typically conceptualized as an information-based process of threat communication (Gibbs 1968; Zimring and Hawkins 1973). Geerken and Gove (1975) formalized the logic of deterrence as a social psychological theory which ushered in the importance of *perceived* beliefs about the costs and risks of punishment. The authors note (p. 503, emphasis added): “[un]like classical economists’ assumptions that men accurately perceive rewards and costs and then act, they expect *a range of accuracy* in prediction of actual rewards and costs.” Embedded in this explanation is the idea that individuals are prone to error and uncertainty in their subjective beliefs about the consequences of engaging in criminal behavior. Thus, for decades criminologists have been attentive to the study of deterrence through the perceptual properties of punishment (Paternoster et al. 1983, 1985; Nagin 2013; Waldo and Chiricos 1972).

Behavioral sciences and behavioral economics, too, have long considered the implications of uncertainty about subjective beliefs, drawing a distinction between risk and uncertainty surrounding this risk, dating back to Knight (1921). Risk is commonly conceptualized as a perceived probability of an event, and ambiguity is the uncertainty surrounding that assessment of probability (i.e., “second order probability”; Camerer and Weber 1992). For instance, consider the example of flipping a fair coin and observing heads. This likelihood is governed by a single parameter—the probability of heads, p . Assuming you believe our assertion that the coin is fair, then you can conclude that the first order parameter, $p = .5$. The subjective belief that the coin is truly fair, meaning you are certain $p = .5$, implies that ambiguity, the second order parameter, is zero; any lack of confidence about the coin’s probability of landing on heads in expectation induces ambiguity.

In an early experiment demonstrating a paradoxical and consistent violation of expected utility theory, Ellsberg (1961) found that individuals consistently prefer simple bets over those that are complicated by

ambiguity. In his “one-urn paradox,” individuals were told that an urn contained 90 balls, 30 of which were red, and the remaining 60 were a mix of black and yellow. Asked to choose between a bet to win a prize by drawing a red ball, or an alternative bet to draw a black ball, most chose the first bet. When asked to choose between a bet where they would win a prize by drawing a non-black ball, or an alternative to draw a non-red ball, most chose the second bet. These choices are inconsistent. The result from the first scenario implies that the largest fraction of balls is red, which also implies the belief that the largest fraction of balls is not black (i.e., red or yellow). This combination of preferences defies the transitivity assumption underlying rational choice through the demonstration of *ambiguity aversion*, the preference for clearly provided information (e.g., the proportion of red balls), even when the expected outcome is the same.

Machina and Siniscalchi (2014) review the substantial evidence regarding ambiguity aversion on decision-making across multiple domains including public policy, finance, and consumer choice. The implications for the role of ambiguity in punishment and deterrence, which by nature involve fuzzy beliefs about punishment risks, are also potentially numerous. For instance, Loughran et al. (2011) posit ambiguity aversion as a mechanism for Sherman’s (1990) idea of generating deterrence from rotating police crackdowns, whereby an otherwise fixed level of certainty of apprehension became variable. More recently, in a set of experiments contrasting expected utility and prospect theory Pickett and colleagues (2020) did not detect a change in attractiveness of a criminal opportunity due to the introduction of ambiguity in arrest risk.

An assumption, often implicit, underlying much of the theoretical rationale for ambiguity as a mechanism of deterrence, such as with rotating police crackdowns, is that individuals are always uniformly *ambiguity averse*. That is, less information about that risk yields more deterrent value for any level of perceived apprehension. However, as pointed out by others (Bleichrodt, Courbage, and Rey 2019; Kocher, Lahno, and Trautmann 2018) this idea is, in fact, inconsistent with Ellsberg’s (1961) original conceptualization. In particular, Kocher et al. (2018) note that in situations with highly probable losses, individuals should in fact be *ambiguity-seeking*.² For instance, Casey and Scholz (1991a, 1991b) tested the effect of ambiguity on tax compliance and evasion using a vignette study among student subjects in which the precision of the probability of detection varied across conditions. The experiments yielded evidence of boundary effects, where participants demonstrated ambiguity aversion and higher compliance as the probability of detection neared zero (the lower boundary) but tended to

become ambiguity-seeking and less compliant when the ambiguous estimate approached more certain detection (the upper boundary). This logic implies that in situations where the probability of detection is very close to 1 (i.e., individuals are facing a near-certain loss), *the injection of ambiguity may in fact be a catalyst*, not a deterrent.

24/7 Sobriety as a Field Test of Ambiguity

Perhaps the most difficult barrier to studying the effects of ambiguity in the decision-making process is properly measuring it. In their review of ambiguity in experimental economics, Krahnert, Ockenfels, and Wilde (2014: 8, emphasis added) describe the challenge of moving from the laboratory to field thusly: “A main challenge for theoretical and empirical studies is the issue of how ambiguity can be defined and how it can be captured. So far, it is *largely unclear how to operationalize the concept of ambiguity in the context of real economic applications.*” Since its introduction into the literature on offender decision-making, criminologists have also struggled with this exact problem (Loughran et al. 2011; Pickett et al. 2015; Pickett & Bushway 2015).³ Prior studies of offending decisions in the presence of ambiguity are limited in that ambiguity is studied either in an artificial lab setting, or in the context of financial decisions. While experimental evaluations of ambiguity in a lab or survey experimental setting allow for direct tests of theoretical constructs with high internal validity, the tradeoff is that those estimates are contrived and do not reflect “real world” decision-making. As such, scholars have stressed the importance of tests based on natural rather than contrived events (Baillon et al. 2018; Camerer 1998; Camerer and Weber 1992; Ellsberg 2011).

The implementation of the 24/7 Sobriety Program in South Dakota and Montana presents an opportunity to measure the impact of naturally occurring state-level variation in ambiguity on decisions across otherwise similar programs. We employ quasi-experimental evaluation methods to evaluate the importance of ambiguity in situ. An experiment randomly varying the definition of a violation that leads to incarceration within a single program contemporaneously would provide the strongest causal evidence but is ethically and legally tenuous. Another study doing so in a simulated environment would have strong internal validity, but potentially weaker external validity. Taken together, evidence from the field and evidence from lab settings are complementary, and both types of studies contribute to the evidence base on decision-making.

The 24/7 program is a particularly apt setting for the study of ambiguity. The original South Dakota 24/7 program was designed to directly reduce alcohol and other drug consumption through high-frequency alcohol and other drug testing with the threat of brief incarceration for failing a test or missing a test without an excuse.⁴ The 24/7 program is grounded in the notion of certainty of detection for violation of probation conditions approaching *unity*, or in the words of program advocates, “[t]he consistency and predictability of punishments make the consequences of bad behavior clear to the offender, reinforcing the need to make better decisions and change behavior” (Swift Certain Fair Resource Center, n.d.).

Beginning in 2010, Montana replicated 24/7 following a similar path from small pilot toward eventual statewide implementation. Montana’s program was intended to mirror South Dakota (Wickum 2017). Both states’ 24/7 programs combine high-frequency alcohol monitoring via breathalyzer or a Secure Continuous Remote Alcohol Monitor (SCRAM) bracelet, with the threat of certain and immediate but brief incarceration—typically 12 to 48 hours without monetary penalty or any record on criminal history—for participants that are found to drink alcohol or skip a test unexcused while assigned to the program. All participants are provided an orientation session where program rules and expectations are reviewed. In both states, a judge may assign individuals with any offense that is associated with alcohol misuse to participate in the program.

Our study focuses primarily on participants in both states that are monitored using twice-daily breathalyzer tests at a court-assigned central location in each county. Participants submit to a preliminary test via a breath alcohol test device that meets the National Highway Traffic Safety Administration evidentiary standard for false negatives and false positives. Confirmatory testing of potential violations further reduces the chance of false positives. In all jurisdictions, breathalyzer participants must travel to a testing location where the test is administered during a fixed window each morning and evening, though the window varies slightly from county-to-county.

A Theoretical Model of the Certainty Effect and Ambiguity

The realization of punishment in the criminal justice system is often multifarious, meted out over the experience of arrest, pretrial detention, and sentence. Each node adds complexity to the perception of deterrence.

Loughran (2019) advises that before investigating systematic departures from rationality predicted by behavioral economics, a necessary condition a priori must be a fully specified rational model (see Camerer and Loewenstein [2004] for further motivation of this point). Nagin (2013) describes how the certainty of punishment can be formally conceptualized:

[T]he certainty of punishment is conceptually and mathematically the product of a series of conditional probabilities—the probability of apprehension given commission of a crime, the probability of prosecution given apprehension, the probability of conviction given prosecution, and the probability of sanction given conviction

Specified this way, the multiple layers of punishment expectations, even if not formally considered by actors in this manner as unique parameters, can carry their own subjective expectation and inherent uncertainty. In other words, doubt about multiple dependent parameters can compound into increased variance around the product, thereby increasing ambiguity around probabilistic expectations of punishment. Consequently, ambiguity surrounding a punishment probability can manifest from multiple sources downstream, which may or may not even reflect how an individual will respond to uncertainty at any given stage.

The 24/7 programs in South Dakota and Montana present a unique opportunity to study ambiguity in that, by design of the intervention, ambiguity can be isolated to a single parameter in the chain of events Nagin describes. The program design of 24/7 has several features which allow us to simplify one's expectation about punishment likelihood. First, given the nature of the testing, the probability of apprehension approaches unity (i.e., participants are tested twice per day, approximately every twelve hours). In both states, all failures result in punishment. That is, there is no uncertainty regarding punishment conditional on failing a test.⁵ The probabilities associated with the steps between apprehension and conviction are effectively invariant at 1. These properties make the dimensions of this ostensibly complex probability of receiving punishment reducible to a single parameter, specifically, the probability of failure given testing.

In South Dakota, participants are informed that any detected level of alcohol results in a violation, thus conditional on drinking, the probability of failure equals 1 with certainty (i.e., no ambiguity). Conversely in Montana, where there is a more lenient violation policy, participants are informed that the violation threshold is 0.02 mg/l, so very close to zero but not zero. While this conveys information about the risk of detection, it also induces

ambiguity about whether or not drinking will result in a failed test and subsequent punishment. In other words, when individuals decide to drink in Montana, the perception of punishment risk is more ambiguous than in South Dakota. This “noise” around the failure likelihood is potential variance in subjective belief about risk, which captures ambiguity as described by Camerer and Weber (1992). This variability between states is the condition which our empirical analysis aims to leverage.

There are several key features of this design that are attractive both empirically and theoretically. First, in any traditional setting, it would be difficult to isolate the (risk) certainty effect from the additional effect of ambiguity, as the two factors are easily conflated, especially with regard to the closeness to the boundary, which would simultaneously affect both. Hence, this major structural challenge of disentangling the risk certainty effect and ambiguity effect is one that our design can uniquely parameterize.

Second, the sharp delineations of testing and punishment certainty put forth to 24/7 participants circumvents other key complications inherent to most perceptual deterrence studies including the questions about the credibility of elicited subjective beliefs, both in general (Dominitz and Manski 1997; Hurd 2009) and specifically related to deterrence (Loughran, Paternoster, and Thomas 2014), as well as key disagreements in the literature regarding the relationship between objective and subjective sanctions (Apel 2013; Kleck et al. 2005; Pogarsky et al. 2018). More specifically, program participants experience a structure that sets most key first-order parameters to be certain or near-certain, including probability of detection and probability of sanction given detection. Participants are explicitly provided information on the other key parameter—the BAC level determining violation—and we observe their behavioral response to the threat of real punitive repercussions, rather than elicited perceptions.

Hypothesis

We define the hypothesis of interest to explain the potential impact of ambiguity when participants face certain detection in both states, and there is ambiguity in punishment in Montana but not South Dakota:

Hypothesis: If participants are ambiguity-seeking, ambiguity positively influences their decision to drink beyond the threshold limit, $P(\text{failure})_{\text{MT}} > P(\text{failure})_{\text{SD}}$.

This hypothesis implies that participants make riskier choices in the presence of ambiguity when facing near-certain detection and punishment, which is an idea informed by insights from the behavioral economics literature. In a variety of contexts, individuals prefer uncertainty when facing losses (Kocher et al. 2018; Machina and Siniscalchi 2014). Further, evidence suggests ambiguity-seeking behavior and lower compliance when perceived likelihood is near the upper boundary when facing likely loss (Casey and Scholz 1991b; Kocher et al. 2018). An empirical finding against this hypothesis would provide evidence to support the practical intent of the program designers in Montana, who sought to minimize violations by preventing intoxication rather than imposing abstention. If participants are ambiguity averse or neutral, their propensity to drink beyond the threshold limit will be reduced or will be unaffected, and we would expect violation rates in Montana to be equal to or less than those in South Dakota, that is $P(\text{failure})_{\text{MT}} \leq P(\text{failure})_{\text{SD}}$.

Data and Methods

Our analysis relies on the similarity of individuals assigned to the 24/7 programs and in the 24/7 programs themselves in all ways except the violation threshold. We combine data provided by the Attorney General's Offices in Montana and South Dakota from the states' statewide 24/7 program data management systems and criminal records databases with field research including visits to 22 county programs during and subsequent to the study period. The states' administrative data provide detailed information on a set of persons with criminal records involving alcohol misuse who have a high risk of recidivism. We observed that the programs are equivalent in design aspects including rules, policies, and funding from site visits during the study period including observations of the testing facilities and program in operation from a sample of county-level program administrators from small, medium, and large population jurisdictions in each state, as well as semi-structured interviews with state program administrators from their respective Offices of the Attorney General (Midgette 2014; Midgette & Kilmer 2021), cross-referenced with depictions in prior literature (Kubas, Kayabas, and Vachal 2017; Stevens 2016). Practical aspects including variation in physical settings of the program, participant characteristics, and implementation fidelity were also comparable, which is consistent with the published history of the two programs (Mabry N.d.; Wickum 2017) and is evident in the similarity of the laws authorizing the program in each state

(Montana 24/7 Sobriety Program Act 2011; South Dakota 24/7 Sobriety Program Act 2007).

This analysis focuses on breathalyzer test information within two states' 24/7 programs that are uniquely capable of informing our understanding of ambiguity in attempts to deter individuals from choosing to engage in prohibited behavior. While there may be other features of the 24/7 program that introduce some uncertainty into the decision-making process for participants, the major design *difference* between the programs in Montana and South Dakota is the blood alcohol threshold that determines a violation. We mitigate several important potential confounding influences by focusing on the probability of a first violation. In subsequent violations, the measurement of deterrent effect is potentially contaminated because participants may update their expectations based on numerous factors including the experience of a violation after drinking and being tested, as well as the punishment itself.

The analytic sample includes all persons that were convicted of a DUI-2 (i.e., a second offense of driving under the influence of alcohol) in South Dakota and Montana and are subsequently assigned to 24/7 for the first time between January 2010 and August 2014 via twice-per-day breathalyzer testing in either state. We focus on DUI-2 offenses because both states include participation in the 24/7 Sobriety program as remediation for DUI-2 by statute and it is the modal category of arrest leading to program assignment in both states. The laws defining DUI-2 are also consistent across states. In both states, DUI-2 is defined as a misdemeanor crime resulting in up to one year in jail, fines of up to \$2,000, and one year of license suspension with the possibility of a restricted license to allow travel to work, substance treatment, and other defined necessary purposes. Both states use a 0.08 blood-alcohol concentration to define the per se threshold of intoxication, both use a 10-year look back period to define a second offense and both define their DUI statutes to include alcohol in combination with other drug use, or use of any drug (including alcohol).

We observe every breathalyzer result for individuals enrolled in 24/7 through year-end 2015. In total, 1.23 million breathalyzer results were recorded for 4,682 participants. Our analysis considers 3,814 participants in both states for whom we observe criminal record information.⁶ Since our goal is to make credible comparisons between participants in 24/7 in Montana and South Dakota, we believe the criminal record information is key to making "apples to apples" comparisons. However, our findings are robust to the inclusion of individuals for whom we do not observe criminal record information. The outcome of interest in our main analysis

is the probability of violation in the first 30, 90, and 180 days of program participation.

Our analysis must confront two important considerations, one practical and one theoretical. Practically, to the extent that alcohol intoxication affects cognitive function, intoxication might confound our empirical results. In this case, cognitive impairment is very unlikely to be a factor in the marginal decision to drink tested here. Given both states test participants approximately every twelve hours, a potential failure can occur in two ways: either a participant can consume any amount of alcohol in the few hours immediately before their test, or the participant can consume multiple beverages any time during the intra-test period. In both cases, the decision to consume alcohol beyond the statutory limit in each state is equivalent.

Theoretically, a participant who consumes a single drink is unlikely to differentiate a 0.02 BAC from zero, and a participant who consumes multiple drinks is likely to fail a test in either state. As a consequence, the part of the decision calculus that we are able to isolate is the distinction between an individual *knowing with certainty* that evidence of a single drink will lead to a failed breathalyzer test (South Dakota) and will definitely be punished as compared to a decision in which the individual perceives that there is *some chance* that by the time the individual tests the BAC from a drink might have dissipated below the 0.02 threshold (Montana) thereby avoiding a violation and the associated punishment.

Analytic Plan

To establish credible counterfactuals (Rosenbaum and Rubin 1985) in a real community supervision setting, we employ two alternative methods to test the robustness of our findings to modeling choices, Mahalanobis distance matching (MDM) and doubly-robust inverse propensity score weighting (IPW). Each method requires a set of observable characteristics with which to establish balance between the treated group and a comparison group drawn from the untreated sample. We evaluated a broad set of theoretically informed candidate measures that may be related to program assignment or violations (Apel and Sweeten 2010): participant demographics associated with differential offending risk (gender and age), criminal history record (the count of prior arrests, separate indicators of prior arrests for violent crime, weapons charges, or drug charges), and community characteristics capturing sociodemographic and socioeconomic variation (a binary metropolitan area indicator, a binary for counties with Native American population centers, poverty rate, median household income, and

unemployment rate), as well as and alcohol availability as measured by the density of on and off-premises alcohol retailers in each participant's county. The inclusion of uninformative or uncorrelated controls can both lead to relatively poor balance and excessive variability in propensity scores (Brookhart et al. 2006; Wyss et al. 2013) and can compound omitted variable bias (Pearl 2000). Additionally, some measures such as unemployment rate, median household income, and poverty rate are highly collinear. To mitigate the risk of assignment model-induced bias, we exclude the Native American population center indicator, median household income, and on-premises retail alcohol outlet density from our final model specification.

We then estimate probability of violation under ambiguity in Montana as compared to the certainty of violation (and subsequent sanction) in South Dakota with the MDM and IPW models. For each time duration, we estimate the average treatment effect on treated (ATT) of ambiguity on our outcome of interest, the probability of violation in the first 30, 90, or 180 days of program participation. In the matching model, treated Montana cases are matched to their nearest neighbor via Mahalanobis-distance one-to-one with replacement to untreated cases in South Dakota.

The MDM matching approach is an alternative estimation strategy to the propensity score-based method that is more commonplace in criminological studies. Pearl (2009) described the potential for bias in point estimates induced by propensity score matching. King and Neilsen (2019) demonstrate conditions under which this concern, as well as a loss of efficiency in standard errors, are empirically evident. The MDM matching estimator also allows for a sensitivity analysis to evaluate how much hidden bias can be present before the qualitative conclusions of the study begin to change (Rosenbaum 2002, 2005). We also implement analytic standard errors proposed by Abadie and Imbens (2006) that are consistent under a range of underlying data distributions. We further present results that leverage a doubly-robust inverse propensity score-weighted model relying on a logit selection model as an alternative estimation strategy to show that our results are consistent across alternative estimators.⁷ The doubly-robust IPW estimator is unbiased if either the propensity score selection model or the model estimating impact is correctly specified (Funk et al. 2011; Huber 1973; Kang and Schafer 2007).

Finally, while there is no direct test of the unconfoundedness assumption required for causal inference from both MDM and IPW, we present two sensitivity analyses to test the robustness of our findings. We first quantify the sensitivity of observed treatment effect to hidden bias by calculating a Γ statistic advised by Rosenbaum (2002). Γ is a crucial element of a selection

on observables analysis because it demonstrates how much hidden bias would have to be present in the omitted selection mechanisms to negate the estimated relationship. In our case, the Γ statistic shows how much bias from unmeasured factors it would take to negate the relationship we observe between the ambiguity condition and the probability of a violation.

We also evaluate program violations for a separate sample of 24/7 participants who are monitored via ankle-worn secure remote continuous alcohol monitoring (SCRAM) devices, rather than the breathalyzer devices that are at the core of this study. Differences in the way that the device detects the presence of alcohol transdermally require a different procedure for determining a program violation. Detection of a violation for the SCRAM version of the program is administered by SCRAM systems in a single central location, and the process does not vary across states—participants receive the same information, face the same threat of sanction, and pay similarly higher daily fees than in-person breathalyzer participants. Thus, the remote alcohol monitor-based program should be virtually identical across states. Central to the validity of our main findings, the differential ambiguity condition does not exist. So, if violation rates are equivalent among SCRAM participants in the two states, it bolsters our claim that the settings are comparable across states and that we have isolated the effect of ambiguity on probability of violations using breathalyzer participants.

Results

Comparison between Montana and South Dakota 24/7 Participants

Fifty-nine percent of Montanans violated by failing an alcohol test in their first 30 days under supervision, compared to 30 percent in South Dakota. Among those who fail, average time to failure is much shorter in the presence of ambiguity, 40 days in Montana versus 109 days in South Dakota. Table 1 displays the unweighted pretreatment characteristics of our analytic sample of DUI-2 24/7 participants in Montana and South Dakota to the characteristics in South Dakota after weighting. Of the nine control variables included in the selection model, balance exists in the unadjusted data among four variables based on the 20 percent standardized bias threshold, and none are beyond the 50 percent threshold Cohen (2013) defined as a “moderate” difference. Roughly 72 percent of observations are male and the average age is 33.5 years. Among the sample, participants in Montana were more likely to have been arrested for a violent crime, drug crime, and

Table 1. Unweighted Sample Characteristics and Standardized Percentage Bias after Matching and Weighting.

Matching Variables	Montana (BAC ≥ 0.02)		South Dakota (BAC > 0)				
	Raw	Raw	Percentage Bias	MDM	Percentage Bias	IPW	Percentage Bias
<i>Individual Characteristics</i>							
Male	0.719	0.733	-3.1	0.726	-1.6	0.708	2.4
Age	33.55	34.88	-11.5	33.30	2.2	33.38	14.5
<i>Prior Arrest History</i>							
Count of Prior Arrests	2.888	3.462	-22.3	2.978	-3.5	2.961	-3.1
Violent Crime Prior	0.319	0.227	20.7	0.317	0.5	0.329	-1.5
Drug Crime Prior	0.290	0.265	5.6	0.290	0	0.296	-1.2
Weapons Crime Prior	0.043	0.024	10.2	0.043	0	0.356	3.7
<i>County Characteristics</i>							
Metropolitan Area	0.362	0.515	-31.2	0.364	-0.5	0.343	3.9
Poverty Rate (percent)	14.83	13.07	45.4	14.33	12.9	16.04	-19.3
Liquor Store Density	0.779	1.025	-41.8	0.780	-0.2	0.712	13.0

*p < 0.05, **p < 0.01, ***p < 0.001

Table 2. MDM and IPW ATT Estimates of Violation Rates by Ambiguity Condition.

Threshold	Model	n	Violation Rate		Ambiguity Impact (ATT)	95 percent Confidence Interval
			≥.02 BAC	>0 BAC		
30 days	MDM		0.588	0.317	0.271	[0.195, 0.347]
	IPW	3,814	0.588	0.304	0.284	[0.233, 0.336]
90 days	MDM		0.588	0.556	0.257	[0.206, 0.308]
	IPW	3,368	0.861	0.582	0.279	[0.238, 0.321]
180 days	MDM		0.861	0.751	0.214	[0.169, 0.259]
	IPW	3,006	0.965	0.784	0.181	[0.154, 0.208]

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; MDM estimated with analytical standard errors based on Abadie and Imbens (2006); IPW estimated via bootstrap to account for propensity weights (Lunceford and Davidian 2004).

weapons crime, though only the difference in violent crime is statistically significant in the unweighted sample. Montana participants also reside in areas with marginally greater alcohol retailer density and higher poverty rates. These potential risk factors suggest that participants in Montana may be on average at higher risk of misconduct, including program violations. Both models were balanced over all model covariates and all cases in Montana—no standardized effect size for any covariate exceeded |20 percent| in the MDM or IPW model.

Impact of Ambiguity in the Alcohol Threshold on the Violation Rate

Table 2 presents the output of the MDM and IPW models predicting violation as a function of the ambiguity introduced by a non-zero violation threshold in Montana. The matching and weighting models produce nearly identical results. The MDM estimates the average causal effect of ambiguity on violations among those facing the higher ambiguity condition in Montana to be a 27.1 percentage point increase, 95 percent CI [0.195, 0.347] in the probability of violation, which is a 46 percent reduction in the compliance rate. These estimates do not differ statistically from those produced by the IPW model, for which the point estimate is marginally larger (ATT = 0.284, 95 percent CI [0.233, 0.336]). The higher probability of violation in Montana remains significant through 180 days in both models, though magnitude declines to 26–28 percentage points at 90 days and 18–21 percentage points at 180 days. This decrease may be evidence of

experiential learning (e.g., through observation of other participants over time) or evidence of differential risk between states among those who are assigned to the program for longer periods. Since the mean and median time on the program in South Dakota is longer than in Montana, those assigned for more than 180 days in Montana may be a higher risk group than those assigned to the program for the same length of time in South Dakota. Nevertheless, impact on violations remains substantively large.

We find that greater ambiguity overwhelms the small difference in defined failure BAC threshold, consistent with our alternative hypothesis. This suggests that ambiguity is a salient feature of the decision to engage in prohibited behavior, and its effect on violations is large and meaningful. We estimate the net effect of ambiguity introduced by the seemingly small policy choice to slightly increase the BAC threshold to roughly double the probability of a violation after 30 days on the program and to increase the probability of a violation by approximately 20 percent at the 180-day mark among the higher-risk pool of long-term participants.

Sensitivity to Unmeasured Confounders

The most important assumption of causal interpretation of estimates generated by MDM and IPW is that, after conditioning on observed confounders, no unmeasured confounder exists (Angrist and Pischke 2009; Rosenbaum 2002). In this specific context, unobserved differences between the 24/7-eligible populations in Montana and South Dakota, which have similar but not identical demographic characteristics, laws, and criminal justice processes may confound our estimates if they are correlated with differential selection into the program in either state. We further constrain our analysis to examine only the probability of a *first* violation among first-time 24/7 participants since the deterrent effect of the program could be differentially affected by the experience of punishment after a violation in the two states. To assess the sensitivity of our findings to this type of hidden bias in program assignment, we perform the sensitivity analysis prescribed by Rosenbaum (2002, 2005) using the Mantel-Haenszel Γ test (Becker and Caliendo 2007). Importantly, benchmarking our estimates using Γ allows us to assess the sensitivity of our estimates to hidden bias.

Based on the 30-day threshold MSM estimator, at $\Gamma = 2.27$ our estimate is statistically significant at $p < .05$. A value of Γ near 1 means a study is potentially very sensitive to hidden bias. Our obtained test statistic implies that some missing confounding factor would need to be at least 2.27 times

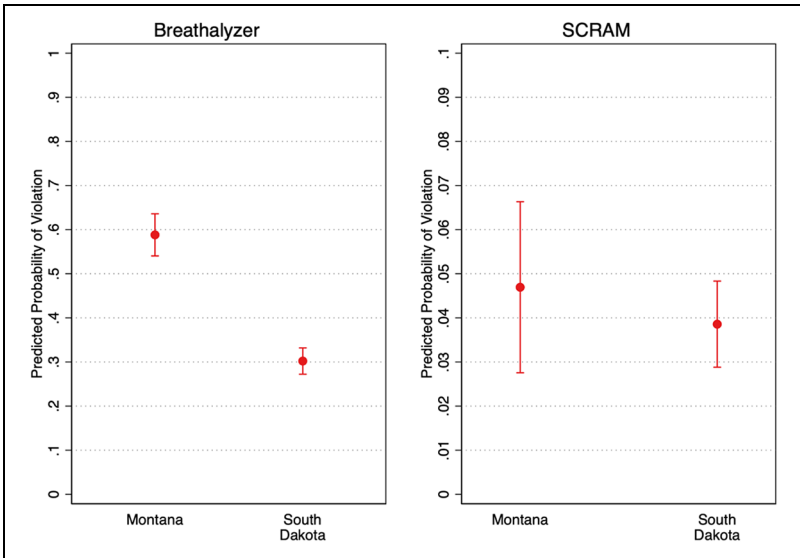


Figure 1. Comparison of 30-day violation rates by state and alcohol supervision technology.

more common in Montana to explain the higher rate of violations in the state. In light of the documentation of the two contexts from qualitative assessments based on field observations, interviews, and document reviews (Midgette 2014; Midgette & Kilmer 2021), in our view it is unlikely that there exists a factor that differentiates selection into Montana's program that can be simultaneously so influential and hidden.

We further examined the sensitivity of our results using SCRAM-based method of alcohol supervision to interrogate the key assumption that program and environmental conditions are equivalent across states. If there were important differences in the program or other risk factors associated with drinking that led to higher risk in Montana than South Dakota, we would expect to see a difference in both breathalyzer and SCRAM violation rates. Using the same Mahalanobis matching method, among DUI-2 24/7 participants on SCRAM we find that the 30-day violation rate among 463 SCRAM participants in Montana is statistically indistinguishable ($p = 0.443$) from the 1,537 participants in South Dakota. This is consistent with the assumption that the states and programs are equivalent but for the ambiguity condition (see Figure 1).⁸

Discussion and Conclusion

The study of individuals' responsiveness to sanctions is a question that is central to both criminal justice policy and criminological theory. Multiple studies using credible research designs show that increasing punishment certainty through strategies like greater police presence materially reduces crime (Di Tella and Schargrotsky 2004; Klick and Tabarrok 2005; J. MacDonald, Fagan, and Geller 2016), even though the relationship between policing and crime is endogenous (Kubrin et al. 2010). That said, the relationship between perceived risk and subsequent offending is routinely demonstrated to be negative, but often quite weak, pushing scholars to embrace insights from behavioral economics to better understand the behavioral mechanisms underlying deterrence and offending decisions more broadly (Loughran 2019; Pogarsky et al. 2018; Thomas, Hamilton, and Loughran 2018).

In this study, we use unique data from the field to estimate the impact of ambiguity on violation of an alcohol sobriety order in community corrections, a direct measure of delinquency in the field. We find support for our hypothesis. Decisions are more prone to error when ambiguity in the likelihood of punishment increases. Ambiguity introduced by changing the failure threshold from zero to .02 BAC leads participants to violate at higher rates under an ostensibly more forgiving condition.

While our findings about ambiguity are novel in criminology, they are consistent with prior research in behavioral economics. In the context of near-certain detection and punishment, ambiguity induces risk seeking for long shots, consistent with Dimmock et al. (2016). Further, when facing near-certain losses, individuals might actually become ambiguity-seeking (Casey and Scholz 1991b; Kocher et al. 2018). Our findings are also consistent with Ellsberg (1961) in that decision heuristics will be informed heavily by clear signals among complex information. Both states clearly define a violation to 24/7 participants; in South Dakota drinking is nominally prohibited by a zero threshold, but in Montana drinking is nominally allowed in small amounts by a 0.02 BAC threshold. Participants are demonstrably worse at distinguishing the threshold for a small amount of alcohol than they are at distinguishing zero tolerance.

The logic that program design features can be low-cost policy levers which might be able to influence compliance is not new and, in fact, is gaining evidence. Many of these design strategies are based on the concept of "nudging" (Thaler and Sunstein 2009) whereby "choice architecture" is employed to motivate key behavioral changes toward compliance. For instance, Fishbane, Ouss, and Shah (2020) observed a large beneficial effect

of nudges intended to remind defendants to appear in court for certain low-level offenses in New York City. Pickett (2018) provides evidence on how providing additional information can reduce intention to drive drunk in a hypothetical experiment.

Our results demonstrate that the potential consequences of ambiguity should be considered in the policy and program design process. Whereas ambiguity may have little influence on decision-making in a context like drunk driving where the probability of detection is very low, our results show that ambiguity tends to increase noncompliance when the probability of detection is very high. Together these studies are complementary in demonstrating that the effect of ambiguity on decision-making is context dependent, and specifically depends on the underlying first order probabilities. In policing and supervision activities where the detection probability is naturally low, we would predict that ambiguity-sensitive individuals would become ambiguity-averse, in which case an additional deterrent effect could be extracted with similar levels of resources through efficient delivery of a stimulus (Kleiman and Kilmer 2009; Sherman 1990). On the other hand, when detection probability is high, ambiguity may have an opposite effect.

Ambiguity is often present in policies, programs, and practices affecting persons under criminal justice supervision and can be a product of a design choice. In the specific case of 24/7, the implementation choice that introduces ambiguity may appear to be in the mutual interest of participants and the jurisdictions operating the program. Participants may be perceived to struggle to maintain strict sobriety expected under a 0 BAC threshold, whereas increasing the threshold to $BAC \geq .02$ allows for some flexibility while still maintaining a low risk of drunk driving or other detrimental consequences of alcohol misuse. The source of ambiguity we interrogate is a consequence of a common practical decision made in community corrections when a gradient of misconduct is feasible: How much detectable prohibited behavior should be allowed before a punitive response is warranted?

Our results show that policy choices that may be intended to introduce flexibility may, under certain conditions, induce more of the prohibited behaviors. Given the prior expectation of lower failure rates under the more lenient .02 threshold, jurisdictions may expect to minimize the number of expensive sanctions delivered (in this case, short stays in jail). In reality, our results, which isolate the contribution of perceived ambiguity, tell a much different story; nearly twice as many violations occur due to the ambiguity that the program rules introduce. This may appear paradoxical to policy and

program architects. On its face, the choice to relax a rule ostensibly allows more flexibility and attempts to differentiate a high-risk behavior (in this case, drinking in excess) from a less risky behavior (drinking at a level that falls short of risking intoxication and its potential consequences). In cases where violations carry expensive consequences for participants and jurisdictions, as is the case in 24/7 Sobriety, drug and treatment courts, and throughout community supervision, this study demonstrates that ambiguity built into systems designed to monitor and deter misconduct may exacerbate unwanted behaviors. Yet, there may be ways to maintain flexibility without the detrimental effects of ambiguity. A simple experiment providing low-cost portable breathalyzers to participants may minimize violations without removing the flexibility Montana's program architects intended. Knowing another drink will push BAC above 0.02 may yield the same decision that knowing a drink will push BAC above zero.

The evidence we present suggests, first, that the basic design of 24/7 may be an effective strategy to deter risky behaviors in community corrections, more so when ambiguity is minimized. This sheds light on the potential theoretical mechanisms at play in Swift, Certain, Fair (SCF) programs such as 24/7 and HOPE. Evaluations of SCF-type programs in multiple states have demonstrated promising crime and substance use reduction effects on participants (e.g., Hamilton et al. 2016; Hawken and Kleiman 2009; Taxman et al. 2003). However, recent research has called into question the universal effectiveness of strategies based on the SCF paradigm (Lattimore et al. 2016; O'Connell, Brent, and Visser 2016), as well as the utility of this certainty-based supervision approach and leading some to argue for the outright abandonment of these types of interventions (Clear and Frost 2014; Cullen, Pratt, and Turanovic 2016). The potential sources of ambiguity in SCF programs that fundamentally rely on certainty to affect behavior are numerous, from minor implementation choices that soften hard and fast rules, to the clarity with which instructions and expectations are conveyed, to core program components such as drug testing. In such programs, ambiguity may result in adverse outcomes.

To test the effect of ambiguity in a field setting, we use methods which rely on the assumption that no latent confounder exists that would substantially change our conclusions. For this reason, we make causal claims with caution and rigorously examine the estimated effect through sensitivity analyses. Generalizations beyond the sampling frame of DUI-2 arrestees in Montana and South Dakota should be done with care. There are few differences in policies between the states. However, Montana law considers DUI-3 to be a misdemeanor, while in South Dakota it is a felony. If the

deterrent effect of a harsher sanction for a subsequent DUI in South Dakota also reduces drinking while on 24/7, our results may be biased. However, prior research has shown that individuals that are arrested for DUI-2 and are assigned to 24/7 experience similar reductions in the probability of re-arrest for DUI-3 in both states (Kilmer and Midgette 2020; Midgette and Kilmer 2015), which suggests the punitiveness of the potential criminal sanction for DUI-3 is not related to the efficacy of 24/7 programs. For this reason, we do not believe the difference in DUI-3 laws affects our conclusions.

Our findings in combination with prior analyses of the impact of 24/7 on DUI (Kilmer et al. 2013, Midgette et al. 2020) are consistent with the dual-process theory of behavior differentiating intuitive “automatic” from reasoning “deliberative” self-control mechanisms (Kahneman 2003). The theoretical framework for 24/7 is founded on the notion that the choice to drink occurs before the choice to drink *then drive*—detering the former controls the risk of the latter. If the decision to drink and the decision to drink *then drive* were made using the same cognitive process, we would expect the large impact of ambiguity on the decision to drink would translate to a similarly large difference in the decision to drink then drive. The collective findings on the effects of 24/7 instead show that the choice to drink, which is made regularly, is greatly influenced by ambiguity, but the less frequent and multifarious decision to drink then drive is not.

The consequence of the ambiguity induced by a seemingly small design choice appears to be massive. By reducing ambiguity, South Dakota was able to achieve reductions in DUI recidivism and impose less punishment. Punishment in the form of incarceration generates potentially large tangible and intangible costs to sanctioned participants, and large monetary costs to jurisdictions. Based on the Vera Institute of Justice Price of Jails Survey (Henrichson, Rinaldi, and Delaney 2015), we estimate the cost to jurisdictions of a day in jail to be, on average, \$186 and for every dollar in incarceration costs, ten dollars in social cost are incurred (McLaughlin et al. 2016). The practical consequences of our findings may be startling for criminal justice policy and practice.

Through the lens of rational choice, the decision to drink and risk jail time in 24/7 happens when the near-certain and immediate expected value of that consumption exceeds the less certain and delayed negative consequence, to which participation in 24/7 adds the threat of jail time. For these individuals, the addition of ambiguity about the likelihood of failure leads to risk seeking with costly consequences that do not appear to be internalized by the decision-maker. While we might assume that the zero-tolerance approach of South Dakota represents a more punitive regime, our results

reveal a different and consequential outcome that recasts the theoretical understanding of punishment risk.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by National Institute of Justice (grant ID: 2015-R2-CX-0016).

ORCID iD

Greg Midgette  <https://orcid.org/0000-0003-2038-2950>

Notes

1. A conclusion across both the economic and criminological literatures is the lack of measurable deterrent effect of more severe sanctions (Chalfin and McCrary 2017; Durlauf and Nagin 2011; Nagin, Cullen, and Jonson 2009; Villettaz, Killias, and Zoder 2006).
2. As an illustration consider the apprehension you might feel about checking a credit card statement after a particularly regrettable spending binge, versus how you feel when checking your banking account when expecting your IRS refund to be electronically deposited. In the former you might choose to avoid seeing the statement at all, versus the latter in which you check your account repeatedly until it shows up.
3. For instance, while Loughran et al. (2011) are credited for first formalizing the concept of ambiguity in criminological discourse, they nonetheless proposed a quixotically coarse measurement of the construct from retrospective stated perceptions, illustrating the difficulty of capturing this parameter in a non-lab setting. Pickett et al. (2014) report results from an experiment in which they measure perceived ambiguity among respondents in two different ways—asking respondents to report a range for a probability versus having them report a level of “sureness.” They found that while the two measures operate differently in some ways, both elicitation methods produced values that were related to overall levels of risk, suggesting the inherent difficulty of differentiating ambiguity and risk perceptions, and, more importantly, that the values were correlated across seemingly disparate questions, suggesting there might be an individual-specific component to how individuals perceive uncertainty and ambiguity.

4. Unlike drug and DUI courts, 24/7 Sobriety does not require participants to attend treatment or work toward other pro-social goals, nor does the program bar such activities; it only mandates desistance from substance use (Long 2009). Participants in 24/7 are ordered to either 1) show up at the county testing facility, typically the county jail, every morning and every evening and blow into a breathalyzer, 2) wear an alcohol monitoring bracelet that tests their sweat every 30 minutes for alcohol, and/or 3) submit to urinalysis or drug patch testing for illegal drugs. Prior research suggests community-level reductions in repeat-DUI and domestic violence arrests (Kilmer et al. 2013).
5. The devices are maintained and calibrated according to the manufacturer's prescribed schedule.
6. We obtained charge-level criminal history records for the entire sample in South Dakota, and for 65 percent ($n = 777$) of the available sample in Montana due to resource constraints. The order of records matched was alphabetical, thus near-random, so should avoid potential issues of statistical selection. We observe statistical balance over the observable characteristics between observations with missing and complete criminal history records. In Montana, administrative records for 24/7 participation were linked to criminal records data by statistical analysts in the Montana Attorney General's Office.
7. Matching estimators are potentially valid for causal inference for those observations that have common support (i.e., successfully matched) and do not alter the relative importance of any matched observation. Weighting-based estimators do not rely on common support, but instead reweight the counterfactual group, when calculating ATT, to achieve balance over the observable characteristics in the data. The intuition underlying the IPW model is to minimize the difference over our covariates between the treated and counterfactual groups to generate a propensity score weighted composite of South Dakota participants that is comparable to Montana participants (Hirano and Imbens 2001; Wooldridge 2010).
8. We also note that violation rates among SCRAM participants are considerably lower than breathalyzer participants due to a combination of the testing technology and potential differences in the sample of 24/7 participants that are assigned to SCRAM versus breathalyzer.

References

- Abadie, A. and G. W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74(1):235-67. doi:10.1111/j.1468-0262.2006.00655.x.
- Angrist, J. and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

- Apel, R. J. 2013. "Sanctions, Perceptions, and Crime: Implications for Criminal Deterrence." *Journal of Quantitative Criminology* 29(1):67-101. doi:10.1007/s10940-012-9170-1.
- Apel, R. J. and G. Sweeten. 2010. "Propensity Score Matching in Criminology and Criminal Justice." Pp. 543-62 in *Handbook of Quantitative Criminology*, edited by A. R. Piquero and D. Weisburd. New York: Springer. doi:10.1007/978-0-387-77650-7_26.
- Baillon, A., Z. Huang, A. Selim, and P. P. Wakker. 2018. "Measuring Ambiguity Attitudes for All (Natural) Events." *Econometrica* 86(5):1839-58. doi:10.3982/ECTA14370.
- Beccaria, C. marchese di. 1764. *On Crimes and Punishments*. 1st ed. Hackett. Retrieved February 27, 2021. (<https://oll.libertyfund.org/title/voltaire-an-essay-on-crimes-and-punishments>).
- Becker, G. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76(2):169-217.
- Becker, S. O. and M. Caliendo. 2007. "Sensitivity Analysis for Average Treatment Effects." *Stata Journal* 7(1):71-83.
- Bleichrodt, H., C. Courbage, and B. Rey. 2019. "The Value of a Statistical Life under Changes in Ambiguity." *Journal of Risk and Uncertainty* 58:1-15.
- Brookhart, M. A., S. Schneeweiss, K. J. Rothman, R. J. Glynn, J. Avorn, and T. Stürmer. 2006. "Variable Selection for Propensity Score Models." *American Journal of Epidemiology* 163(12):1149-56.
- Camerer, C. F. 1998. "Prospect Theory in the Wild: Evidence from the Field." Social Science Working Paper No. 1037, California Institute of Technology, Pasadena, CA.
- Camerer, C. F. and M. Weber. 1992. "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity." *Journal of Risk and Uncertainty* 5(4): 325-70. doi:10.1007/BF00122575.
- Casey, J. T. and J. T. Scholz. 1991a. "Beyond Deterrence: Behavioral Decision Theory and Tax Compliance." *Law & Society Review* 25(4):821-43. doi:10.2307/3053871.
- Casey, J. T. and J. T. Scholz. 1991b. "Boundary Effects of Vague Risk Information on Taxpayer Decisions." *Organizational Behavior and Human Decision Processes* 50(2):360-94. doi:10.1016/0749-5978(91)90027-Q.
- Chalfin, A. and J. McCrary. 2017. "Criminal Deterrence: A Review of the Literature." *Journal of Economic Literature* 55(1):5-48. doi:10.1257/jel.20141147.
- Clarke, R. V. 1995. "Situational Crime Prevention." *Crime and Justice* 19:91-150. doi:10.1086/449230.
- Clear, T. R. and N. A. Frost. 2014. *The Punishment Imperative: The Rise and Failure of Mass Incarceration in America*. New York: New York University Press.
- Cohen, J. 2013. *Statistical Power Analysis for the Behavioral Sciences*. Milton Park, England: Routledge.
- Cook, P. J. 2016. "Behavioral Science Critique of HOPE." *Criminology & Public Policy* 15(4):1155-61. doi:10.1111/1745-9133.12256.

- Cornish, D. and R. V. Clarke. 1986. *The Reasoning Criminal: Rational Choice Perspectives on Offending*. The Hague, the Netherlands: Springer-Verlag.
- Cullen, F. T., T. C. Pratt, and J. J. Turanovic. 2016. "It's Hopeless." *Criminology & Public Policy* 15(4):1215-27. doi:10.1111/1745-9133.12260.
- Dimmock, S. G., R. Kouwenberg, O. S. Mitchell, and K. Peijnenburg. 2016. "Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence." *Journal of Financial Economics* 119(3):559-77. doi:10.1016/j.jfineco.2016.01.003.
- Di Tella, R. and E. Schargrodsky. 2004. "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces after a Terrorist Attack." *American Economic Review* 94(1):115-33. doi:10.1257/000282804322970733.
- Dominitz, J. and C. F. Manski. 1997. "Using Expectations Data to Study Subjective Income Expectations." *Journal of the American Statistical Association* 92(439): 855-67. doi:10.1080/01621459.1997.10474041.
- Durlauf, S. N. and D. S. Nagin. 2011. "Imprisonment and Crime." *Criminology & Public Policy* 10(1):13-54. doi:10.1111/j.1745-9133.2010.00680.x.
- Ellsberg, D. 1961. "Risk, Ambiguity, and the Savage Axioms." *The Quarterly Journal of Economics* 75(4):643-69. doi:10.2307/1884324.
- Ellsberg, D. 2011. "Notes on the Origins of the Ellsberg Urns (Introduction to the Symposium Issue)." *Economic Theory* 48:221-27.
- Fishbane, A., A. Ouss, and A. K. Shah. 2020. "Behavioral Nudges Reduce Failure to Appear for Court." *Science* 370(6517):eabb6591.
- Funk, M. J., D. Westreich, C. Wiesen, T. Stürmer, M. A. Brookhart, and M. Davidian. 2011. "Doubly Robust Estimation of Causal Effects." *American Journal of Epidemiology* 173(7):761-67. doi:10.1093/aje/kwq439.
- Geerken, M. R. and W. R. Gove. 1975. "Deterrence: Some Theoretical Considerations." *Law & Society Review* 9(3):497-513. doi:10.2307/3053169.
- Gibbs, J. 1968. "Crime, Punishment, and Deterrence." *Southwestern Social Science Quarterly* 48(4):515-31.
- Hamilton, Z., C. M. Campbell, J. van Wormer, A. Kigerl, and B. Posey. 2016. "Impact of Swift and Certain Sanctions." *Criminology & Public Policy* 15(4): 1009-72. doi:10.1111/1745-9133.12245.
- Hawken, A. and M. Kleiman. 2009. *Managing Drug Involved Probationers with Swift and Certain Sanctions: Evaluating Hawaii's HOPE: Executive Summary*. Washington, DC: National Criminal Justice Reference Services.
- Henrichson, C., J. Rinaldi, and R. Delaney. 2015. *The Price of Jails: Measuring the Taxpayer Cost of Local Incarceration*. Washington, DC: Vera Institute of Justice.
- Hirano, K. and G. W. Imbens. 2001. "Estimation of Causal Effects Using Propensity Score Weighting: An Application to Data on Right Heart Catheterization." *Health Services and Outcomes Research Methodology* 2(3):259-78. doi:10.1023/A:1020371312283.

- Huber, P. J. 1973. "Robust Regression: Asymptotics, Conjectures and Monte Carlo." *Annals of Statistics* 1(5):799-821. doi:10.1214/aos/1176342503.
- Hurd, M. D. 2009. "Subjective Probabilities in Household Surveys." *Annual Review of Economics* 1(1):543-62. doi:10.1146/annurev.economics.050708.142955.
- Kahneman, D. 2003. "A Perspective on Judgment and Choice: Mapping Bounded Rationality." *American Psychologist* 58(9):697-720. doi:10.1037/0003-066X.58.9.697.
- Kahneman, D. and A. Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2):263-91. doi:10.2307/1914185.
- Kang, J. D. Y. and J. L. Schafer. 2007. "Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data." *Statistical Science* 22(4):523-39. doi:10.1214/07-STS227.
- Kilmer, B. and G. Midgette. 2020. "Deterring Crime: Insights from an Individual-level Analysis of 24/7 Sobriety." *Journal of Policy Analysis and Management* 39(3):801-34.
- Kilmer, B., N. Nicosia, P. Heaton, and G. Midgette. 2013. "Efficacy of Frequent Monitoring with Swift, Certain, and Modest Sanctions for Violations: Insights from South Dakota's 24/7 Sobriety Project." *American Journal of Public Health* 103(1):e37-43.
- King, G. and R. Nielsen. 2019. "Why Propensity Scores Should Not Be Used for Matching." *Political Analysis* 27(4):435-54. doi:10.1017/pan.2019.11.
- Kleck, G., B. Sever, S. Li, and M. Gertz. 2005. "The Missing Link in General Deterrence Research*." *Criminology* 43(3):623-60. doi:10.1111/j.0011-1348.2005.00019.x.
- Kleiman, M. A. 2009. *When Brute Force Fails: How to Have Less Crime and Less Punishment*. Princeton, NJ: Princeton University Press.
- Kleiman, M. and B. Kilmer. 2009. "The Dynamics of Deterrence." *Proceedings of the National Academy of Sciences* 106(34):14230. doi:10.1073/pnas.0905513106.
- Klick, J. and A. Tabarrok. 2005. "Using Terror Alert Levels to Estimate the Effect of Police on Crime." *Journal of Law and Economics* 48(1):267-79.
- Knight, F. H. 1921. *Risk, Uncertainty, and Profit*. Boston, MA: Houghton-Mifflin Company.
- Kocher, M. G., A. M. Lahno, and S. T. Trautmann. 2018. "Ambiguity Aversion Is Not Universal." *European Economic Review* 101:268-83. doi:10.1016/j.euroecorev.2017.09.016.
- Krahnen, J. P., P. Ockenfels, and C. Wilde. 2014. *Measuring Ambiguity Aversion: A Systematic Experimental Approach*. Berlin, German: Verein für Socialpolitik/ German Economic Association.
- Kubas, A., P. Kayabas, and K. Vachal. 2017. *Does the 24/7 Sobriety Program Positively Influence Driver Behaviors in North Dakota?* Upper Great Plains Transportation Institute. Retrieved February 27, 2021. (<https://www.ugpti.org/resources/reports/downloads/dp-296.pdf>).

- Kubrin, C. E., S. F. Messner, G. Deane, K. McGeever, and T. D. Stucky. 2010. "Proactive Policing and Robbery Rates across U.S. Cities*." *Criminology* 48(1): 57-97. doi:10.1111/j.1745-9125.2010.00180.x.
- Lattimore, P. K., D. L. MacKenzie, G. Zajac, D. Dawes, E. Arsenault, and S. Tueller. 2016. "Outcome Findings from the HOPE Demonstration Field Experiment." *Criminology & Public Policy* 15(4):1103-41. doi:10.1111/1745-9133.12248.
- Long, L. 2009. "The 24/7 Sobriety Project." *Public Laws* 17:2.
- Loughran, T. A. 2019. "Behavioral Criminology and Public Policy." *Criminology & Public Policy* 18(4):737-58.
- Loughran, T. A., R. Paternoster, A. Chalfin, and T. Wilson. 2016. "Can Rational Choice Be Considered a General Theory of Crime? Evidence from Individual-level Panel Data." *Criminology* 54(1):86-112. doi:10.1111/1745-9125.12097.
- Loughran, T. A., R. Paternoster, A. R. Piquero, and G. Pogarsky. 2011. "On Ambiguity in Perceptions of Risk: Implications for Criminal Decision Making and Deterrence." *Criminology: An Interdisciplinary Journal* 49(4):1029-61. doi:10.1111/j.1745-9125.2011.00251.x.
- Loughran, T. A., R. Paternoster, and K. J. Thomas. 2014. "Incentivizing Responses to Self-report Questions in Perceptual Deterrence Studies: An Investigation of the Validity of Deterrence Theory Using Bayesian Truth Serum." *Journal of Quantitative Criminology* 30(4):677-707. doi:10.1007/s10940-014-9219-4.
- Lunceford, J. K. and M. Davidian. 2004. "Stratification and Weighting via the Propensity Score in Estimation of Causal Treatment Effects: A Comparative Study." *Statistics in Medicine* 23(19):2937-60. doi:10.1002/sim.1903.
- Mabry, A. N.d. *24/7 Sobriety Program*. Pierre: South Dakota Office of the Attorney General. Retrieved February 27, 2021. (<https://atg.sd.gov/docs/247ppt.pdf>).
- MacDonald, J., J. Fagan, and A. Geller. 2016. "The Effects of Local Police Surges on Crime and Arrests in New York City." *PLoS One* 11(6):e0157223.
- Machina, M. J. and M. Siniscalchi. 2014. "Chapter 13—Ambiguity and Ambiguity Aversion." Pp. 729-807 in *Handbook of the Economics of Risk and Uncertainty*, edited by M. Machina and K. Viscusi. Vol. 1. North-Holland, the Netherlands. doi:10.1016/B978-0-444-53685-3.00013-1.
- Matsueda, R. L., D. A. Kreager, and D. Huizinga. 2006. "Deterring Delinquents: A Rational Choice Model of Theft and Violence." *American Sociological Review* 71(1):95-122. doi:10.1177/000312240607100105.
- McLaughlin, M., C. Pettus-Davis, D. Brown, C. Veeh, and T. Renn. 2016. *The Economic Burden of Incarceration in the U.S.* Working Paper No. CI072016, Concordance Institute for Advancing Social Justice, St. Louis, MO.
- Midgette, G. 2014. Monitoring with swift, certain, and moderate sanctions to reduce alcohol-related crime: The South Dakota 24/7 Sobriety Program. Santa Monica

- (CA): RAND RGSD-339. Retrieved February 27, 2021. (https://www.rand.org/pubs/rgs_dissertations/RGSD339.html).
- Midgette, G. and B. Kilmer. 2021. *Can novel "swift-certain-fair" programs work outside of pioneering jurisdictions? An analysis of 24/7 Sobriety in Montana*. Retrieved February 27, 2021. (https://www.rand.org/pubs/working_papers/WR1083.html).
- Montana 24/7 Sobriety Program Act, H.B. 0106, 2011 Montana Legislature (2011) Retrieved February 27, 2021. (<https://leg.mt.gov/bills/2011/billhtml/HB0106.htm>).
- Nagin, D. S. 2013. "Deterrence in the Twenty-first Century." *Crime and Justice* 42(1):199-263. doi:10.1086/670398.
- Nagin, D. S., F. T. Cullen, and C. L. Jonson. 2009. Imprisonment and Reoffending. Pp. 115-200 in *Crime and Justice: A Review of Research*. Vol. 38, edited by Michael Tonry. Chicago, IL: University of Chicago Press.
- Nagin, D. S., R. M. Solow, and C. Lum. 2015. "Deterrence, Criminal Opportunities, and Police." *Criminology: An Interdisciplinary Journal* 53(1):74-100. doi:10.1111/1745-9125.12057.
- National Institute of Justice. 2015. *CrimeSolutions Program Profile: South Dakota's 24/7 Sobriety Project*. Retrieved February 27, 2021. (<https://crimesolutions.ojp.gov/programdetails?id=404#eb>).
- O'Connell, D. J., J. J. Brent, and C. A. Visser. 2016. "Decide Your Time." *Criminology & Public Policy* 15(4):1073-102. doi:10.1111/1745-9133.12246.
- Paternoster, R. 2010. "How Much Do We Really Know about Criminal Deterrence Centennial Symposium: A Century of Criminal Justice - Crimes and Punishments." *Journal of Criminal Law & Criminology* 100(3):765-824.
- Paternoster, R., L. E. Saltzman, G. P. Waldo, and T. G. Chiricos. 1983. "Estimating Perceptual Stability and Deterrent Effects: The Role of Perceived Legal Punishment in the Inhibition of Criminal Involvement." *The Journal of Criminal Law and Criminology* 74(1):270-97. doi:10.2307/1143322.
- Paternoster, R., L. E. Saltzman, G. P. Waldo, and T. G. Chiricos. 1985. "Assessments of Risk and Behavioral Experience: An Exploratory Study of Change." *Criminology* 23(3):417-36. doi:10.1111/j.1745-9125.1985.tb00348.x.
- Pearl, J. 2000. *Causality: Models, Reasoning and Inference*. Vol. 9. Cambridge, MA: Cambridge University Press, pp. 10-11.
- Pearl, J. 2009. "Letter to the Editor." *Statistics in Medicine* 28:1415-24.
- Pickett, J. T. and S. D. Bushway. 2015. "Dispositional Sources of Sanction Perceptions: Emotionality, Cognitive Style, Intolerance of Ambiguity, and Self-Efficacy." *Law and Human Behavior* 39(6):624.
- Pickett, J. T. 2018. "Using Behavioral Economics to Advance Deterrence Research and Improve Crime Policy: Some Illustrative Experiments." *Crime & Delinquency* 64(12):1636-59. doi:10.1177/0011128718763136.

- Pickett, J. T., J. C. Barnes, T. Wilson, and S. Patrick Roche. 2020. "Prospect Theory and Criminal Choice: Experiments Testing Framing, Reference Dependence, and Decision Weights." *Justice Quarterly* 37(6):1140-68.
- Pickett, J. T., T. A. Loughran, and S. Bushway. 2015. "On the Measurement and Properties of Ambiguity in Probabilistic Expectations." *Sociological Methods & Research* 44(4):636-76.
- Pickett, J. T., T. A. Loughran, and S. Bushway. 2016. "Consequences of Legal Risk Communication for Sanction Perception Updating and White-collar Criminality." *Journal of Experimental Criminology* 12(1):75-104.
- Pogarsky, G., S. P. Roche, and J. T. Pickett. 2018. "Offender Decision-making in Criminology: Contributions from Behavioral Economics." *Annual Review of Criminology* 1(1):379-400. doi:10.1146/annurev-criminol-032317-092036.
- Rosenbaum, P. R. 2002. "Sensitivity to Hidden Bias." Pp. 105-70 in *Observational studies*. New York: Springer.
- Rosenbaum, P. R. 2005. "Sensitivity Analysis in Observational Studies." Pp. 1809-14 in *Encyclopedia of Statistics in Behavioral Science*, edited by Brian S. Everitt and David C. Howell. Hoboken, NJ: Wiley.
- Rosenbaum, P. R. and D. B. Rubin. 1985. "Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score." *The American Statistician* 39(1):33-38. doi:10.1080/00031305.1985.10479383.
- Rubinstein, A., O. Morgenstern, H. W. Kuhn, and J. von Neumann. 2007. *Theory of Games and Economic Behavior: 60th Anniversary Commemorative Edition*. United Kingdom: Princeton University Press.
- Sherman, L. W. 1990. "Police Crackdowns: Initial and Residual Deterrence." *Crime and Justice* 12:1-48.
- South Dakota 24/7 Sobriety Program Act, H.B. 1072, 82nd Session. 2007. Retrieved February 27, 2021. (<https://sdlegislature.gov/#/Session/Bill/21236/112990>).
- Stevens, J. C. 2016. "Is Montana's '24/7 Sobriety Program' Detering Drunk Drivers?" University of Montana Undergraduate Theses and Professional Papers. Retrieved February 27, 2021. (<https://scholarworks.umt.edu/cgi/viewcontent.cgi?article=1084&context=utpp>).
- Swift Certain Fair Resource Center. n.d. *Programs-swift Certain Fair Resource Center*. Retrieved February 27, 2021. (<https://scfcenter.org/programs/>).
- Taxman, F. S., D. Reedy, K. Moline, M. Ormand, and C. Yancey. 2003. *Strategies for the Drug-involved Offender*. College Park, MD: University of Maryland Center for Applied Policy Studies, Bureau of Government Research.
- Thaler, R. H. and C. R. Sunstein. 2009. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New York, NY: Penguin Books.

- Thomas, K. J., B. C. Hamilton, and T. A. Loughran. 2018. "Testing the Transitivity of Reported Risk Perceptions: Evidence of Coherent Arbitrariness." *Criminology* 56(1):59-86. doi:10.1111/1745-9125.12154.
- Thomas, K. J., T. A. Loughran, and B. C. Hamilton. 2020. "Perceived Arrest Risk, Psychic Rewards, and Offense Specialization: A Partial Test of Rational Choice Theory." *Criminology* 58(3):485-509. doi:10.1111/1745-9125.12243.
- Villettaz, P., M. Killias, and I. Zoder. 2006. "The Effects of Custodial vs. Non-custodial Sentences on Re-offending: A Systematic Review of the State of Knowledge." *Campbell Systematic Reviews* 2(1):1-69. doi:10.4073/csr.2006.13.
- Waldo, G. P. and T. G. Chiricos. 1972. "Perceived Penal Sanction and Self-reported Criminality: A Neglected Approach to Deterrence Research." *Social Problems* 19(4):522-40. doi:10.1525/sp.1972.19.4.03a00090.
- Wickum, L. 2017. *24/7 Sobriety Program: The Montana Story*. Montana Highway Patrol. Retrieved February 21, 2021. (<https://www.mtleague.org/wp-content/uploads/2017/10/24-7SobrietyPresentation.pdf>).
- Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT press.
- Wyss, R., C. J. Girman, R. J. LoCasale, M. Alan Brookhart, and T. Stürmer. 2013. "Variable Selection for Propensity Score Models when Estimating Treatment Effects on Multiple Outcomes: A Simulation Study." *Pharmacoepidemiology and Drug Safety* 22(1):77-85.
- Zimring, F. E. and G. J. Hawkins. 1973. *Deterrence: The Legal Threat in Crime Control*. Chicago: University of Chicago Press.

Author Biographies

Greg Midgette is an assistant professor in the Department of Criminology and Criminal Justice at the University of Maryland. His research interests include drug policy, the role of program and policy design in decision-making, and the impacts of criminal justice policies on public health.

Thomas A. Loughran is a professor in the Department of Sociology and Criminology and the School of Public Policy at Pennsylvania State University. His research interests include responses to sanctions and decision-making, developmental and life course criminology, and methods for analyzing longitudinal data.

Sarah Tahamont is an assistant professor in the Department of Criminology and Criminal Justice at the University of Maryland. Broadly, her research interests include corrections, causal inference, longitudinal patterns of criminal justice contact, and methodological advances in Criminology and Criminal Justice.