Experimental Economics (2023) 26:304–338
https://doi.org/10.1007/s10683-022-09755-w

ORIGINAL PAPER

Ambiguity and enforcement

Evan M. Calford[1] · **Gregory DeAngelo**[2]

Received: 22 April 2021 / Revised: 18 March 2022 / Accepted: 22 March 2022 /
Published online: 20 April 2022
© The Author(s) 2022

Abstract
Law enforcement officials face numerous decisions regarding their enforcement choices. One important decision, that is often controversial, is the amount of knowledge that law enforcement distributes to the community regarding their policing strategies. Assuming the goal is to minimize criminal activity (alternatively, maximize citation rates), our theoretical analysis suggests that agencies should reveal (shroud) their resource allocation if criminals are uncertainty seeking, and shroud (reveal) their allocation if criminals are uncertainty averse. We run a laboratory experiment to test our theoretical framework, and find that enforcement behavior is approximately optimal given the observed non-expected utility uncertainty preferences of criminals.

Keywords Law enforcement · Strategic uncertainty · Experimental economics · Ambiguity aversion

JEL Classification K42 · C72 · D81 · D82

1 Introduction

In the 18th century Jeremy Bentham proposed the *panopticon* as a mechanism for using monitoring uncertainty to influence the behavior of the monitored (Bentham, 1843). Bentham's panopticon was a circular penitentiary where a single centrally located observer could watch the inmates in each cell without themselves being observed by the inmates. Inmates would be unaware of when or if the observer was watching their cell and, Bentham argued, would therefore behave as if they were

✉ Evan M. Calford
Evan.Calford@anu.edu.au

Gregory DeAngelo
gregory.deangelo@cgu.edu

[1] Research School of Economics, Australian National University, Canberra, Australia

[2] Department of Economic Sciences, Claremont Graduate University, Claremont, USA

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

always being watched. Evaluating Bentham's hypothesis through the lens of modern utility theory, the success of the panopticon may depend on the uncertainty preferences of the monitored: an inmate who is sufficiently uncertainty averse might indeed behave as Bentham suggests, but an uncertainty seeking inmate might be relatively undeterred by an unknown probability of being monitored at any given instant.[1] If, instead, inmates were informed of the (truthful) instantaneous probability of being monitored the uncertainty seeking inmate may exhibit more compliant behavior because the revelation of the monitoring probability ameliorates some of the inmate's uncertainty.

Bentham's panopticon illustrates the central question of this paper: when should a monitor reveal, or shroud, their monitoring probabilities? The decision to reveal or shroud monitoring probabilities depends not only on the uncertainty attitudes of the monitored, but is also affected by the incentives and motivations of the monitor. The monitor may seek to minimize non-compliant behavior, or they may seek to maximize the number of citations given for non-compliant behavior.[2]

The applications of strategic manipulation of monitoring uncertainty are varied, although we focus on law enforcement as our leading example.[3] Law enforcement officers weigh their primary responsibility of "enforcing the laws that are enacted by elected officials ...and ...interpreted by the courts," with other important duties including crime prevention (USDOJ, n.d.). Common policing strategies may support one goal yet hinder others. For example, hidden speed traps or unmarked vehicles have the potential to detect individuals that are behaving recklessly without providing forewarning about the presence of enforcement, while the use of signposted speed traps and marked patrol cars might allow the citizen to adjust their behavior to evade the law or even discourage individuals from breaking the law at all. Additionally, law enforcement activities could be driven by financial motives, especially if law enforcement revenues comprise a significant share of the financial inflows to a community (Goldstein et al. (2020); Kantor et al. (2017)).

To explore the relationship between the use of unmarked police vehicles and law enforcement revenues, we present Fig. 1 which displays a negative relationship

[1] Farago et al. (2008) find experimental evidence that criminals are more risk-taking than either students or entrepreneurs. Uncertainty aversion has been noted in Harel and Segal (1999), Lochner (2007) and DeAngelo and Charness (2012). However, Lochner (2007) notes that young males that are engaged in criminal behavior are responsive to law enforcement behavior in a manner that is consistent with expected utility theory. Finally, Mungan and Klick (2014) show theoretically that forfeiture of illegal gains generate behavior that is consistent with risk-aversion amongst would-be criminals.

[2] The latter incentive can be induced when the monitor has a citation quota or when the monitor shares in fine revenues, for example.

[3] The decision about whether to announce enforcement behavior is not unique to policing. Indeed, the IRS decides whether to provide information on auditing probabilities via the announcement of targeted sections of the tax code, while weighing the conflicting goals of incentivizing tax compliance and revenue recovery (Andreoni et al. (1998) provide an overview of the economics of tax compliance). Several other examples of announced versus surprise monitoring environments exists. For example, a manager who supervises production at multiple plants, the Environmental Protection Agency air and water monitors, or the health department can choose whether to announce their planned inspection rate at each location, or not.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

between the percentage of unmarked cars and the dependence of a location on revenues from law enforcement fines and fees.[4]

Figure 1 is consistent with the logic underlying Bentham's panopticon. When the fine-to-tax ratio is low, police officers are focused on crime prevention: the use of undercover vehicles increases uncertainty about the level of traffic monitoring which causes uncertainty averse citizens to behave more cautiously. In contrast, when the fine-to-tax ratio is high, police officers have an incentive to write more citations: the use of marked vehicles reduces uncertainty about the level of traffic monitoring which may lead to more traffic violations and increased revenue. However, drawing causal inferences from this data is extremely challenging. In each jurisdiction, the decision of whether to reveal or shroud monitoring activities will depend on the cultural, historical, legal and institutional environment in which the decision makers are embedded.[5] Nevertheless, the data represented in Fig. 1 is, to the best of our knowledge, the "best" observational data available on this topic.

It is precisely because of the lack of data that allows for clean identification regarding the revelation of monitoring activities that we turn to a theoretical and experimental study of the question. We abstract away from the confounding factors mentioned above and use a unifying framework that can be broadly applied.[6] Indeed, there is a considerable body of similar work related to compliance and enforcement including Cason et al. (2021); Slemrod (2019) that we contribute to in this research.[7] Notably, Salmon and Shniderman (2019) study the use of ambiguous punishment in a framed tax compliance experiment and found only limited evidence that subjects exhibited risk aversion. We build a theoretical model to identify the equilibrium effects of the revelation or shrouding of monitoring probabilities, and then test the equilibrium predictions in a laboratory experiment. Importantly our model does not assume expected utility, and our experimental results include evidence of non-neutral ambiguity preferences.

For concreteness, we present our model within the context of police monitoring and enforcement of speeding violations. We implement our model by positing that a driver has a choice of two roads, and that law enforcement may allocate their

[4] The percentages of marked and unmarked police vehicles are sourced from the 2013 Law Enforcement Management and Administrative Statistics database, while revenues from taxes and law enforcement fines and fees were obtained from the 2013 US Census data on local and government finances about government revenues. In total, 1606 cities are included in the data. The size of the bubbles in Fig. 1 correspond to the number of observations.

[5] For example, police monitoring varies across both geographic (signposted speed cameras in New South Wales, Australia and hidden speed cameras in Victoria, Australia) and temporal boundaries (hidden speed cameras were banned in Arizona in 2010).

[6] We focus on law enforcement monitoring as our leading example throughout the paper. However, in our experimental implementation we used a supervisor/worker framing to prevent subjects' existing preferences for breaking or abiding by laws from influencing our results. With only minor modifications, our model could easily be adapted to study the revelation of monitoring in other important and interesting environments, including auditing.

[7] Cason et al. (2021) provides an overview with respect to compliance with environmental regulations, and Slemrod (2019) provides a broader empirical overview of tax compliance and enforcement. See also the literature review in Sect. 2 for further details.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

monitoring resources across both roads. The more resources placed on a given road, the greater the probability that a driver who speeds on that road will be caught and punished. The following describes the stages of the experiment:

Stage 1: Enforcement Officer chooses monitoring probabilities m_A and m_B for the first and second road, and chooses whether to reveal m_A and m_B or only reveal $m = m_A + m_B$.
Stage 2 : The Driver decides to either speed on the first road (choice A), speed on the second road (choice B) or not speed (choice C).
Stage 3 : Officer and Driver payoffs are obtained.

There are two payoff schemes that are implemented for the drivers, using a between treatment design, labeled the *Prob* and *EV* treatments. If a driver chooses A, then they either (*Prob*) earn 100 points with probability $0.9 - m_A$ and earn 0 with probability $0.1 + m_A$, or (*EV*) earn $90 - 100m_A$. Similarly, B returns either (*Prob*) 100 points with probability $0.9 - m_B$ and nothing otherwise, or (*EV*) $90 - 100m_B$ points. If the driver chooses C they receive either (*Prob*) 100 points with probability 0.5 and nothing otherwise, or (*EV*) 50 points. In each case, the *EV* treatment pays the expected value of the *Prob* treatment lottery.

The officer pays a constant marginal cost of monitoring and can decide whether to reveal the monitoring probabilities (m_A and m_B) to the driver, or only reveal the total ($m_A + m_B$) monitoring probability.[8] We study two different payoff structures for the enforcement officer: one where the officer strictly prefers the driver to not speed and choose C (i.e. the officer wishes to minimize the number of speeding violations), and one where the officer's payoff is increasing in the monitoring probability on the road chosen by the driver (i.e. the officer earns revenue from catching a speeding driver).

This design allows us to address our central research questions. (1) How does monitoring uncertainty affect behavior? (2) Does an enforcer manipulate monitoring uncertainty to induce favorable behavior among the monitored? (3) Is there a distinction between enforcers with a revenue incentive and enforcers with a prevention incentive?

The experimental design also allows the underlying components of uncertainty preference, risk preference and ambiguity preference, to be partially separated. We define a risky environment to be one where probabilities are known, such that there is uncertainty over outcomes, and an ambiguous environment to be one where probabilities are unknown such that there is uncertainty over probabilities. Thus, naturally, uncertainty (i.e. the absence of certainty) can be categorized as the union of risk and ambiguity. In the *EV* treatment the agent faces certainty when monitoring probabilities are observable, and faces uncertainty when monitoring probabilities are unobservable. Therefore, the difference in driver behavior between the two monitoring environments in the *EV* treatment identifies the union of risk and ambiguity preferences. In the *Prob* treatment the agent faces risk when monitoring probabilities are

[8] We also include treatments where the observability of monitoring is exogenously determined.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

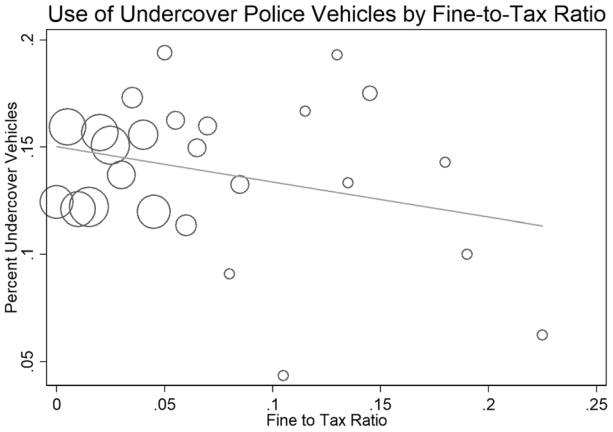


Fig. 1 Share of Government Revenues from Law Enforcement Fines and Fees & the Use of Unmarked Police Vehicles. $N = 1606$, bubble size corresponds to the number of observations in each bin. Data from the 2013 Law Enforcement Management and Administrative Statistics database and 2013 US Census

observable, and faces uncertainty when monitoring probabilities are unobservable. Therefore, the difference in driver behavior between the two monitoring environments in the *Prob* treatment identifies ambiguity preferences.[9] Given this structure of the experimental design, we infer preferences from driver behavior rather than measure preferences via a separate measurement task. We return to these themes in Sect. 3.

Our key theoretical result is to identify when law enforcement should *shroud* monitoring activity and when law enforcement should *reveal* monitoring activity as a function of the underlying preferences of officers and drivers, which is summarized in Table 1. When the officer is motivated by revenue (prevention) incentives then, in equilibrium, they will reveal (shroud) their monitoring strategy when the driver is uncertainty averse and shroud (reveal) their monitoring when the driver is uncertainty seeking. Intuitively, the officer discourages law breaking (by shrouding when facing an uncertainty averse driver) when motivated by prevention incentives but encourages law breaking (by shrouding when facing an uncertainty seeking driver) when motivated by revenue incentives. The more people who break the law, the more people that can be caught.[10]

[9] Formally, this identification requires an assumption of Expected Utility over risk. In an online appendix, we consider the effects of non-expected utility risk preferences that are generated by non-linear probability weighting. We simulate the effects of a Prelec weighting function, for typical parameter values, and find that probability weighting can explain only small deviations from Expected Utility in our environment. We conclude that, while non-expected utility risk preferences may explain a portion of the deviations from EU that is observed in our data, the majority of the observed deviations from Expected Utility are not caused by non-linear probability weighting.

[10] For clarity of argument we refer to the case where officers prefer to catch offenders, rather then prevent offenses from occurring, as the case where officers are seeking to maximize fine revenues. We recognize, but do not include in our model, that officers may have other non-pecuniary incentives that could generate a similar preference structure.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Our experimental results are broadly consistent with our model's equilibrium predictions. When monitoring levels are observed officers monitor at near equilibrium rates and drivers typically best respond, particularly in the *EV* treatment. When monitoring is unobserved, we find evidence that drivers do not exhibit uncertainty neutrality. In cases where the aggregate monitoring level is high, we observe evidence of uncertainty aversion. On the other hand, in cases where the aggregate monitoring level is low, we observe evidence of uncertainty seeking. This pattern of behavior could be explained by either heterogenous uncertainty preferences across subjects, or that some subjects exhibit prospect-theory-like uncertainty preferences.[11] Given the observed behavior of drivers it is optimal for officers to always reveal their monitoring strategies, and officers do reveal their monitoring levels approximately 70% of the time.

While our theoretical model is necessarily highly stylized the intuition behind the equilibrium is both intuitive and broadly applicable. When monitoring agents are uncertainty averse (uncertainty seeking), increasing uncertainty in the level of monitoring at any given location or instant will cause the agents to undertake safer (riskier) actions. Knowing this, the level of monitoring uncertainty can be manipulated to induce favorable behavior from the monitored agents. While our theoretical results apply to a specific domain, our experimental results suggest that the underlying logic does indeed generalize to at least one related domain; specifically, the domain where utility maximizing agents are replaced with human decision makers.

The paper proceeds as follows. Section 2 identifies connections between our paper and previous work on monitoring and enforcement in the domains of law enforcement and auditing, as well as existing work on the role of strategic uncertainty and ambiguity aversion. Section 3 describes our theoretical model and experimental design. Section 4 discusses the implementation of our experimental design. Section 5 presents the results of the experiment. Section 6 discusses the results and concludes the paper.

2 Literature review

This work contributes to two distinct research areas. As such, we outline the state of the literature in each of these research areas separately. We start by first discussing the literature on law enforcement, where we highlight various enforcement approaches that have been pursued in practice and those instances where these approaches have been examined within the lab. Second, we discuss the literature on ambiguity aversion and strategic uncertainty, focusing on the small experimental literature.

[11] We thank an anonymous referee for help in clarifying this argument.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Table 1 Law enforcement choice to either reveal or shroud monitoring probabilities, in equilibrium, as a function of the officer's incentive structure (column) and the driver's uncertainty preferences (row)

	Revenue incentives	Prevention incentives
uncertainty aversion	**Reveal**	Shroud
uncertainty seeking	Shroud	**Reveal**

Bold entries indicate the modal behavior in our experimental implementation of the model.

2.1 Compliance and enforcement literature

Research using observational data highlights the tension for law enforcement agencies in deciding whether to announce their enforcement strategies. Namely, does announcing policing activities or shrouding them from citizens enhance the likelihood of deterring law violators? Announcing the locations of checkpoints or utilizing signage that indicates the heavy presence of law enforcement in certain regions has the potential to deter proscribed activities. Indeed, as noted by the National Highway Traffic Safety Administration, "although forewarning the public might seem counterproductive to apprehending violators, it actually increases the deterrent effect." Alternatively, many law enforcement agencies have opted to utilize unannounced law enforcement practices. Indeed, major law enforcement agencies such as the Los Angeles, New York and San Francisco police departments do not announce their checkpoints. Moreover, nearly 90% of law enforcement agencies in the U.S utilize unmarked law enforcement vehicles (Bureau of Justice Statistics, 2013).

The tension between whether announcing enforcement strategies or not is impacting public safety has received some attention. Lazear (2006) highlights the role of the cost of learning for citizens that are being monitored by law enforcement. In short, he notes that "for high cost learners, and when the monitoring technology is inefficient, it is better to announce what will be tested. For efficient learners, deemphasizing the test itself is the right strategy. This is analogous to telling drivers where the police are posted when police are few. At least there will be no speeding on those roads. When police are abundant or when the fine is high relative to the benefit from speeding, it is better to keep police locations secret, which results in obeying the law everywhere." [12] Eeckhout et al. (2010) examine the use of arbitrary and publicized enforcement crackdowns (radar machines), noting that the marginal benefit of crackdowns is quite near to the marginal social cost. Banerjee et al. (2019) examine roadway checkpoints in India, and note that fixed location checkpoints are not advisable, as drivers learn about the locations of checkpoints and engage in strategic driving behavior to avoid checkpoints.[13] Lacey et al. (1999) examine a checkpoint program in Tennessee, whereby almost 900 checkpoints were well-advertised

[12] Uncertainty over the application of the law could occur inadvertently, though. As DeAngelo and Owens (2017) note, changes in legal threshold result in disparate application of the laws, depending on officer experience, which could be generating further uncertainty about the application of laws.

[13] See also Olken and Barron (2009).

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

on the television, radio and in newspapers, finding that the checkpoints can be attributed to an over 20% reduction in drunk driving fatal crashes relative to locations where checkpoints are not utilized. However, and as noted above, there appears to be a trend amongst law enforcement agencies to withhold advanced, public knowledge about law enforcements' whereabouts, especially with regards to checkpoints.

The issue of alerting the public of the location of concentrated law enforcement or not has been further complicated by technological innovations. For example, radar detection technology has enabled drivers to identify when and where speed traps are being set up. And more recently, smartphones and apps (such as *Waze*[14]) have enabled drivers to real-time crowd-source information about the location of law enforcement activities - including DWI check points, speed traps, and speed cameras - to other individuals using the app.[15]

As adjustments to law enforcement decisions to announce or veil their checkpoint locations and timing have occurred, a common discussion relating to law enforcement decisions is whether they are being conducted for purposes of enhancing safety, revenues, or both. For example, Carpenter et al. (2015) discuss policing for profit with regards to civil asset forfeiture laws and the protection of personal property. Makowsky and Stratmann (2009, 2011) and DeAngelo et al. (2019) examine fiscal incentives of law enforcement agencies on public safety, finding that police are sensitive to principal-agent issues that can result in disproportionate law enforcement presence in regions that generate greater revenues to the principle. Recent lawsuits have been brought against the city of Buffalo and state of Missouri noting that, among other things, checkpoints aimed to generate revenues, not enhance safety. However, other analyses (e.g, DeAngelo and Hansen (2014)) have shown that roadway safety officers have considerable effects on public safety.

Regardless of the objective function of the law enforcement agency, agencies will aim to identify proscribed activities. The use of strategic uncertainty appears to be prevalent in recent years for large law enforcement agencies. However, it is also likely present in smaller law enforcement agencies, although less well-documented in media outlets. For example, as law enforcement agencies experience budget cuts, they could utilize strategic uncertainty so as to obscure knowledge of staffing cuts to the general public. Since perceptions of police presence matter for reducing crime (Vollaard and Hamed, 2012), it could be the case that strategic uncertainty can maintain higher levels of perception of police presence.

Although it is difficult to exclusively identify the effect of uncertainty using observational data, the role of uncertainty in legal issues has minimally been explored in the laboratory. DeAngelo and Charness (2012) explore the role of uncertainty in enforcement, noting that subjects facing treatments with identical expected

[14] www.waze.com.

[15] In response to this functionality, the NYPD sent a cease-and-desist letter to *Google*, the parent company of *Waze* that the functionality must be removed from the app immediately, as "individuals who post the locations of DWI checkpoints may be engaging in criminal conduct since such actions could be intentional attempts to prevent and/or impair the administration of the DWI laws and other relevant criminal and traffic laws." In addition to *Waze*, other apps, such as *MrCheckpoint*, have also started appearing to notify the public of DWI checkpoints.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

costs but containing uncertainty over the enforcement regime that will be faced are less likely to engage in proscribed behavior. Harel and Segal (1999) explore the role of uncertainty in the legal system by exploring conduct of defendants that face bench (certainty) versus jury (uncertainty) trials.[16] Finally, Grechenig et al. (2010) examine uncertainty pertaining to contributions to a public good, finding that subjects are willing to punish even when there is uncertainty about the degree of pro-social conduct of a person. Moreover, the punishment associated with the uncertainty does not support higher levels of cooperation and results in net welfare decreases. However, a void still remains in examining the choice to use uncertainty as an enforcement strategy.

The theoretical model presented below can also be, in the abstract, applied to audit and compliance schemes. There has been several recent experimental studies of various compliance mechanisms. For example, Cason et al. (2016) and Gilpatric et al. (2015) use laboratory experiments to study the use of audit tournaments, and Gilpatric et al. (2011) also study endogenous audit mechanisms where the probability of audit depends on the reports of others. Cason and Gangadharan (2006) and Clark et al. (2004) study the use of leverage, where subjects are moved between two different inspection groups that differ with respect to the probability of inspection and severity of punishment. Our experimental design differs from each of these by allowing subjects, in the role of officers, to select whether monitoring levels are known or unknown. Finally, DeAngelo and Gee (2020) highlight that punishment can only be imposed when non-compliance is detected. They specifically explore differences in group versus peer endogenous and exogenous monitoring, noting that when monitoring is an endogenous choice groups fail to implement group monitoring, resulting in large decreases in contributions to a public good. Alternatively, peer monitoring results in similar levels of public good provision when either endogenously chosen or exogenously imposed.

2.2 Strategic uncertainty literature

There is a growing literature that considers the effects of ambiguity aversion and strategic uncertainty in games. Among the theoretical literature Hanany et al. (2020), who study incomplete information games with ambiguity averse players, is the most relevant to our experimental design. An application of the Hanany et al. (2020) framework to our environment could be constructed by allowing the Officer to condition their monitoring strategy on a privately observable ambiguous signal. The Hanany et al. (2020) framework can be applied to a very broad class of games, but requires making an assumption regarding the type of ambiguity preferences held by the agents in the game.[17] In contrast, we are able to take advantage of the symmetric nature of our game and allow for a broader interpretation of ambiguity aversion, but our approach does not generalize easily to other games.

[16] See also Baker et al. (2004).

[17] The assumption required is really an assumption about the nature of the ambiguous signal and the beliefs agents hold regarding the signal. Once a belief structure is specified, there are natural restrictions imposed on the ambiguity preferences of the agents.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

There also exists a broader theoretical literature on ambiguity aversion in games, chiefly focusing on games with complete information.[18] Bade (2011) and Riedel and Sass (2014) allow agents to construct ambiguous strategies, while Azrieli and Teper (2011) and Grant et al. (2016) provide alternative treatments of games with incomplete information.

There is a growing experimental literature on the role of ambiguity aversion and strategic uncertainty in games. Early contributions from Camerer and Karjalainen (1994) and Bohnet and Zeckhauser (2004) both find that subjects prefer to play games against known-probability objective randomization devices (e.g. coin flips) than against other people, reflecting an aversion to strategic uncertainty.

In a series of papers, David Kelsey and Sara le Roux (Kelsey and le Roux, 2015, 2017, 2018) find that ambiguity preferences affect strategic behavior, that there is little difference in the amount of ambiguity perceived in the behavior of foreign versus domestic opponents, that ambiguity has a stronger effect on behavior in games than on behavior in ball-and-urn tasks, and that there is often little correlation between expressed ambiguity preferences across environments. Relatedly, Eichberger et al. (2008) establish that subjects find "grannies" to be a greater source of strategic ambiguity than game theorists.

Both Li et al. (2019) and Ivanov (2011) elicit beliefs over opponent's strategies and, using different theoretical underpinnings, use the reported beliefs to infer uncertainty preferences. In both cases they find evidence of ambiguity aversion in games. Calford (2020) elicits subject uncertainty preferences via ball-and-urn tasks and correlates uncertainty preferences with behavior in games, finding a correlation between uncertainty aversion and strategic behavior. In a separate treatment, Calford (2020) induces beliefs about an opponent's uncertainty preferences and finds that subjects best respond to their opponent's safe strategy more often when their opponent is uncertainty averse.

While it is clear that uncertainty preferences affect behavior in games, and Calford (2020) finds that subjects rationally change their behavior when faced with an uncertainty averse opponent, there is currently no experimental evidence regarding the decision to withhold strategic information from an opponent to manipulate the opponent's reaction to strategic uncertainty.

3 Theoretical model and experimental design

We study a sequential game between two players, an officer and a driver. We begin by describing the game tree where, for analytical convenience, we model the game as a three-stage game with the officer having the move at the first and second stage, and the driver having the final move.

In the first stage, the officer chooses a total level of monitoring, $m \in [0, 1.8]$, and an information structure $I \in \{Obs, Unobs\}$. At the second stage, the officer chooses

[18] See Dow and da Costa Werlang (1994), Eichberger and Kelsey (2000), Eichberger and Kelsey (2014), Lo (1996, 2009) and Marinacci (2000) for complete information examples.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

how to partition their monitoring across two roads, A and B, such that $m_A + m_B = m$, where $m_A \in [0, 0.9]$ and $m_B \in [0, 0.9]$ denote the monitoring levels across the two roads, respectively. The officer's strategy, therefore, can be summarized by the vector $M = (m, m_A, m_B, I)$, subject to the constraint that $m_A + m_B = m$.

In the third stage, the information available to the driver depends on the officer's choice of I. If $I = Obs$ then the driver can observe m, m_A, and m_B; if $I = Unobs$ then the driver can observe only m, but not m_A or m_B. An action for the driver is a choice of either A, B or C, such that a strategy for the driver can be summarized by a pair of functions $D^{Obs} : [0, 0.9]^2 \to \{A, B, C\}$ and $D^{Unobs} : [0, 1.8] \to \{A, B, C\}$ that reflect the driver's decision in the case where monitoring is observed, or unobserved, respectively. The driver's choice of A or B can be interpreted as the driver speeding on the respective road, and C being a choice of not speeding.

Within this game tree we study four (two times two) distinct payoff structures. There are two variants of the driver's payoffs: one where the driver's payoff is deterministic (the *EV* treatment), and one where the driver's payoff is probabilistic (the *Prob* treatment). The payoffs in each case are equivalent for a risk neutral Expected Utility maximizing driver. There are also two variants of the officer's payoffs. In the first, the officer strictly prefers the driver choosing C over either A or B. In the second, the officer weakly prefers the driver choosing either A or B over C. We refer to these two cases as the crash minimization (*CrashMin*) treatment and revenue maximization (*RevMax*) treatments, respectively. The payoff functions are displayed in Tables 2 and 3.

For brevity, we analyze the four payment structures jointly.[19] In addition, we are interested in three different information structures. First, the full game as described above. Second, the game restricted such that $I = Obs$, and third, the game restricted such that $I = Unobs$. The second and third cases are subgames of the original game and, given that we use a subgame perfect solution concept, do not require any special treatment.

3.1 Bespoke design considerations

3.1.1 Incorporating two options for "speeding"

A, potentially, surprising element of the experimental design outlined above is that we provide the driver with two options for speeding, A and B, rather than just a binary choice between speeding or not speeding. The reason for doing so is that, in equilibrium, monitoring uncertainty cannot be sustained in our structure when driver choice is binary.

Consider the following amendment to our experimental design. The driver may only select roads A or C, the officer selects a single monitoring variable m_A, the officer may choose to either announce or shroud m_A, and the officer seeks to prevent the driver from speeding subject to minimizing monitoring costs. Suppose that, in the case that m_A is announced, the game has a unique equilibrium value of $m_A = m_A^*$.

[19] Note that whenever the game is implemented, the payoff functions that are in use are common knowledge. The differing payoff structures do not induce a Bayesian game.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

In equilibrium, m_A^* must be the smallest level of monitoring that makes speeding unattractive to the driver.

Next, suppose that the officer elects to shroud monitoring. The driver can deduce that, despite monitoring being unobservable, the level of monitoring will never be above m_A^*: if monitoring were above m_A^*, then the officer would be better off by reducing monitoring to m_A^* and revealing that monitoring is m_A^*. Thus, the driver believes that monitoring will be at, or below, m_A^* when monitoring is shrouded and the driver may speed. Given this, the officer prefers to reveal the monitoring level.

The experimental design outlined above, allowing the driver to speed on either road *A* or road *B*, separates the total level of monitoring from the distribution of monitoring resources. In doing so, the design constructs an environment where uncertainty over the distribution of monitoring resources can persist in equilibrium.

3.1.2 The safety of a "safe" option

Compare the role of the safe option, the driver choosing *C*, across the *EV* and *Prob* treatments. In the *EV* treatment, *C* pays a certain 50 points. In the case where monitoring is unobserved, choosing *C* removes uncertainty over payoffs (relative to choosing *A* or *B*). In the *Prob* treatment, *C* pays 100 points with probability $\frac{1}{2}$. In the case where monitoring is unobserved, choosing *C* removes uncertainty over probabilities (relative to choosing *A* or *B*).

In each treatment, *C* is relatively safer than choosing *A* or *B* when monitoring is unknown. That *C* is safer, but not necessarily safe, is consistent with the underlying real-world phenomena that is being modeled. Reckless driving, speeding and ignoring road rules, is obviously risky. Obeying all speed limits and road rules reduces the risks of driving, but does not render driving a "safe" activity—even if all drivers obeyed all relevant road rules, accidents and traffic fatalities would still occur.

Note, in particular, that the relationship between road *C* and roads *A* and *B* in the *Prob* treatment, when monitoring is unobserved, is designed to reflect the typical Ellsberg urn experiments that are used to measure ambiguity preferences (see Halevy (2007), for example). In the Ellsberg urn experiments, ambiguity preference is elicited by asking a subject to choose between a bet on a ball drawn from an urn with a known composition and a bet on a ball drawn from an urn with an unknown composition. Here, in a similar fashion, road *C* generates a known probability of winning 100 points while roads *A* and *B* generate an unknown probability of winning 100 points. Thus, driver behavior in the *Prob* treatment, when monitoring is unobserved, can be used to infer the ambiguity preferences of drivers.[20]

3.1.3 The inclusion, or exclusion, of preference measurement tasks

The experimental design choices outlined in the previous subsections are structured such that the direction of ambiguity, and uncertainty, preference can be directly

[20] We require an assumption of expected utility over risk. See footnote 9.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Table 2 Driver payoffs, in points. $\{x_1, p_1; x_2, p_2\}$ denotes the lottery which pays x_1 points with probability p_1 and x_2 points with probability p_2

	A	B	C
Probabilistic treatment	$\{100, 0.9 - m_A; 0, 0.1 + m_A\}$	$\{100, 0.9 - m_B; 0, 0.1 + m_B\}$	$\{100, 0.5; 0, 0.5\}$
Expected value treatment	$100 \times (0.9 - m_A)$	$100 \times (0.9 - m_B)$	50

Table 3 Officer payoffs, in points. $\mathbb{1}(A)$ is an indicator variable that equals 1 if the driver chooses road A, and 0 otherwise

Revenue Maximization treatment	$20 + 100m_A\mathbb{1}(A) + 100m_B\mathbb{1}(B) - 20(m_A + m_B)$
Crime Minimization treatment	$80 - 40(\mathbb{1}(A) + \mathbb{1}(B)) - 20(m_A + m_B)$

inferred from driver behavior. An alternative design approach, not taken here, would be to use external preference measurement tasks and then correlate behavior in the measurement tasks with behavior in the game of interest.

The prior experimental research that has compared behavior in games to ambiguity preferences as measured in individual decision tasks, outlined in Sect. 2.2, has found that measures of ambiguity preferences are often source dependent. That is, comparisons of ambiguity preference in games to ambiguity preferences in individual decision tasks are not necessarily stable either across subjects or different classes of games. This notion of source dependent ambiguity preference can be traced back to, at least, Fox and Tversky (1995).

For example, consider the following two hypothetical subjects. Subject A is ambiguity averse, and reveals this ambiguity preference in an individual measurement task. However, subject A also believes that they are extremely skilled at predicting the behavior of others. Thus, when playing a game, subject A does not *subjectively* experience ambiguity, and holds very strong and precise beliefs about the expected behavior of others. Subject A would appear to be ambiguity neutral while playing games. On the other hand, Subject B, who is also ambiguity averse, is very uncertain about how others will behave in a game. In fact, subject B subjectively perceives a game to be even more ambiguous than an Ellsberg urn. Thus, subject B exhibits stronger ambiguity aversion in games than individual Ellsberg urn tasks, while subject A has the opposite preference pattern. The interpretation of source preference being driven by ambiguity perception is compatible with, for example, recent experimental evidence that separately identifies both how much ambiguity an agent perceives in a given environment and the ambiguity preferences of the agent (Baillon et al., 2018; Li et al., 2019).

A distinct argument against using a separate individual preference measurement task is that there may be cross task contamination between the game and the preference measurement task. Recently, Baillon et al. (Forthcoming) have outlined the empirical difficulties in observing ambiguity preference in environments where subjects make decisions across multiple ambiguity measurement tasks. They find that

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

approximately one-half of subjects behave as if they are able to use the multiple tasks to construct a hedge against ambiguity at the experiment level.

Similar arguments, albeit using different theoretical underpinnings, can be found in the literature on the measurement of risk preferences. Recently, Holzmeister and Stefan (2021) have argued that differences in measured risk preference across measurement tasks are associated with subjects' subjective *perceived* riskiness across tasks. Thus, an empirically valid measurement of risk preferences would require an environment in which subjects perceived the degree of riskiness to be approximately equal to the riskiness of our monitoring game. Further, in order to make meaningful comparisons between a risk preference measurement task and behavior in the monitoring game would require a task that allows for parametric estimation of risk preferences. Charness et al. (2013) refer to such tasks as "complex" tasks, and question whether they are useful for cross-domain prediction.

Given these difficulties of observing and correlating preferences across multiple sources and tasks, we do not use an individual preference measurement task. In light of the above discussion, the preferences that are identified in this paper should be interpreted as source specific (the response to uncertainty generated by a strategic monitoring decision of another person).

3.1.4 The modeling approach and solution concept

We use a non-standard modeling approach that is particularly suited to the experimental implementation of games of imperfect information where agents might have non Expected Utility preferences. Here we outline the approach and discuss potential applications, and we defer the details to the online Appendix. In environments where an assumption of EU can be justified, a game such as ours can easily be modeled using standard techniques. However, generalizing Bayesian Nash Equilibrium to accommodate ambiguity averse agents presents several challenges. Worryingly, equilibrium can, in general, be sensitive to the choice of preference model and belief updating rules.[21]

Our modeling approach is to impose restrictions, via a set of axioms, on the mapping from the observable level of monitoring to the utility of the driver.[22] The axioms are model free, but are consistent with both Expected Utility and standard

[21] Should agents be modeled as holding Maxmin Expected Utility preferences (Gilboa and Schmeidler, 1989), Choquet Expected Utility preferences (Schmeidler, 1989), Smooth Ambiguity Averse preferences (Klibanoff et al., 2005), or something else? Depending on the choice of preference structure, there may also be a variety of plausible belief updating rules, including Full Bayesian updating or the Dempster-Shafer rule (see Section 7 of Machina and Siniscalchi (2013) for a detailed discussion). Recently, Hanany et al. (2020) provided an analysis of equilibrium satisfying sequential optimality for agents with Smooth Ambiguity preferences, and conclude that sequential rationality can pin down a specific updating rule. Hanany et al. (2020) focus on providing a specific modeling approach that is applicable to a wide array of games, while we instead apply a more general axiomatic preference structure to our specific game.

[22] Alternatively, as is standard in the decision theory literature, we could begin with a preference relation and then define axioms over the preference relation. Given the game theoretic context here, it seems more natural to place restrictions on the utilities instead.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

models of ambiguity aversion. Effectively, we use our axioms to convert the game from a three stage game of imperfect information to a two stage game of perfect information and we can therefore use Subgame Perfect Nash Equilibrium to solve the game. The axioms have direct and testable implications for driver behavior, which we test for below.

Consequently, the equilibrium will only pin down actions that are observable to the driver: for example, the choice of m_A and m_B are unrestricted when monitoring is unobserved, beyond the mechanical constraint that $m_A + m_B = m$. This degree of freedom in the equilibrium is a natural consequence of allowing for ambiguity averse agents: if ambiguity aversion affects play along the equilibrium path of a game, then there must exist ambiguity along the path.

This axiomatic approach is not restricted to our particular game, and could be applied more generally, although in some games it may not be possible to formulate a set of axioms that are both palatable and tractable. In our environment, we are able to exploit the natural symmetry that exists across roads A and B to generate a satisfactory set of axioms.

3.2 Driver preferences and best response

We impose as little structure on the preferences of the driver as possible, assuming only monotonicity, symmetry, non-triviality and continuity of the driver's utility function. Specifically, we do not assume expected utility and do not make assumptions about the nature of the driver's preferences over uncertainty.

The symmetry assumption ensures that driver preferences depend only on the degree of monitoring, and not on the label of the roads, and monotonicity guarantees that drivers prefer roads with less monitoring. Non-triviality assumes that drivers will prefer to speed when monitoring is minimal, and prefer not to speed when monitoring is maximal. Finally, continuity ensures that a point of indifference exists. Details are provided in the Online Appendix. We denote the best response correspondence of the driver by a pair of mappings from monitoring to a subset of roads: $b^{\text{Obs}} : [0, 0.9]^2 \to 2^{\{A,B,C\}}$ and $b^{\text{Unobs}} : [0, 1.8] \to 2^{\{A,B,C\}}$.

Lemma 1 *Assume that the driver's utility function satisfies Monotonicity, Symmetry, Non-triviality and continuity. Then*

$$b^{\text{Obs}}(m_A, m_B) = \begin{cases} C & \text{if } \min\{m_A, m_B\} > 0.4 \\ A & \text{if } m_A < 0.4, m_A < m_B \\ B & \text{if } m_B < 0.4, m_B < m_A \\ \{A, B\} & \text{if } m_A = m_B < 0.4 \\ \{A, C\} & \text{if } m_A = 0.4 < m_B \\ \{B, C\} & \text{if } m_B = 0.4 < m_A \\ \{A, B, C\} & \text{if } m_A = m_B = 0.4 \end{cases} \tag{1}$$

and there exists a value $0 < \overline{m} < 1.8$ *such that*

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

$$b^{\text{Unobs}}(m) = \begin{cases} C & \text{if } m > \overline{m} \\ \{A, B\} & \text{if } m < \overline{m} \\ \{A, B, C\} & \text{if } m = \overline{m}. \end{cases} \tag{2}$$

Proof See Online Appendix. □

The driver's best response correspondence is intuitive. Consider first the case of observed monitoring: It is immediate, given the payoff functions in Table 2, that the point of indifference occurs when $m_A = m_B = 0.4$. When the road with the least monitoring presents a better proposition than choosing C (i.e. $m_A < \min\{0.4, m_B\}$ or $m_B < \min\{0.4, m_A\}$), then the driver should choose the least monitored road. On the other hand, when C provides a better proposition than either road (i.e. $0.4 < \min\{m_A, m_B\}$), then the driver should choose C.

For the case of unobserved monitoring, the assumptions on driver preferences guarantee the existence of a monitoring level $\overline{m}$ such that the driver is indifferent between speeding or not speeding. When monitoring is above this reference point C is preferred, and when monitoring is below the reference point A or B is preferred.[23] The indifference point for the driver, $\overline{m}$, provides a measure of the uncertainty preferences of the driver. A risk and ambiguity neutral driver will be indifferent when $\overline{m} = 0.8$. In both the *EV* and *Prob* treatments ambiguity aversion will lower the indifference point ($\overline{m} < 0.8$) and ambiguity seeking will raise the indifference point ($\overline{m} > 0.8$). Expected Utility risk preferences have no effect on $\overline{m}$ in the *Prob* treatment, while risk aversion lowers, and risk seeking raises, the indifference point in the *EV* treatment.[24]

3.3 Officer behavior and equilibrium characterization

Given the simple cutoff rule best response behavior of drivers, the optimal strategy for the officer is the intuitive one: in the *CrimMin* treatment, monitor just enough to prevent speeding; in the *RevMax* treatment, monitor up to the maximal level such that the driver still speeds. Therefore, in equilibrium, no drivers speed in *CrimMin* and all drivers speed in *RevMax*, yet monitoring levels in the two treatments are indistinguishable.[25]

Beginning with the *CrimMin* treatment, the officer strictly prefers to induce the driver to choose road C because the payoff penalty when a driver chooses roads A

[23] Note that the symmetry assumption ensures that the driver is always indifferent between A and B.

[24] Note that for drivers who are risk seeking but ambiguity averse, or risk averse but ambiguity seeking, the effects of uncertainty preference are indeterminate and depend on the relative strength of risk and ambiguity preferences.

[25] This implies an interesting corollary. Suppose that an outside observer wished to determine whether the officer has *RevMax* or *CrimMin* preferences. The outside observer can only do so by observing the driver behavior, and not the officer behavior, because the officer behavior is indistinguishable in both cases. If drivers are observed to speed, the officer is revenue focused. If drivers do not speed, then the officer is safety focused.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

or B (40 points) is larger than the maximal monitoring cost (36 points). Therefore, given the driver's best response, $m_A, m_B \geq 0.4$ in the observed monitoring case and $m_A + m_B \geq \overline{m}$ in the unobserved monitoring case. Because monitoring is costly, the officer wishes to minimize monitoring subject to this constraint. In the observed monitoring case we have an equilibrium at $m_A^* = m_B^* = 0.4$ where the driver chooses road C and in the unobserved monitoring case we have an equilibrium at $m_A^* + m_B^* = \overline{m}$ where the driver chooses road C.[26]

In the *RevMax* treatment the marginal revenue from additional monitoring exceeds the marginal cost when the additional monitoring does not alter the driver's behavior. Therefore, the officer wishes to monitor up to the point where the driver switches to choosing road C. Therefore, in equilibrium, $m_A^* = m_B^* = 0.4$ in the observed monitoring case and $m_A^* + m_B^* = \overline{m}$ in the unobserved monitoring case.[27]

When the officer can choose whether to fully reveal their monitoring strategy or not, the officer should reveal their monitoring if and only if their payoff is higher in the observed monitoring case. The payoff difference between the observed and unobserved case depends on whether $\overline{m}$ is greater or less than 0.8. In the *RevMax* (*CrimMin*) treatment officer payoffs are increasing (decreasing) in monitoring – therefore the officer will reveal their strategy when $\overline{m} < 0.8$ in the *RevMax* treatment and when $\overline{m} > 0.8$ in the *CrimMin* treatment.

We formalize the above intuition in the Online Appendix, and we summarize the full set of equilibrium predictions across all treatments in Tables 4, 5 and 6.

Note that when monitoring is observable, the equilibrium is identical in the *Prob* and *EV* treatments. Further, conditional on $\overline{m}$, when monitoring is unobservable, the equilibrium is also identical in the *Prob* and *EV* treatments. Therefore, the only distinction between the two treatments arises from the potential for the driver to have different attitudes towards uncertainty across the two treatments.

The experimental design is similar to the theoretical framework developed here. Therefore, we hypothesize that behavior across the two treatments should differ only to the extent that driver preferences differ across the two treatments. We focus our hypotheses on two key metrics: the total level of monitoring, m, and the probability that a driver chooses not to speed, $\Pr(C)$.

Our first three hypotheses are direct implications of Lemma 1.

Hypothesis 1 When monitoring is observed (a) total monitoring, m, and (b) the probability of not speeding, $\Pr(C)$, are the same across the *Prob* and *EV* treatments.

[26] In the case of continuous monitoring the total monitoring level in equilibrium is unique. However, if the officer's strategy space is discrete, as is the case in our experimental implementation, then there exists an equilibrium where the driver chooses A or B when $m_A + m_B = \overline{m}$, and the officer chooses total monitoring to be slightly above $\overline{m}$.

[27] In the case of continuous monitoring the total monitoring level in equilibrium is unique. However, if the officers strategy space is discrete, as is the case in our experimental implementation, then there exists an equilibrium where the drive chooses C when $m_A + m_B = \overline{m}$, and the officer chooses total monitoring to be slightly below $\overline{m}$.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Hypothesis 2 When monitoring is observed (a) $\Pr(C) = 0$ when $m > 0.8$ and (b) $\Pr(C) = 1$ when $m < 0.8$.

Hypothesis 3 The probability of choosing C, $\Pr(C)$, is non-decreasing in monitoring for all treatments.

Our final hypothesis is an equilibrium condition on the officer's behavior, and follows immediately from the results presented in Table 4.

Hypothesis 4 (a) The officer chooses to reveal monitoring in the *RevMax* treatment and shroud monitoring in the *CrimMin* treatment if the driver is uncertainty averse. (b) The officer chooses to shroud monitoring in the *RevMax* treatment and reveal monitoring in the *CrimMin* treatment if the driver is uncertainty seeking.

4 Experimental implementation

All experimental sessions were conducted at the Vernon Smith Experimental Economics Laboratory (VSEEL) at Purdue University using undergraduate students, recruited via ORSEE (Greiner, 2015), as subjects. The experimental interface, implemented using oTree (Chen et al., 2016), was tested and refined using pilot sessions, and the data from the pilot sessions are not included in our analysis.[28] As we discussed in the previous section, subjects were divided across two roles, drivers and officers, although neutral language was used throughout the experiment to avoid potentially loaded terms, such as "driver" or "police".[29] Additionally, subjects participated in a single treatment, *Prob* or *EV*. The *Prob* treatment was conducted across 6 sessions using 110 subjects and the *EV* treatment was conducted across 4 sessions with 72 subjects.[30] The mean age of our subjects is just over 21 years old, just over half are Engineering or Science majors, approximately 45% are female, and they had taken 2.0 economics courses on average.[31] No subject participated in more than one session. Figure 2 provides a visual explanation of the overall experimental design.

It is apparent from Fig. 2 that there were a few design differences between the *Prob* and *EV* treatments. Before outlining the differences, we discuss the design philosophy that lead to the design choices. Recall that Sect. 3.2 describes the equilibrium as a function of driver preferences. The mapping from preferences to equilibrium is invariant to the differences in design, including whether payment is in points or binary lotteries, between the *Prob* and *EV* treatments. We therefore view

[28] The pilot sessions were chiefly used to assess subject comprehension of the experiment and, as a function of subject feedback and the pilot data, we simplified the experimental design for the main sessions.

[29] The roles were described to the subjects as 'worker' and 'supervisor'.

[30] The number of subjects is not balanced across treatments because we observe each subject in both roles in the *EV* treatment, but only observe subjects in one role in the *Prob* treatment.

[31] A breakdown of subject characteristics by treatment is provided in the Online Appendix.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Table 4 Officer's equilibrium choice of $I \in \{\text{Obs}, \text{Unobs}\}$ as a function of the officer's incentive structure (column) and the driver's uncertainty preferences (row)

	RevMax treatment	*CrimMin* treatment
$\overline{m} < 0.8$	$I = \text{Obs}$	$I = \text{Unobs}$
$\overline{m} > 0.8$	$I = \text{Unobs}$	$I = \text{Obs}$

Table 5 Equilibrium monitoring levels as a function of the observable information structure. The same equilibrium monitoring levels hold in the case where the officer chooses I and the cases where I is fixed endogenously

	$I = \text{Obs}$	$I = \text{Unobs}$
Equilibrium monitoring level	$M = (0.8, 0.4, 0.4)$	$M = (\overline{m}, m_A, m_B)$ with $m_A + m_B = \overline{m}$

$M = (m, m_A, m_B)$ indicates that the officer monitors road A with probability m_A, road B with probability m_B and that total monitoring is m.

Table 6 Equilibrium driver action observed on the equilibrium path

	RevMax treatment	*CrimMin* treatment
Equilibrium on-path driver action	A or B	C

the differences in design between the two treatments as providing us with a form of a robustness check: if the results are sensitive to minor design changes that are, theoretically, inconsequential then our theoretical framework is not well suited to addressing our research questions.

The theoretical predictions depend on the indifference point, $\overline{m}$, of the driver. Clearly, as outlined above, the change from payment in points in the *EV* treatment to payment in binary lotteries in the *Prob* treatment will, and is intended to, affect $\overline{m}$. It is also possible that some of the other, more minor, changes between treatments could affect $\overline{m}$. If such changes do occur, then it is not possible to attribute differences in $\overline{m}$ across the two treatments to the change in payment method.

The *Prob* treatment consisted of 36 rounds, consisting of 9 blocks with 4 rounds each, with subjects remaining in the same role (officer or driver) for the entire experiment, with an exchange rate of 100 points = \$1.20 and average earnings of \$24.63 (including a \$5 show up fee). The *EV* treatment consisted of 48 rounds, consisting of 12 blocks of 4 rounds each, with subjects switching roles after 24 rounds, an exchange rate of 100 points = \$0.90, and average earnings of \$25.21 (including a \$5 show up fee).[32] In both treatments opponents were randomly re-matched every period and subjects were paid the sum of their earnings across all rounds.

[32] We had subjects play both roles in the *EV* treatment to avoid concerns that subjects in the worker role may become bored and lose focus in the simpler environment. Exchange rates were varied to ensure that total payments across all rounds would be approximately equal across treatments.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

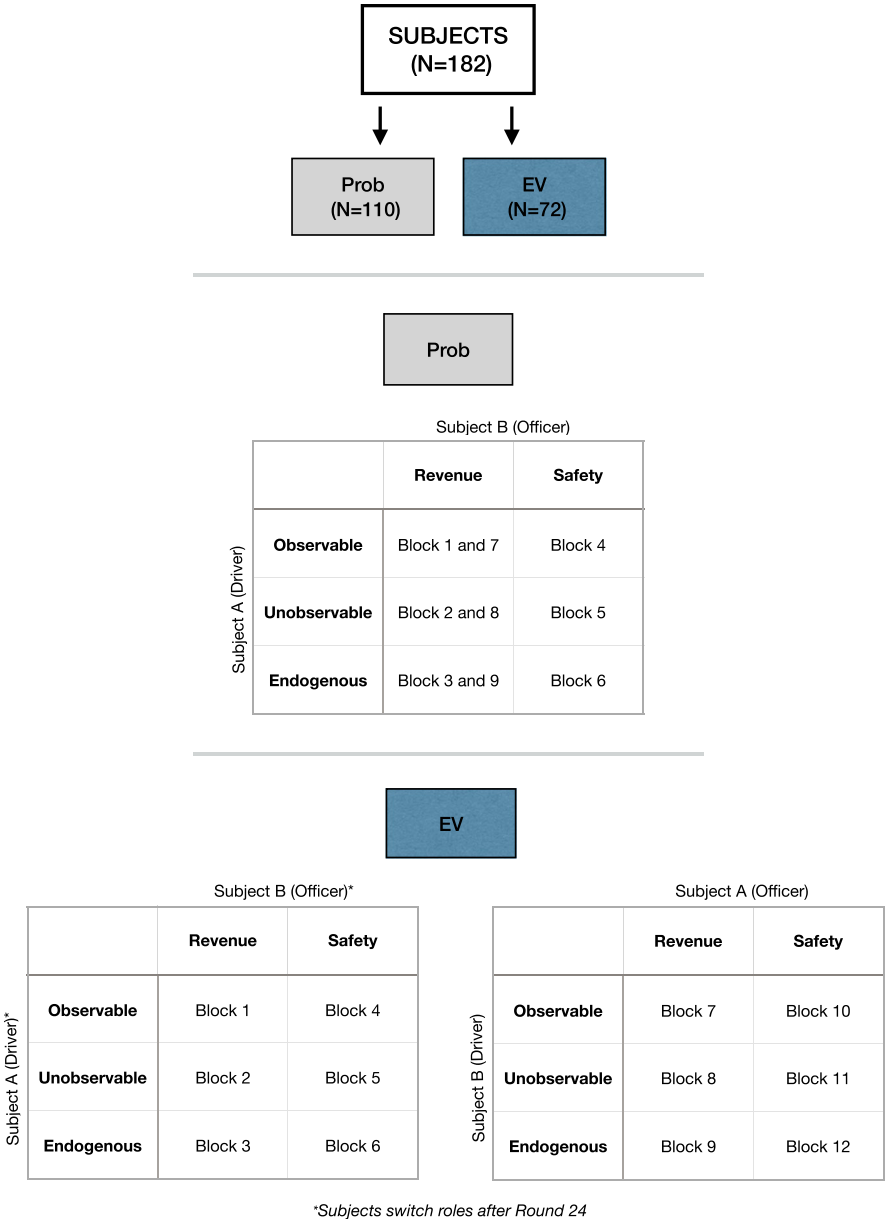


Fig. 2 Visual explanation of experimental treatments, illustrating the order of treatments for a single session of each of the *Prob* and *EV* treatments. The order of blocks was varied between sessions, and each block consisted of four rounds

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

As seen in Fig. 2, subjects participated in 6 different games within a session. Each of the games had one of the *CrimMin* or *RevMax* officer payoff schemes and one of the three information structures (I = Obs, I = Unobs or endogenous I). Rounds were grouped into blocks of four rounds, with the game remaining constant within blocks but changing between blocks. The assignment of games to blocks varied across sessions, with the requirement that subjects played one block of I = Obs and one block of I = Unobs before playing the corresponding block with endogenous I. The blocks were ordered in this fashion because the I = Obs and I = Unobs treatments are each subgames of the endogenous information treatment; by using this ordering, subjects experienced both subgames before playing the full game.

Supervisors were provided with the opportunity to choose the monitoring level for each of option A and B. They were informed that their choice of A and B will impact their own payoff, as well as that of the worker. The specific impact that the choices of A and B would have on the supervisor's payoff depended on which payoff scheme that the supervisor faced, as supervisor's were informed that there were two possible payoff schemes. In the first payoff scheme, increases in the monitoring level for option A increased the supervisor's payoff if the worker chose option A, but reduced the supervisor's payoff if the worker chose option B or C. Alternatively, in the second payoff scheme, supervisors were informed that increases in monitoring for option A or B resulted in a reduction in the supervisor's payoff the worker chose any option, A, B or C.

Workers were provided with the opportunity to choose an option, A, B or C, which were described as "tasks." The workers were informed that only their payoffs would be visible when choosing their task, and not the payoff for the supervisor. The workers were further informed that there were three informational environments that would display their payoffs associated with options A, B or C. In the first environment, the worker would be able to see their exact payoff from their task choice. In the second environment, the worker would only be able to see a range of possible payoffs should they choose option A or B, but would know their exact payoff if they chose option C. In the third environment, the worker was informed that the supervisor would have control over whether the worker's were in the first or second informational environment. Finally, subjects were informed that the experiment would consist of multiple rounds and that their final payment would be calculated by adding together the earnings from each round, converting their earning from experimental currency units to dollars, and then adding a $5 show-up payment.

The instructions, which are provided in an Online Appendix, were augmented with printed color-coded payoff tables. Subjects could use the tables to quickly look up payoffs given a conjecture of both their own and their opponents' actions. Subjects received a copy of the instructions and payoff tables, and the instructions were read aloud to subjects, at the beginning of each session. Subjects were able to ask clarifying questions at any time; subjects addressed their questions to the experimenter quietly, and the experimenter (after verifying that the question was indeed a clarifying question) then announced both the question and answer to the entire session.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

5 Results

In the spirit of backwards induction, we begin by estimating a logit choice model that summarizes driver behavior and testing the assumptions we imposed on driver preferences. Next, we present heat maps that summarize the monitoring behavior of officers. Finally, we test our empirical hypothesis and discuss factors that might influence the observed pattern of driver behavior.

5.1 Driver behavior

The theoretical analysis and equilibrium predictions developed in Sect. 3.2 are deterministic. To bring the model to the data, we need to allow for stochastic choice. Given that our theoretical model is entirely non-parametric, and generalizes several possible functional forms for the utility of drivers, we model driver choice purely as a function of the observable level of monitoring. Even though we use a reduced form stochastic choice model, our assumptions over driver preferences imply testable implications for the estimated parameters.

To compare the *Obs* and *Unobs* cases on the same scale, we define the following function:

$$V^I(m_A, m_B) = \begin{cases} \min\{m_A, m_B\} & \text{if } I = \text{Obs} \\ \frac{m_A + m_B}{2} & \text{if } I = \text{Unobs.} \end{cases}$$

$V^I(m_A, m_B)$ denotes the expected level of monitoring that will be faced by the driver, conditional on the driver choosing optimally between A and B. Clearly, when monitoring is observed the driver should choose the road with less monitoring; when monitoring is unobserved the expected monitoring on either road is given by $\frac{m}{2} = \frac{m_A + m_B}{2}$.

We consider the following binary choice model

$$y_{i,r} = \begin{cases} 1 & \text{if } \beta_0 + \beta_1 \mathbb{1}(\text{Obs}) + \beta_2 V^I_{i,r}(m_A, m_B) + \beta_3 \mathbb{1}(\text{Obs}) V^I_{i,r}(m_A, m_B) + \epsilon_{i,r} \geq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where $y^I_{i,r}$ is an indicator variable equal to 1 if subject i chooses option C in round r, and equal to 0 if the subject chooses options A or B, $\mathbb{1}(\text{Obs})$ is an indicator variable equal to 1 if the subject observes the monitoring choices of the officer, and $\epsilon_{i,r}$ is a mean zero error term following a logistic distribution. We estimate Equation 3 independently for the between subject $t = Prob$ and $t = EV$ treatments. Data from the within subject crash minimization and revenue maximization treatments are pooled, and we estimate generalized population averaged parameters (across subjects) with robust standard errors.[33] The results are presented in Table 7. In the

[33] We tested for order effects by including order effect dummy variables and found no significant results. We did not include the order effects in the regressions below, although doing so does not change the results.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Online Appendix we re-run the analysis with demographic controls: the parameter estimates are essentially unchanged.

While our empirical model is a reduced form model, our axiomatic assumptions over driver preferences imply a set of testable implications on the estimated parameters. The tests are summarized in Table 8. Monotonicity implies that the coefficients on $V^I_{i,r}(m_A, m_B)$ should be positive: as monitoring increases, C should be chosen more often. Symmetry only generates a testable assumption for the case of observed monitoring: when the least monitored road has a monitoring rate of 0.4, $V^I_{i,r}(m_A, m_B) = 0.4$, the driver should be indifferent between choosing C or not. Non-triviality generates two implications for each treatment. When $V^I_{i,r}(m_A, m_B) = 0$ the driver should prefer not to choose C, and when $V^I_{i,r}(m_A, m_B) = 0.9$ the driver should prefer C.

Table 9 presents the results of the tests outlined in Table 8. Each equation in Table 9 represents a Null hypothesis that is tested against the natural Alternative. For the Monotonicity and Non-triviality tests, rejection of the Null hypothesis is consistent with the underlying behavioral assumption, while for Symmetry rejection of the Null is inconsistent with the behavioral assumption. As Table 9 shows, the data is consistent with Monotonicity and Non-triviality for both the *EV* and *Prob* treatments and thus also provides support for Hypothesis 3. The data is consistent with Symmetry for the *EV* treatment, but not for the *Prob* treatment.

The Symmetry assumption requires, for the case where monitoring is observable, that the driver is indifferent between choosing C, which pays 100 points with probability 0.5, and not choosing C when the better option of A and B pays 100 points with probability 0.5.[34] In other words, the driver should be indifferent between C and not C when $\min\{m_A, m_B\} = 0.4$ and, in the logistic choice model estimated above, this implies that the driver should select C with probability 0.5 when $\min\{m_A, m_B\} = 0.4$. However, in the *Prob* treatment drivers select C with a predicted probability of only 0.34 in this case (with a 95% confidence interval of [0.28, 0.39]), and are predicted to select C with probability 0.5 when $\min\{m_A, m_B\} = 0.433$.[35] The violation of symmetry is highly significant, but not economically large. The slight deviation from Symmetry is likely the result of some drivers in the *Prob* treatment exhibiting spiteful preferences by not selecting the option that the officer intended the driver to choose. We do not observe the same effect in the *EV* treatment because subjects switched roles in that treatment, allowing each subject the opportunity to "control" the environment in at least some rounds.

The violation of symmetry for the *Prob* treatment but not the *EV* treatment also provides some evidence against Hypothesis 1, that behavior should be identical across the two treatments when monitoring is observed. Further, as is documented in Fig. 3 below, there is greater heterogeneity in the *Prob* than EV treatment as well.

[34] For the *EV* treatment, replace the statements "100 points with probability 0.5" with "50 points".

[35] For comparison, in the *EV* treatment, drivers select C with a predicted probability of 0.47 when $\min\{m_A, m_B\} = 0.4$ (with a 95% confidence interval of [0.43, 0.51]), and are predicted to select C with probability 0.5 when $\min\{m_A, m_B\} = 0.402$.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Table 7 Population average GEE parameter estimates of equation 3, with robust standard errors in parentheses.

	EV treatment	Prob treatment
Constant	-3.94^{***}	-3.09^{***}
	(0.41)	(0.33)
$\mathbb{1}$(Obs)	-30.77^{**}	-5.80^{***}
	(8.99)	(1.14)
$V^I_{i,r}(m_A, m_B)$	9.43^{***}	6.00^{***}
	(0.91)	(0.72)
$\mathbb{1}(\text{Obs})V^I_{i,r}(m_A, m_B)$	77.01^{**}	14.52^{***}
	(22.28)	(2.77)
Observations	1728	1980
Subjects	72	55

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 8 Testable implications of our axiomatic assumptions over preferences

	Unobserved monitoring	Observed monitoring
Monotonicity	$\hat{\beta}_2 > 0$	$\hat{\beta}_2 + \hat{\beta}_3 > 0$
Symmetry		$\hat{\beta}_0 + \hat{\beta}_1 + 0.4(\hat{\beta}_2 + \hat{\beta}_3) = 0$
Non Triviality	$\hat{\beta}_0 < 0$	$\hat{\beta}_0 + \hat{\beta}_1 < 0$
	$\hat{\beta}_0 + 0.9\hat{\beta}_2 > 0$	$\hat{\beta}_0 + \hat{\beta}_1 + 0.9(\hat{\beta}_2 + \hat{\beta}_3) > 0$

In our model of stochastic choice, the natural definition of uncertainty preferences is that an uncertainty averse agent should choose the more uncertain option less often than an uncertainty neutral agent. We write $\Pr^I_t(C|V)$ to denote the probability of a driver choosing C, in treatment t, under information scheme I, when the expected level of monitoring is V.

Definition 1 A driver is

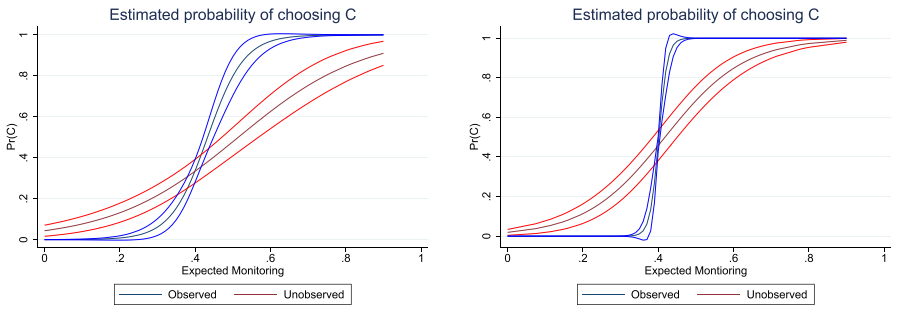
- uncertainty neutral if $\Pr^{\text{Obs}}_{\text{EV}}(C|V) = \Pr^{\text{Unobs}}_{\text{EV}}(C|V)$;
- uncertainty averse if $\Pr^{\text{Obs}}_{\text{EV}}(C|V) < \Pr^{\text{Unobs}}_{\text{EV}}(C|V)$;
- uncertainty seeking if $\Pr^{\text{Obs}}_{\text{EV}}(C|V) > \Pr^{\text{Unobs}}_{\text{EV}}(C|V)$.

For any fixed level of V then the expected value of choosing optimally between A and B is the same across the *Obs* and *Unobs* treatments and, therefore, the utility of choosing either A or B must be equal across treatments for an uncertainty neutral driver. Given our symmetry assumptions, the choice problem faced by a driver in the *Unobs* treatment is a mean preserving spread of the choice problem faced in the *Obs* treatment: an uncertainty averse (seeking) driver dislikes the mean preserving spread and therefore chooses C more (less) often in the *Unobs* case.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Table 9 p-values for tests of parameter restrictions associated with Monotonicity, Symmetry and Non-triviality

	EV treatment	Prob treatment
$\hat{\beta}_2 \leq 0$	$p < 0.001$	$p < 0.001$
$\hat{\beta}_2 + \hat{\beta}_3 \leq 0$	$p < 0.001$	$p < 0.001$
Monotonicity satisfied?	Yes	Yes
$\hat{\beta}_0 + \hat{\beta}_1 + 0.4(\hat{\beta}_2 + \hat{\beta}_3) = 0$	$p = 0.094$	$p < 0.001$
Symmetry satisfied?	Yes	No
$\hat{\beta}_0 \geq 0$	$p < 0.001$	$p < 0.001$
$\hat{\beta}_0 + \hat{\beta}_1 \geq 0$	$p < 0.001$	$p < 0.001$
$\hat{\beta}_0 + 0.9\hat{\beta}_2 \leq 0$	$p < 0.001$	$p < 0.001$
$\hat{\beta}_0 + \hat{\beta}_1 + 0.9(\hat{\beta}_2 + \hat{\beta}_3) \leq 0$	$p < 0.001$	$p < 0.001$
Non-triviality satisfied?	Yes	Yes



(a) Probabilistic treatment **(b)** Expected Value treatment

Fig. 3 Estimated (logit) driver choice probabilities, with 95% confidence intervals

In the *Prob* treatment we identify the effect of ambiguity aversion, independent of risk aversion, under the assumption that subjects have EU preferences over risk.[36]

Definition 2 A driver is

- ambiguity neutral if $\Pr^{\text{Obs}}_{\text{Prob}}(C|V) = \Pr^{\text{Unobs}}_{\text{Prob}}(C|V)$;
- ambiguity averse if $\Pr^{\text{Obs}}_{\text{Prob}}(C|V) < \Pr^{\text{Unobs}}_{\text{Prob}}(C|V)$;
- ambiguity seeking if $\Pr^{\text{Obs}}_{\text{Prob}}(C|V) > \Pr^{\text{Unobs}}_{\text{Prob}}(C|V)$.

Figure 3 plots the implied predicted choice probabilities as a function of monitoring levels with the data for the $I = \text{Obs}$ case in blue and $I = \text{Unobs}$ in red (with 95% confidence intervals).

[36] The Online Appendix discusses the case where this assumption fails.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Figure 3b shows the driver behavior for the *EV* treatment. When the monitoring probabilities are observed the empirical driver choice probabilities are nearly optimal: choosing C whenever $V(A,B) \geq 0.4$ and choosing A or B otherwise. This is strong support for Hypothesis 2 in the *EV* treatment. Figure 3a shows the driver behavior for the *Prob* treatment. The support for Hypothesis 2 is weaker here than for the *EV* treatment, with choice probabilities being close to deterministic only when monitoring is outside the range [0.3,0.5].

The most remarkable feature of Fig. 3 is that, per definitions 1 and 2, we observe uncertainty and ambiguity seeking when $V(A,B) > 0.4$ and uncertainty and ambiguity aversion when $V(A,B) < 0.4$. Thus, in aggregate, we observe substantial deviations from both uncertainty and ambiguity neutrality. Notice that for high values of $V(A, B)$, and particularly in the *EV* treatment, $\Pr^{Obs}(C|V)$ is close to 1 and, conversely, $\Pr^{Obs}(C|V)$ is close to 0 for low values of $V(A, B)$. This implies a mechanical restriction on the structure of uncertainty and ambiguity preferences that can be observed: when facing high levels of monitoring in the *Unobs* treatment, an uncertainty or ambiguity averse driver will behave in the same manner as an uncertainty or ambiguity neutral driver (selecting C). Conversely, when facing low monitoring in the *Unobs* treatment, an uncertainty or ambiguity seeking driver will behave in the same manner as an uncertainty or ambiguity neutral driver (selecting A or B). This limits the ability to detect uncertainty and ambiguity aversion for high levels of monitoring (and uncertainty and ambiguity seeking for low levels of monitoring). Despite this, there are two candidate explanations for the behavioral patterns displayed in Fig. 3.

First, uncertainty and ambiguity preference could be governed by a prospect-theory-like reflection effect: uncertainty and ambiguity aversion for gains and uncertainty and ambiguity seeking for losses. Note that, when monitoring is high, the payoff from speeding is below the reference level of selecting the safe option and, when monitoring is low, the payoff from speeding is above the safe reference level.[37] Fig. 4 presents this argument graphically, emphasizing that uncertainty and ambiguity preferences are reflected around the reference point of $V(A,B) = 0.4$. While the Figure is suggestive, we did not design our experiment to test for prospect theory preferences and cannot, therefore, rule out alternative explanations.

Second, the observed pattern of behavior is also consistent with between subject preference heterogeneity. In particular, it may be the case that no individual driver exhibits reference dependence, yet aggregate preferences do generate reference dependent behavior. Suppose that some fraction of subjects are uncertainty or ambiguity averse and some fraction are uncertainty or ambiguity seeking. Following the arguments above, the uncertainty or ambiguity averse subjects will primarily determine aggregate deviations from uncertainty or ambiguity neutrality when

[37] The location of the reference point at $V(A,B) = 0.4$ is induced by the expected value of option C remaining constant, and providing a payment that is identical to choosing either road A or B when $m_A = m_B = 0.4$.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

$V(A, B) < 0.4$, and the uncertainty seeking subjects will primarily determine aggregate deviations from uncertainty neutrality when $V(A, B) > 0.4$.[38]

5.2 Officer behavior

Overall, the officer behavior approximates a best response to the driver behavior outlined in the previous section. For rounds where monitoring is observed, and driver behavior is close to the Nash equilibrium behavior, this implies that officer behavior is very close to the Nash equilibrium prediction ($m_a = m_b \in \{0.39, 0.4\}$ in the revenue maximization treatment and $m_a = m_b \in \{0.4, 0.41\}$ in the crash minimization treatment). Figures 5 and 6 show kernel density estimates of the probability density function of the officer's monitoring decision in the *RevMax* and *CrashMin* treatments, respectively. In each figure, the *Prob* treatment is shown in panel (a) and the *EV* treatment in panel (b). The densities are remarkably high around the equilibrium monitoring probabilities – particularly in the revenue maximization *EV* treatment (Fig. 5b).[39]

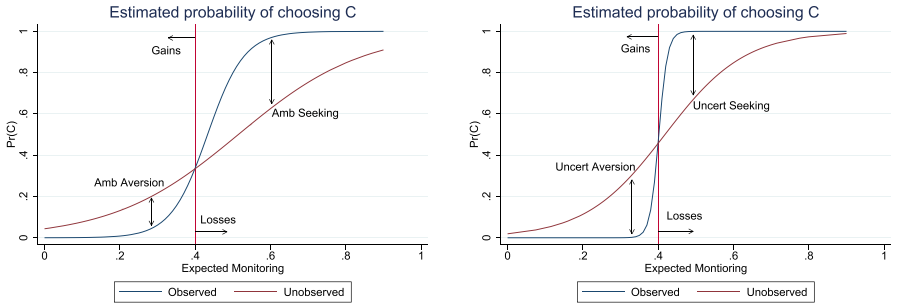
Figures 7 and 8 show kernel density estimates of the probability density function of the officer's monitoring decision for rounds where monitoring was not observed by the driver in the *RevMax* and *CrashMin* treatments, respectively. In each figure, the *Prob* treatment is shown in panel (a) and the *EV* treatment in panel (b). There is substantially more variation in the choice of monitoring levels when monitoring is unobserved, coupled with a substantial shift in mass towards lower monitoring for the revenue maximization case and a shift in mass towards greater monitoring for the crime minimization case. These changes in monitoring behavior, relative to the observed monitoring rounds, are a rational response to the uncertainty preferences of the drivers. In the revenue maximization case, where the equilibrium involves the officer inducing the driver to choose option A or B, we have $V(A, B) < 0.4$ and the reduction in monitoring is a rational response to the driver's uncertainty aversion in the gains domain. In the crash minimization case, where the equilibrium involves the officer inducing the driver to choose option C, we have $V(A, B) > 0.4$ and the increase in monitoring is a rational response to the driver's uncertainty seeking in the loss domain.[40] We also note that in the crash minimization rounds there are some officers who engage in zero monitoring. Setting $m_A = m_B = 0$ can, for example, be justified by an officer who exhibits a large amount of uncertainty aversion with

[38] It may also be that decision making simply becomes noisier in the face of uncertainty (i.e. choice probabilities are biased towards 50% in the unobserved treatments). However, if decision noise were driving the effect, we would expect the red and blue curves in Fig. 4 to intersect at the point where $\Pr(C) = 0.5$ (rather than at $V(A, B) = 0.4$ as predicted by prospect theory or the aggregation of heterogenous uncertainty preferences). For the *EV* treatment the point of intersection coincides with both predictions, but for the *Prob* treatment the point of intersection matches the prediction of prospect theory rather than the decision noise prediction.

[39] There are data points that are not at (0.4, 0.4) in the *RevMax EV* treatment, but they are scattered and are washed out in the figure by the massive density of data at the equilibrium prediction.

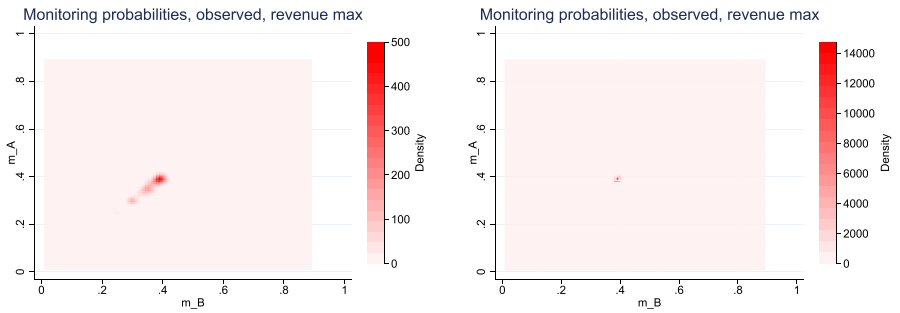
[40] This analysis holds irrespective of whether the driver behavior is caused by prospect theory style preferences or by some other phenomena.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press



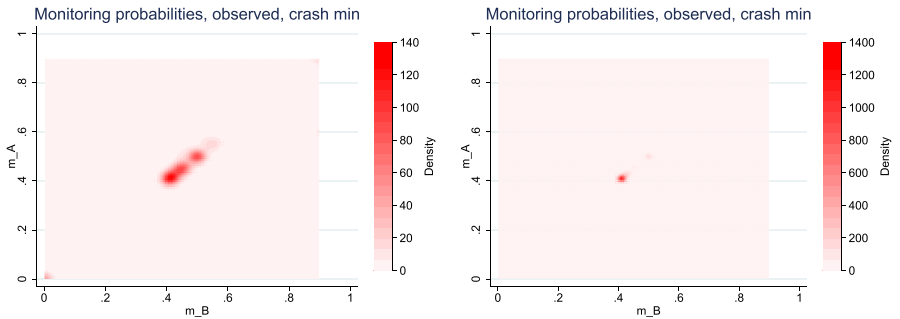
(a) Probabilistic treatment **(b)** Expected Value treatment

Fig. 4 Reference dependent uncertainty preferences



(a) Probabilistic treatment **(b)** Expected Value treatment

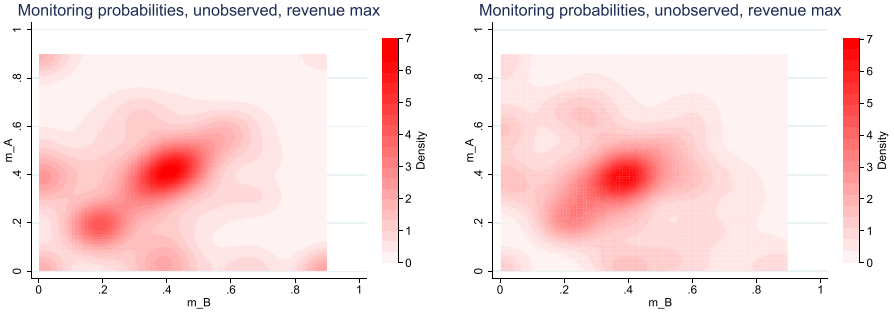
Fig. 5 Officer monitoring probabilities, observed monitoring, revenue maximization treatment



(a) Probabilistic treatment **(b)** Expected Value treatment

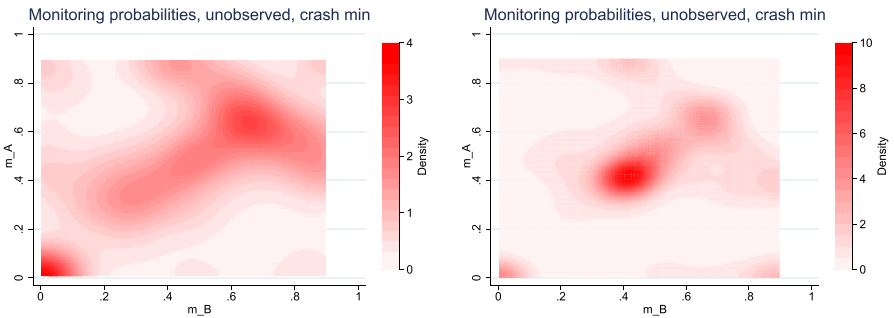
Fig. 6 Officer monitoring probabilities, observed monitoring, crash minimization treatment

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press



(a) Probabilistic treatment **(b)** Expected Value treatment

Fig. 7 Officer monitoring probabilities, unobserved monitoring, revenue maximization treatment



(a) Probabilistic treatment **(b)** Expected Value treatment

Fig. 8 Officer monitoring probabilities, unobserved monitoring, crash minimization treatment

respect to the driver's response to positive monitoring levels because it guarantees a payoff of at least 40 points.

Hypothesis 1 postulates that, when monitoring is observed, behavior should be identical across the *Prob* and *EV* treatments. To test for the equivalence of officer behavior across treatments, we regress the total monitoring level on a treatment indicator for *Prob*, clustering standard errors at the subject level. In aggregate, the level of monitoring across the two treatments is almost indistinguishable: 0.814 in the *Prob* treatment and 0.812 in *EV* treatment ($p = 0.884$). When the *RevMax* and *CrimMin* treatments are isolated, we observe a small and insignificant difference in the *CrimMin* case and a small but statistically significant difference in the *RevMax* case. Monitoring in the *Prob* treatment is further from the equilibrium value of 0.8 in each case.[41] Thus, on average, Hypothesis 1 is supported for officers, but we

[41] For the *CrimMin* case, monitoring levels are 0.892 in the *Prob* treatment and 0.863 in *EV* treatment ($p = 0.402$). For the *RevMax* case, monitoring levels are 0.728 in the *Prob* treatment and 0.764 in *EV* treatment ($p = 0.037$).

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

observe greater variation in monitoring in the *Prob* treatment, a result that is also apparent by comparing panels (a) in figures Figures 5 and 6 to the respective panels (b).

Officers earn higher payoffs in rounds where drivers can observe the monitoring probabilities, under both the crash minimization and revenue maximization payment schemes. In the revenue maximization rounds the driver's uncertainty and ambiguity aversion in the gains domain implies that officers must lower their monitoring levels and therefore lower their fine revenue when monitoring is unobserved. In the crash minimization rounds the driver's uncertainty and ambiguity seeking in the loss domain implies that officers must increase their costly monitoring to prevent crashes when monitoring is unobserved. Given this outcome, Hypothesis 4 predicts that we should expect the officer to reveal their choice of monitoring, under both payment schemes, when given the option. Table 10 demonstrates that the officers do, indeed, reveal their strategy more often than not. Thus, while we do not observe data that is in full concordance with Hypothesis 4, there is some support for the hypothesis.

6 Discussion and conclusion

Citizens often comply with legal rules, but in the instances when they do not, we assign legal agents to identify and sanction non-compliance. However, the public service actors that we charge with the authority of managing such enforcement agencies typically have conflicting incentives. On the one hand, they aim to deter illegal activities and enhance public safety. Alternatively, they might face pressure to raise government revenues by establishing a monitoring environment that encourages the commission of illegal acts (Graham and Makowsky, 2021; Makowsky et al., 2019). Such a monitoring environment could include announcing or shrouding the presence of enforcement activities.

Various auditing, enforcement and monitoring policies have been examined extensively in the literature. The law enforcement literature, for example, typically revolves around policing for profit versus safety. While most of this research attempts to understand the effect of various law enforcement incentives (either endogenous or imposed) on public safety outcomes, observational studies struggle to understand the reason that one law enforcement strategy is more effective than another. Stated differently, the impact of a law enforcement strategy on public safety is a function of the law enforcer's incentives, the enforcement agency's decisions to announce or shroud their enforcement practices and the preferences of citizens. Given that, of these three factors influencing the effectiveness of law enforcement strategies, the law enforcement agency's decision to announce or shroud their enforcement practices is often the only observable behavior, empirical analyses using real-world data cannot fully determine the effectiveness of such law enforcement policies. While Fig. 1 is suggestive of the relationship between fiscal incentives and law enforcement decisions to announce or shroud their strategies, Shaw (2015) notes that "if you wanted to know if your local law enforcement was policing well or unfairly relative to other similar jurisdictions, there is little data available for you to assess that."

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Table 10 Proportion of rounds in which the Officer revealed their monitoring strategy to the Drivers, with 95% confidence intervals in brackets

	Revenue Max	Crash Min
Prob treatment	0.64	0.74
	[0.60, 0.68]	[0.71, 0.77]
EV treatment	0.67	0.74
	[0.63, 0.71]	[0.70, 0.77]

To overcome these issues, we utilize a lab environment to disentangle the impact of each of these three factors that influence the effectiveness of law enforcement strategies. In doing so, we identify how specific environments impact the decision to engage in riskier decisions by citizens. Simultaneously, we are able to identify the conditions that lead an enforcement agent to announce or shroud their intended enforcement practices. The nexus of these decisions yields important insights regarding ongoing debates surrounding issues of public safety. For example, California Vehicle Code 40802 does not permit "unjustified speed limit traps", which are defined as sections of highway with a lower speed limit that is not justified by a traffic survey conducted within the past five years. Hidden speed cameras were utilized in identifying speeding motorists in Arizona, but this was outlawed by the Governor in 2010. In addition, the Revised Code of Washington 46.08.065 prohibits the use of unmarked vehicles by public agencies, with exceptions only being made for Washington state patrol for general undercover or confidential investigative purposes, as approved by the chief of the Washington state patrol. These laws evolved out of concern that many communities and law enforcement agencies voiced about the use of marked or unmarked police cars.[42] From a public safety viewpoint, the tension between using marked vehicles, or not, pertains to whether drivers are more responsive to the presence of a marked vehicle (risk; over time, with marked vehicles, the drivers learn the distribution of monitoring) or the concern that unmarked vehicles are patrolling (ambiguity; drivers cannot learn the distribution of monitoring). Our findings note that an agency concerned with increasing public safety should indeed reveal their enforcement activities (i.e. utilize marked vehicles) to deter ambiguity seeking citizens.

Our laboratory experiment, and theoretical framework, allow us to provide new insight about these ongoing debates. Relying on a laboratory environment enables us to overcome the endogeneity of the enforcement strategy while also being able to discern the ambiguity and risk attitudes of the enforced subjects. The results of our work display two distinct equilibria for law enforcement: (1) revealing enforcement strategies when facing uncertainty averse citizens and revenue-enhancing incentives and (2) revealing enforcement strategies when facing uncertainty seeking citizens and safety-enhancing incentives. We note that these equilibria are consistent with an underlying prospect theory structure of uncertainty preferences among the representative citizen. When facing safety-enhancing incentives officers will, in

[42] See Preusser Research Group (2015) for a more in depth discussion of these trade-offs.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

equilibrium, monitor to such an extent that only uncertainty seeking citizens commit violations. Therefore, the choice to reveal monitoring is made to reduce violations among this subset of the population. On the other hand, when facing revenue-maximization incentives officers will, in equilibrium, monitor to such an extent that the marginal offender is uncertainty averse. Therefore, the choice to reveal monitoring is made to increase violations among this subset of the population.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10683-022-09755-w.

Acknowledgments We would like to thank Glenn Harrison, John List and the seminar audiences at UC Davis and Purdue University as well as the George Mason University School of Law Experimental Law and Economics Workshop, the Australian New Zealand Workshop on Experimental Economics, and the Public Choice Society Meetings for helpful comments. This study was supported by funds from the Krannert School of Management, Purdue University, and received IRB approval from Purdue University. The replication material for the study is available at https://doi.org/10.7910/DVN/VHS9OT.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature, 36*(2), 818–860.

Azrieli, Y., & Teper, R. (2011). Uncertainty aversion and equilibrium existence in games with incomplete information. *Games and Economic Behavior, 73*, 310–317.

Bade, S. (2011). Ambiguous act equilibria. *Games and Economic Behavior, 71*, 246–260.

Baillon, A., Huang, Z., Selim, A., & Wakker, Peter P. (2018). Measuring ambiguity attitudes for all (natural) events. *Econometrica, 86*(5), 1839–1858.

Baillon, A, Halevy, Y, & Li, C. Randomize at your own risk: on the observability of ambiguity aversion. *Econometrica*, Forthcoming.

Baker, T., Harel, A., & Kugler, T. (2004). The virtues of uncertainty in law: An experimental approach. *Iowa Law Review, 89*, 871.

Banerjee, A., Duflo, E., Keniston, D., & Singh, N. (September 2019). The efficient deployment of police resources: Theory and new evidence from a randomized drunk driving crackdown in India. Working Paper 26224, National Bureau of Economic Research. URL http://www.nber.org/papers/w26224.

Bentham, J. (1843). *The Works of Jeremy Bentham, Volume 4 (Panopticon, Constitution, Colonies, Codification)*.

Bohnet, I., & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior and Organization, 55*, 467–484.

Bureau of Justice Statistics. (2013). Law enforcement management and administrative statistics. Technical report, United States Department of Justice, Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Calford, E. M. (2020). Uncertainty aversion in game theory: Experimental evidence. *Journal of Economic Behavior and Organization, 176*, 720–734.

Camerer, Colin F., & Risto Karjalainen (1994). Ambiguity aversion and non-additive beliefs in non-cooperative games: Experimental evidence. In M. Bertrand & M. J. Machina (Eds.), *Models and experiments in risk and rationality* (pp. 325–358). Netherlands: Springer.

Carpenter, D, Knepper, L, Erickson, A, & McDonald, J. (November 2015). *Policing for profit: The abuse of civil asset forfeiture. Technical report, Institute for Justice*.

Cason, T, Friesen, L, & Gangadharan, L. (2016). Regulatory performance of audit tournaments and compliance observability. *European Economic Review, 85*(C), 288–306. https://EconPapers.repec.org/RePEc:eee:eecrev:v:85:y:2016:i:c:p:288-306.

Cason, Timothy N., & Gangadharan, L. (2006). An experimental study of compliance and leverage in auditing and regulatory enforcement. *Economic Inquiry, 44*(2), 352–366. https://doi.org/10.1093/ei/cbj019. URL https://onlinelibrary.wiley.com/doi/abs/10.1093/ei/cbj019

Cason, Timothy N., Friesen, Lana, & Gangadharan, Lata (2021). Complying with environmental regulations. In A. Chaudhuri (Ed.), *A research agenda in experimental economics* (pp. 69–92). UK: Edward Elgar.

Charness, G., Gneezy, U., & Imas, A. (2013). Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior and Organization, 87*, 43–51.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree - an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance, 9*, 88–97.

Clark, J, Friesen, L, & Muller, A. (2004). The good, the bad, and the regulator: An experimental test of two conditional audit schemes. *Economic Inquiry, 42*(1), 69–87. https://doi.org/10.1093/ei/cbh045. URL https://onlinelibrary.wiley.com/doi/abs/10.1093/ei/cbh045

DeAngelo, G., & Charness, G. (2012). Deterrence, expected cost, uncertainty and voting: Experimental evidence. *Journal of Risk and Uncertainty, 44*(1), 73–100.

DeAngelo, G., & Gee, L. K. (2020). Peers or police?: The effect of choice and type of monitoring in the provision of public goods. *Games and Economic Behavior, 123*, 210–227.

DeAngelo, G., & Hansen, B. (2014). Life and death in the fast lane: Police enforcement and traffic fatalities. *American Economic Journal: Economic Policy, 6*(2), 231–57. https://doi.org/10.1257/pol.6.2.231http://www.aeaweb.org/articles?id=10.1257/pol.6.2.231

DeAngelo, G, & Owens, EG. (2017). Learning the ropes: General experience, task-Specific experience, and the output of police officers. *Journal of Economic Behavior & Organization, 142*(C), 368–377. https://doi.org/10.1016/j.jebo.2017.08.00https://ideas.repec.org/a/eee/jeborg/v142y2017icp368-377.html.

DeAngelo, G, Harvey, A, & Mungan, M. (2019). Policing for profit: The political economy of law enforcement. *Working Paper*.

Dow, J., & Werlang, S. R. (1994). Nash equilibrium under Knightian uncertainty: Breaking down backward induction. *Journal of Economic Theory, 64*, 305–324.

Eeckhout, J, Persico, N, & Todd, PE. (June 2010). A theory of optimal random crackdowns. *American Economic Review, 100*(3), 1104–35. https://doi.org/10.1257/aer.100.3.1104. URL http://www.aeaweb.org/articles?id=10.1257/aer.100.3.1104

Eichberger, J., & Kelsey, D. (2000). Non-additive beliefs and strategic equilibria. *Games and Economic Behavior, 30*, 183–215.

Eichberger, J., & Kelsey, D. (2014). Optimism and pessimism in games. *International Economic Review, 55*(2), 483–505.

Eichberger, J., Kelsey, D., & Schipper, B. C. (2008). Granny versus game theorist: Ambiguity in experimental games. *Theory and Decision, 64*, 333–362.

Farago, K., Kiss, O., & Boros, J. (2008). Risk-taking in entrepreneurs, compared to criminals and students: The role of uncertainty and stakes. *The Journal of Socio-Economics, 37*(6), 2231–2241.

Fox, Craig R., & Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics, 51*, 585–603.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics, 18*, 141–153.

Gilpatric, SM., Vossler, CA., & McKee, M. (2011). Regulatory enforcement with competitive endogenous audit mechanisms. *The RAND Journal of Economics, 42*(2), 292–312. ISSN 07416261. URL http://www.jstor.org/stable/23046800.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Gilpatric, S. M., Vossler, C. A., & Liu, L. (2015). Using competition to stimulate regulatory compliance: A tournament-based dynamic targeting mechanism. *Journal of Economic Behavior and Organization, 119*, 182–196.

Goldstein, R., Sances, M. W., & You, H.-Y. (2020). Exploitative revenues, law enforcement, and the quality of government service. *Urban Affairs Review, 56*(1), 5–31. https://doi.org/10.1177/1078087418791775

Graham, S. R., & Makowsky, M. D. (2021). Local government dependence on criminal justice revenue and emerging constraints. *Annual Review of Criminology, 4*(1), 311–330. https://doi.org/10.1146/annurev-criminol-061020-021824

Grant, S., Meneghel, I., & Tourky, R. (2016). Savage games. *Theoretical Economics, 11*, 641–682.

Grechenig, K, Nicklisch, A, & Thöni, C. (2010). Punishment despite reasonable doubt - a public goods experiment with sanctions under uncertainty. *Journal of Empirical Legal Studies, 7*(4), 847–867. https://doi.org/10.1111/j.1740-1461.2010.01197.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1740-1461.2010.01197.x

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association, 1*(1), 114–125.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica, 75*(2), 503–536.

Hanany, E., Klibanoff, P., & Mukerji, S. (2020). Incomplete information games with ambiguity averse players. *American Economic Journal: Microeconomics, 12*(2), 135–187.

Harel, A., & Segal, U. (1999). Criminal law and behavioral law and economics: observations on the neglected role of uncertainty in deterring crime. *American Law and Economic Review, 1*(1), 276–312.

Holzmeister, F., & Stefan, M. (2021). The risk elicitation puzzle revisited: Across-methods (in)consistency? *Experimental Economics, 24*, 593–616.

Ivanov, A. (2011). Attitudes to ambiguity in one-shot normal-form games: An experimental study. *Games and Economic Behavior, 71*, 366–394.

Kantor, S, Kitchens, C, & Pawlowski, S. (September 2017). Civil asset forfeiture, crime, and police incentives: Evidence from the comprehensive crime control act of 1984. Working Paper 23873, National Bureau of Economic Research. URL http://www.nber.org/papers/w23873.

Kelsey, D., & le Roux, S. (2015). An experimental study on the effect of ambiguity in a coordination game. *Theory and Decision, 79*(4), 667–688.

Kelsey, D., & le Roux, S. (2017). Dragon slaying with ambiguity: Theory and experiments. *Journal of Public Economic Theory, 19*(1), 451.

Kelsey, D., & le Roux, S. (2018). Strategic ambiguity and decision making: An experimental study. *Theory and Decision, 5*, 12.

Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica, 73*(6), 1849–1892.

Lacey, JH., Jones, RK., & Smith, RG. (1999). Evaluation of checkpoint tennessee: Tennessee's statewide sobriety checkpoint program. Technical report, U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC 20590.

Lazear, E. (2006). Speeding, terrorism, and teaching to the test. *Quarterly Journal of Economics, 121*, 1029–1061. https://doi.org/10.1162/qjec.121.3.1029

Li, C., Turmunkh, U., & Wakker, P. P. (2019). Trust as a decision under ambiguity. *Experimental Economics, 22*, 51–75.

Lo, K. C. (1996). Equilibrium in beliefs under uncertainty. *Journal of Economic Theory, 71*, 443–484.

Lo, K. C. (2009). Correlated Nash Equilibrium. *Journal of Economic Theory, 144*, 722–743.

Lochner, L. (2007). Individual perceptions of the criminal justice system. *American Economic Review, 97*(1), 444–460.

Machina, M. J., & Siniscalchi, M. (2013). Ambiguity and ambiguity aversion. In M. J. Machina & W. K. Viscusi (Eds.), *Handbook of the economics of risk and uncertainty.* UK: Newnes.

Makowsky, Michael D., & Stratmann, T. (March 2009). Political economy at any speed: What determines traffic citations? *American Economic Review, 99*(1), 509–27. https://doi.org/10.1257/aer.99.1.509. URL http://www.aeaweb.org/articles?id=10.1257/aer.99.1.509

Makowsky, Michael D., & Stratmann, T. (2011). More tickets, fewer accidents: How cash-strapped towns make for safer roads. *The Journal of Law & Economics, 54*(4), 863–888. ISSN 00222186, 15375285. URL http://www.jstor.org/stable/10.1086/659260.

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press

Makowsky, M. D., Stratmann, T., & Tabarrok, A. (2019). To serve and collect: The fiscal and racial determinants of law enforcement. *The Journal of Legal Studies, 48*(1), 189–216. https://doi.org/10.1086/700589

Maarinacci, Massimo. (2000). Ambiguous games. *Games and Economic Behavior, 31*, 191–219.

Mungan, M., & Klick, J. (2014). Forfeiture of illegal gains, attempts, and implied risk preferences. *Journal of Legal Studies, 43*(1), 137–153.

Olken, BA., & Barron, P. (June 2009). The Simple Economics of Extortion: Evidence from Trucking in Aceh. *Journal of Political Economy, 117*(3), 417–452. URL https://ideas.repec.org/a/ucp/jpolec/v117y2009i3p417-452.html.

Preusser Research Group, Inc. (January 2015). Process overview of the high-visibility enforcement programs targeting handheld device users in California and Delaware. Technical report, National Highway Traffic Safety Administration.

Riedel, F., & Sass, L. (2014). Ellsberg games. *Theory and Decision, 76*(4), 469–509.

Salmon, Timothy C., & Shniderman, A. (2019). Ambiguity in criminal punishment. *Journal of Economic Behavior & Organization, 163,* 361–376. ISSN 0167-2681. https://doi.org/10.1016/j.jebo.2019.05.014. URL https://www.sciencedirect.com/science/article/pii/S0167268119301532.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica, 57*(3), 571–587.

Shaw, Emily. (November 2015). Help uncover the nation's dependence on local court fees and fines. Online (Sunshine Foundation). URL https://sunlightfoundation.com/2015/11/30/help-uncover-the-nations-dependence-on-local-court-fees-and-fines/.

Slemrod, J. (December 2019). Tax compliance and enforcement. *Journal of Economic Literature, 57*(4), 904–54. https://doi.org/10.1257/jel.20181437. URL https://www.aeaweb.org/articles?id=10.1257/jel.20181437

USDOJ. *Community Relations Services Toolkit for Policing*. https://www.justice.gov/crs/file/836401/download, n.d.

Vollaard, B., & Hamed, J. (2012). Why the police have an effect on violent crime after all: Evidence from the british crime survey. *The Journal of Law and Economics, 55*(4), 901–924. https://doi.org/10.1086/666614

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

https://doi.org/10.1007/s10683-022-09755-w Published online by Cambridge University Press