Information and Policing

Shota Ichihashi*

July 17, 2024

Abstract

Agents decide whether to commit crimes based on their heterogeneous returns from crimes, or their types. The police possess some information about these types. The police search agents, without commitment, in order to detect crimes subject to a search capacity constraint. The deterrent effect of policing is lost when the police have full information about types. The information structure that minimizes the crime rate is informative about the agents' types but uninformative about their equilibrium behavior. The result extends to the case in which the police have partial commitment power or when they endogenously choose search capacity at a cost.

*Queen's University, Department of Economics. Email: shotaichihashi@gmail.com. I thank Moritz Meyer-ter-Vehn for providing valuable comments on an early draft of this paper. For valuable suggestions and comments, I thank Yeon-Koo Che, Christopher Cotton, Navin Kartik, Qingmin Liu, Ruslan Momot, Alex Smolin, and the participants of seminars and talks at Osaka University, Columbia Business School, Toulouse School of Economics, University of Michigan (Ross School of Business), and Queen's University.

1 Introduction

Law enforcement agencies are increasingly relying on data and algorithms to predict crimes (Perry, 2013; Brayne, 2020). They use a variety of data sources, such as criminal records, social media posts, financial records, and local environmental information. Private vendors, such as Palantir and PredPol, also offer predictive algorithms to police departments. This trend, which is driven by the pursuit of more effective law enforcement, has raised a number of concerns. As a case in point, the EU's proposed regulation "Artificial Intelligence Act" classifies a certain use of artificial intelligence to predict crime as a prohibited practice.[1]

Motivated by recent discussions, I study how information available to law enforcement affects its ability to deter crimes, when the law enforcer cannot commit to how to use this information. I examine this question from the perspective of information design.

The model consists of a unit mass of agents and a law enforcement, called the "police." The agents face heterogeneous returns on crimes. At the outset, the police observe information about the type of each agent according to an exogenous signal structure. Each agent decides whether to commit a crime, and *simultaneously*, the police search agents subject to a search capacity constraint.

The simultaneous-move assumption implies that the police cannot commit to a predetermined search strategy. This assumption captures a situation in which agents decide whether to commit crimes—such as illegal parking, tax fraud, or drug trafficking—and then individual officers or auditors search agents to detect these crimes without directly observing the agents' actions. Under this assumption, although the police may be concerned about the overall crime rate (i.e., the number of agents committing crimes), in equilibrium, the police will focus solely on uncovering crimes instead of deterring them, taking the crime rate as given.

I begin with a preliminary observation that information restriction is necessary for crime deterrence: If the police have full information about the agents' types, every agent commits a crime with probability 1 in any equilibrium. The fully informed police will allocate search resources only to the agents who are most likely to commit crimes. This search strategy fails

[1] See https://www.europarl.europa.eu/news/en/press-room/20230609IPR96212/meps-ready-to-negotiate-first-ever-rules-for-safe-and-transparent-ai.

to deter any crime: Each agent is either not searched at all, and thus commits a crime, or is searched with a positive probability but still commits a crime as predicted by the police.

The main result is the characterization of the signal structure that minimizes a crime rate. To derive this signal structure, I first solve a relaxed problem, in which the agents do not observe their types, and we choose a joint signal structure for the police and the agents to minimize a crime rate. The relaxed problem has a solution in which the police receive no information and randomly search agents, and each agent learns only whether their type exceeds some cutoff. The relaxed problem identifies a lower bound of possible crime rates that can arise in the original problem. I then turn to the original problem—where the agents observe their types—and construct a signal structure for the police that attains the same outcome as in the relaxed problem.

The key property of this crime-minimizing signal structure is that it is partially informative about the agents' types, yet reveals no information about the agents' equilibrium behavior. By not informing the police about the agents' behavior, this signal structure minimizes the distortion seen in the full information benchmark, where information erodes the deterrence effect of searches. This result suggests that predicting crimes could be counterproductive when law enforcement lacks commitment power regarding how to act based on the prediction.

In practice, a law enforcement agency has some level of commitment power in choosing a search strategy. In Section 3.3, we show that the crime-minimizing signal continues to be relevant in such a situation. Specifically, suppose that a designer (e.g., a police organization) possesses some noisy signal of agents' types and can commit to a search strategy that is measurable with respect to the signal. The designer can also provide additional information to the police (e.g., individual officers), who, without commitment, will use this information to further allocate the search capacity set by the designer. I show that the designer's optimal information policy features the crime-minimizing signal structures of the baseline model.

The above results assume that the police face an exogenously given search capacity. In Section 4, we allow the police to endogenously choose a total search capacity at a cost: For example, an individual officer might exert different levels of search effort depending on the available information. Under a certain condition, the crime-minimizing signal structure of the baseline model continues to minimize crime when search capacity is endogenous. I also provide an example in which this result fails, and a crime-minimizing signal structure enables the police to identify a fraction of the agents who will abstain from crimes in equilibrium.

The contribution of this paper is to offer a new application of information design, specifically, the issue of what information law enforcement should be endowed with. By characterizing the information structure that minimizes crime, the paper shows how the attempt to predict crimes could backfire under law enforcement's lack of commitment to the use of information. Theoretically, the paper studies an information design problem with a rich space of players and strategies and a constraint on feasible information structures. It illustrates how a solution technique based on a relaxed problem facilitates the characterization of the optimal information structure.

Related work. The paper relates to the literature on Bayesian persuasion and information design (see Kamenica (2019) and Bergemann and Morris (2019) for surveys). Several papers, such as Lazear (2006), Eeckhout, Persico, and Todd (2010), and Hernández and Neeman (2022), study the intersection of Bayesian persuasion and law enforcement. These papers study how to disclose information to criminals or players to deter their socially undesirable actions. In contrast, this paper studies what information a law enforcer should have about such players. Methodologically, it addresses an information design problem with a continuum of players, states, and actions (for the police), which seems to be less understood than a single-player Bayesian persuasion problem (see Smolin and Yamashita (2022) for a discussion). The information design literature provides conditions under which an optimal signal takes a tractable form, such as monotone partitional signals, censorship policies, and nested intervals (e.g., Guo and Shmaya 2019; Dworczak and Martini 2019; Kolotilin, Mylovanov, and Zapechelnyuk 2022). The crime-minimizing signal structure does not belong to these classes of signals. Finally, the relaxed problems examined in this paper relate to Bayesian persuasion with moral hazard, in which the state distribution is endogenously determined by a player's action (Rodina, 2017; Boleslavsky and Kim, 2018; Zapechelnyuk, 2020; Hörner and Lambert, 2021).

This paper also relates to the economic literature on crime and policing, which starts from

Becker (1968). The question of what information about agents should or should not be used for policing is discussed in the context of racial profiling (Knowles, Persico, and Todd 2001; Persico and Todd 2005; Bjerk 2007; Persico 2009). In terms of the timing and payoffs of the game, my paper builds closely on Persico (2002), who studies whether requiring the police to adopt a fairer search strategy reduces crime. Instead of constraining police behavior, I restrict information available to police and study what information renders policing effective. A model of predictive enforcement is also studied by Che, Kim, and Mierendorff (2023), who consider a bandit model that captures the endogenous generation and use of information for law enforcement. To focus on the role of information in policing, the model abstracts away from other important considerations, such as the design of judicial systems, richer responses by potential criminals and victims, as well as the "fairness" of predictive algorithms (e.g., Curry and Klumpp 2009; Cotton and Li 2015; Jung, Kannan, Lee, Pai, Roth, and Vohra 2020; Vasquez 2022; Liang, Lu, and Mu 2022).

2 Model

The model consists of police and a unit mass of agents, indexed by $i \in [0,1]$. Each agent i has some underlying returns on crime, or *type*, $x_i \in [0,1]$. Each agent observes their type. We may interpret an agent's type as reflecting the individual characteristics that affect their returns on crime (e.g., legal earning opportunities) or crime opportunities specific to certain locations or times. The agents' types are independently and identically drawn from distribution function $F \in \Delta[0,1]$, which has a positive density f and is commonly known.[2] We use $\mathbb{E}_F[\cdot]$ for the expectation operator under F. We also use $F(\cdot|\tilde{x} \leq c)$ for the conditional distribution of F on $[0,c]$ and $\mathbb{E}_F[\cdot|\tilde{x} \leq c]$ for the corresponding expectation operator.

The police learn about each agent's type according to a *signal structure* (S,π), which consists of a set S of signals and a collection $\pi = \{\pi(\cdot|x)\}_{x\in[0,1]}$ of conditional distributions $\pi(\cdot|x) \in \Delta S$ over signals for each type x. For each agent $i \in [0,1]$, the police observe a signal $s_i \in S$ drawn from distribution $\pi(\cdot|x_i)$. Conditional on types, signals are independent across

[2] We write ΔX for the set of all probability distributions on a set X. See, e.g., Sun (2006) for a formal treatment of a continuum of independent random variables and the corresponding law of large numbers.

agents. The signal structure is exogenous and commonly known, but only the police observe realized signals.

Given the signal structure, the police and the agents play the following simultaneous-move game: Each agent decides whether or not to commit a crime, and simultaneously, the police allocate search resources across agents. Specifically, the police choose a *search strategy* $p : S \to [0,1]$, where $p(s)$ is the probability of searching agents with signal $s \in S$. The police have a measure $\overline{P} \in (0,1)$ of searches to allocate; thus, the police can choose a search strategy $p(\cdot)$ if and only if the total mass of searches does not exceed $\overline{P}$, i.e.,

$$\int_0^1 \int_S p(s)\pi(ds|x)F(dx) \leq \overline{P}. \tag{1}$$

The payoff of an agent from committing a crime is $x - \rho$, where x is the agent's type and $\rho \in [0,1]$ is the search probability that the agent faces. The payoff of not committing a crime is 0. This payoff specification is equivalent to the following richer setup: The payoff of committing a crime is $U(y,\rho)$, which strictly increases in an agent's type y; strictly decreases in search probability ρ; and has a threshold search probability $\hat{p}(y)$ that solves $U(y,\hat{p}(y)) = 0$ for each type y (without loss, the payoff of not committing a crime is normalized to 0). An agent will commit a crime if $\hat{p}(y) > \rho$. But once we redefine the agent's type as $x = \hat{p}(y)$, this richer setup and our original setup lead to the same set of best responses by agents under any search strategy. This setup subsumes $U(y,\rho) = (1-\rho)y - L\rho$, i.e., an agent enjoys the returns on crime if they are not searched, but incurs a loss of L if they are searched.

To describe and discuss the police's payoff, we define the *mass of successful searches* as the mass of agents who commit crimes and are searched by the police. Also, we define the *crime rate* as the mass of agents who commit crimes.

As for the police's payoff, we only assume that it is a function of the mass of successful searches and the crime rate, and strictly increases in the mass of successful searches. We do not impose any assumption on how the police prioritize crime detection over crime reduction. However, to simplify the exposition, we hereafter adopt a particular specification in which the police's payoff is equal to the mass of successful searches. Moving from the general specification to this simple one does not change the set of equilibria for the following reason:

Under the simultaneous-move assumption, the police's behavior cannot directly affect the agents' actions and thus the crime rate. Therefore, even if the police's payoff depends on the crime rate, in equilibrium, the police act *as if* their payoffs depend only on the mass of successful searches.

The solution concept is Bayesian Nash equilibrium, which we refer to as *equilibrium*. To minimize possible case classifications, we assume that the primitives—i.e., the distribution of returns on crime and the police's search capacity—satisfy the following:

Assumption 1. The primitives, F and $\overline{P}$, satisfy

$$\overline{P} < \int_0^1 xF(dx). \tag{2}$$

The assumption means that the police do not have enough resources to fully deter crime: To attain a crime rate of 0, the police would have to search each agent i with a probability of at least x_i, but the inequality $\overline{P} < \int_0^1 xF(dx)$ implies that such a search strategy violates the search capacity constraint. As a result, a positive mass of agents, facing search probabilities strictly below x_i, prefer to commit crimes.

Our main focus is on a signal structure that minimizes the equilibrium crime rate.

Definition 1. A *crime-minimizing signal structure* is a signal structure that has an equilibrium with the lowest crime rate across all signal structures and equilibria. The corresponding equilibrium is called a *crime-minimizing equilibrium*.

The police's equilibrium search strategy is different from a crime-minimizing strategy because, even though the police might care about reducing crime, they act to maximize the number of successful searches due to the simultaneous-move assumption. As a result, providing the police with more information does not necessarily reduce crime in equilibrium. The crime-minimizing signal structure restricts the police's information so that the equilibrium search strategy exhibits maximum deterrence. We may view the crime-minimizing signal structure as the choice of a social planner who seeks to reduce crimes and regulates the information available to the police.

2.1 Discussion on the Timing Assumption

The results of this paper hinge on the assumption that the police and agents move simultaneously. One implication of this assumption is that the police cannot commit to a search strategy in advance. This section provides a partial justification for this assumption and discuss how it can be relaxed.

First, we may view a signal structure as a predictive algorithm used within a law enforcement agency. In such a context, predictions generated by an algorithm (i.e., realized signals) would be invisible to the public and too complex to describe in advance, making it difficult for the police to commit to a predetermined search strategy.

Second, in line with the literature on decentralized law enforcement, such as Persico (2002) and Porto et al. (2013), we may view the "police" not as an organization but as individual officers and auditors. The simultaneous-move assumption then arises from the idea that the action of an individual officer does not directly influence the decisions of potential criminals.

Meanwhile, in practice, law enforcement likely has some commitment power regarding the allocation of search resources. To accommodate such a situation, we study the case of partial commitment in Section 3.3 and show that the crime-minimizing signal structure continues to be part of the crime-minimizing policy.

2.2 Preliminary Analysis: The Fully Informed Police

To illustrate that restricting the police's information is necessary for crime deterrence, we begin with the analysis of the fully informed police. Specifically, suppose that the signal structure is such that for every type $x \in [0,1]$, $\pi(\cdot|x)$ places probability 1 on $s = x$.

Theorem 0. *Suppose that the police have full information. An equilibrium exists, and in any equilibrium, almost every agent commits a crime with probability* 1.

Proof. First, we construct an equilibrium. Suppose that the police adopt a search strategy $p^*(x) = \frac{x\overline{P}}{\mathbb{E}_F[\tilde{x}]}$. This search strategy is feasible because it induces the total search capacity of $\int_0^1 p^*(x)F(dx) = \overline{P}$. Assumption 1 implies that $\overline{P} < \mathbb{E}_F[\tilde{x}]$, which means that $p^*(x) < x$ for all $x \in [0,1]$. As a result, every agent commits a crime with probability 1. The police then find it optimal to choose any search strategy that exhausts search capacity $\overline{P}$, including $p^*(\cdot)$. Thus we obtain an equilibrium.

Second, take any equilibrium. The police cannot choose a search strategy such that $p(x) \geq x$ for (almost) every x, because by Assumption 1, it violates the search capacity constraint. Thus, the set $X \triangleq \{x \in [0,1] : x > p(x)\}$ has a positive mass, and any type in X commits a crime with probability 1. Let Y be the set of types that commit crimes with probability strictly below 1. If Y has a positive mass, the police must be allocating a positive mass of searches to types in Y, because otherwise they would commit crimes. We then obtain a contradiction, because the police could increase the mass of successful searches by shifting search probabilities from types in Y to X.[3] Thus, the set Y has measure zero, i.e., almost every agent commits a crime with probability 1. $\square$

The intuition for Theorem 0, which appears in the existing work such as Persico (2002), is that the police's ability to predict crimes—combined with their incentive to uncover crimes and a lack of commitment in search strategy—eliminates the deterrence effect of policing.[4]

In contrast to the equilibrium crime rate, the police's equilibrium strategy is not unique. For example, suppose that $F = U[0,1]$ (i.e., the uniform distribution on $[0,1]$) and $\overline{P} = \frac{1}{4}$. In the proof of Theorem 0, we constructed an equilibrium in which the police adopt search strategy $p(x) = \frac{x}{2}$ for every $x \in [0,1]$. The following is another equilibrium: The police search any type $x \leq \frac{1}{\sqrt{2}}$ with probability x and never search any type $x > \frac{1}{\sqrt{2}}$. In this equilibrium, each type $x \leq \frac{1}{\sqrt{2}}$ is indifferent between committing a crime and not, yet breaks ties for committing a crime. Indeed, if types below $\frac{1}{\sqrt{2}}$ did not commit crimes, the police would profitably deviate and search types above $\frac{1}{\sqrt{2}}$ instead. In general, a strategy profile is an equilibrium if and only if (i) the search capacity $\overline{P}$ is allocated across agents in a way that almost every agent weakly prefers to commit a crime, and (ii) almost every agent commits

[3] Note that the police's deviation would be profitable even if the police's payoffs depend on a crime rate, because the deviation does not change the agents' actions (and thus the crime rate) due to the simultaneous-move assumption.

[4] More specifically, Persico (2002) considers two groups of agents with group-specific type distributions. Persico (2002) provides sufficient conditions—in terms of these type distributions and the police's search capacity—under which allowing the police to tailor search rates to groups will increase the equilibrium crime rate relative to constraining the police to adopt a uniform search rate. See also Goldman and Pearl (1976) for a related result.

a crime with probability 1.

3 Crime-Minimizing Signal Structure

We now turn to our main focus, the crime-minimizing signal structure. The analysis consists of two steps. First, we study a relaxed problem, in which the agents do not directly observe their types, and we design a joint information structure for the police and the agents in order to minimize the equilibrium crime rate. Because the incentive constraints of the agents are relaxed, the resulting crime rate becomes a lower bound of the possible crime rates in the original problem. Second, we turn to the original setup and construct a signal structure for the police that attains this lower bound in an equilibrium.

3.1 Relaxed Problem

To define the relaxed problem, we modify the model as follows: The agents know the type distribution F but do not directly observe their types. Instead, the information of the agents and the police is determined by a *joint signal structure*, (S_P, S_A, π). Here, S_P and S_A are the sets of signals for the police and agents, respectively, and $\pi = \{\pi(\cdot|x)\}_{x\in[0,1]}$ is the collection of conditional probability distributions on $S_P \times S_A$ for each type. If agent i has type x_i, the police observe s_i^P and agent i observes s_i^A, where $(s_i^P, s_i^A) \sim \pi(\cdot|x_i)$. The rest of the game remains the same: Each agent i observes s_i^A and decides whether to commit a crime, and simultaneously, the police choose a search strategy to maximize the mass of successful searches. The following result characterizes a crime-minimizing joint signal structure.

Lemma 1. *In the relaxed problem, the following joint signal structure minimizes a crime rate: The police learn no information, e.g., $S_P = \{\phi\}$, and each agent learns whether their type exceeds cutoff $\hat{c} \in (0,1)$ that uniquely solves*

$$\mathbb{E}_F[\tilde{x}|\tilde{x} \le \hat{c}] = \overline{P}. \tag{3}$$

In equilibrium, the police search every agent with probability $\overline{P}$, and each agent commits a crime if and only if their type exceeds $\hat{c}$.

Proof. Take any joint signal structure and any equilibrium. The proof consists of three steps. First, by the revelation principle of information design (e.g., Bergemann and Morris 2019), we can replace the agents' signals with action recommendations.[5] Thus, we set $S_A = \{crime, not\}$ and assume that in equilibrium, each agent commits a crime after observing signal *crime* and not after observing signal *not*.

Second, we replace the police's signal with an uninformative signal, such as $S_P = \{\phi\}$. The police take the agents' strategies as given and allocate search capacity $\overline{P}$ to maximize successful searches. Thus, compared to the original joint signal structure, replacing the police's signal with $S_P = \{\phi\}$ reduces the mass of successful searches. In other words, a larger fraction of search capacity $\overline{P}$ goes to the agents who observe signal *not* and a lower fraction goes to the agents who observe signal *crime*. As a result, the obedience constraints of the agents get relaxed, ensuring that the agents continue to follow action recommendations.[6]

The relaxed problem now reduces to Bayesian persuasion: An agent receives a payoff of $x - \overline{P}$ from committing a crime, and we disclose information about type $x \sim F$, in the form of action recommendations, to the agent in order to minimize the probability of committing a crime. As shown in Kamenica and Gentzkow (2011), the solution is to disclose whether type x exceeds a cutoff $\hat{c}$ defined by $\mathbb{E}_F[\tilde{x}|\tilde{x} \leq \hat{c}] = \overline{P}$, which deters the types below $\hat{c}$ from committing a crime. Type distribution has a density and satisfies Assumption 1, so the cutoff $\hat{c} \in (0, 1)$ exists and is unique. □

Lemma 1 implies that in our original setup, the minimum crime rate is attained if all types below $\hat{c}$ abstain from crimes. This outcome cannot arise if the agents observe their types and the police have no information, because the resulting random search induces types between $\overline{P}$ and $\hat{c}$ to commit crimes. However, the next section shows that we can provide the police with partial information to attain the same crime rate.

---

[5] Formally, take any joint signal structure and equilibrium. Let $a_i(y)$ be the equilibrium probability that agent i takes action $a \in \{crime, not\}$ after observing signal $y \in S_A$; $\pi_{AP}(\cdot|x, S)$ be the conditional distribution of an agent's signal when their type is x and the police's signal is in S; and $\pi_P(\cdot|x)$ be the conditional distribution of the police's signal given type x. We define the new signal structure as $\hat{\pi}(S \times \{a\}|x) = \pi_P(S|x) \int_0^1 \int_{S_A} a_i(y)\pi_{AP}(dy|x, S)di$ for each $S \subset S_P$ and action recommendation $a \in \{crime, not\}$.

[6] The obedience constraint for signal *crime* (resp. signal *not*) is that an agent's expected payoff from committing a crime conditional on signal *crime* is non-negative (resp. non-positive). The obedience constraints refer to the constraints for both signals.

3.2 Characterizing the Crime-Minimizing Signal Structure

We now turn to the original problem, in which the agents observe their types. First, we define a class of signal structures:

Definition 2. For each $c \in [0, 1]$, the *truth-or-noise signal structure with cutoff* c, denoted by (S_c, π_c), is the following signal structure: The signal space S_c is $[0, c]$; for each $x \leq c$, distribution $\pi_c(\cdot|x)$ draws $s = x$ with probability 1; and for any $x > c$, distribution $\pi_c(\cdot|x)$ is independent of x and equals $F(\cdot|\tilde{x} \leq c)$.

To understand the truth-or-noise signal structures, consider the police's posterior belief on an agent's type induced by a signal drawn from (S_c, π_c) (see Figure 1). The signal can be the "truth" (i.e., an agent's type is below c and the signal is equal to their true type) or a "noise" (i.e., an agent's type is above c and the signal is a realization of a random draw from $F(\cdot|\tilde{x} \leq c)$). As a result, the posterior belief places positive probabilities both on types below and above c. Specifically, the posterior induced by signal $s \in [0, c]$ contains a point mass $F(c)$ on type s and a mass $1 - F(c)$ of types distributed according to $F(\cdot|\tilde{x} > c)$.[7] The posterior beliefs (indexed by their point masses) are distributed according to $F(\cdot|\tilde{x} \leq c)$ and average to prior distribution F.

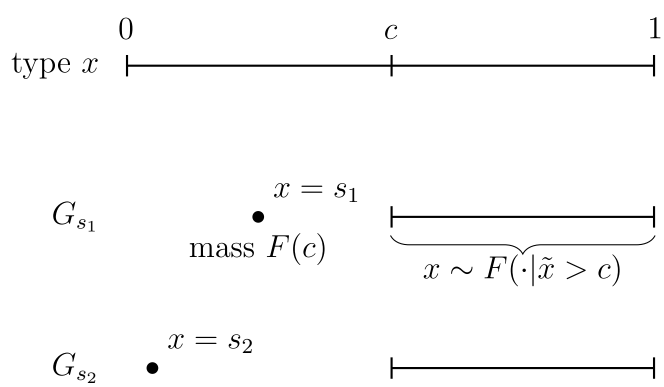

Figure 1: Posterior distribution G_s of types conditional on signal $s = s_1, s_2$ under (S_c, π_c).

[7]To see this, let $f(s|\tilde{x} \leq c)$ denote the conditional density associated with $F(s|\tilde{x} \leq c)$ evaluated at $x = s$. Roughly, the "probability" with which signal s is realized is $f(s|\tilde{x} \leq c)$. The "probability" of the joint event—in which the realized signal and the true type are both equal to s—is equal to the probability of $x \leq c$, which is $F(c)$, multiplied by the "probability" of the true type s conditional on $x \leq c$, which is $f(s|\tilde{x} \leq c)$. Therefore, the posterior probability of type s conditional on signal s is $\frac{F(c) f(s|\tilde{x} \leq c)}{f(s|\tilde{x} \leq c)} = F(c)$.

Theorem 1. *Let $\hat{c} \in (0, 1)$ denote the cutoff defined by equation (3). The truth-or-noise signal structure with cutoff $\hat{c}$ is a crime-minimizing signal structure.*

Proof. It suffices to show that signal structure $(S_{\hat{c}}, \pi_{\hat{c}})$ has an equilibrium in which types above $\hat{c}$ commit crimes and types below $\hat{c}$ do not. Consider the following strategy profile: The police adopt search strategy $p^*(s) = s$ for every $s \in [0, \hat{c}]$, and each agent commits a crime if and only if their type exceeds $\hat{c}$. We show that this strategy profile is an equilibrium. First, from the police's perspective, each agent is committing a crime with probability $1-F(\hat{c})$ conditional on any signal. Thus the police are indifferent across all search strategies that exhaust search capacity $\overline{P}$. Search strategy p^* indeed exhausts the search capacity because of equation (3), i.e., $\overline{P} = \mathbb{E}_F[\tilde{x}|\tilde{x} \leq \hat{c}]$. The strategy of each agent is also optimal: Any agent with type $x \leq \hat{c}$ knows that the police will observe signal $s = x$ and search them with probability x, so the agent is indifferent and be willing to abstain from a crime. Agents with types above $\hat{c}$ will be searched with probability at most $\hat{c}$, so they commit crimes. Hence the strategy profile described above is an equilibrium. □

The crime-minimizing signal structure is partially informative about the agents' types but uninformative about their equilibrium behavior. We provide intuition for each property. First, the crime-minimizing signal structure prevents the police from predicting crimes by equalizing the likelihood of crimes across all signals. This property minimizes the distortion highlighted by Theorem 0, where the police, which cannot commit to a search strategy, end up concentrating their search resources on agents most likely to commit crimes. Various signal structures have this first property, including the one that discloses no information. Second, the crime-minimizing signal structure reveals partial information and enables the police to reduce wasteful searches: The signals are differentiated according to the lowest possible types, and the equilibrium search rates are tailored to these types. Hence the police will never search agents with a probability greater than what is minimally necessary to deter crimes.

Remark 1 (Multiplicity of Equilibria). The game induced by the crime-minimizing signal structure in Theorem 1 has other equilibria. Indeed, for any crime rate $r \in [1-F(\hat{c}), 1]$, there exists an equilibrium in which the crime rate is r and the police continue to adopt the same search strategy as in the crime-minimizing equilibrium. In this equilibrium, any type above $\hat{c}$ commits a crime with probability 1, and any type below $\hat{c}$ commits a crime with probability $1 - \frac{1-r}{F(\hat{c})}$. Such an equilibrium exists, because all types below $\hat{c}$ are indifferent between committing a crime and not under the police's search strategy. Therefore, if we were to explicitly model an information designer who aims to minimize a crime rate, we should think of the designer as considering partial implementation as opposed to, e.g., adversarial equilibrium selection (Bergemann and Morris, 2019).

Remark 2 (Multiplicity of Crime-Minimizing Signals). The crime-minimizing signal is not unique in the relaxed problem or in the original problem. The relaxed problem has multiple solutions, such as the signal structures in Lemma 1 and Theorem 1. The original problem also has multiple solutions. As an example, consider $F = U[0,1]$ and $\overline{P} = \frac{1}{4}$. By solving equation (3), we obtain $\hat{c} = \frac{1}{2}$. One crime-minimizing signal structure that is different from the one in Theorem 1 is the following: Signal $s \in [0, \frac{1}{2}]$ is realized with probability 1 if and only if the true type is s or $1-s$. Conditional on signal s, the posterior belief places equal probability on the events $x = s$ and $x = 1 - s \geq \hat{c}$. Noting that $F(\hat{c}) = \frac{1}{2}$, we can apply the proof of Theorem 1 verbatim and conclude that this new signal structure attains the same crime rate as the one described in Theorem 1.

3.3 The Case of Partial Commitment

Theorem 1 assumes that the police cannot commit to a search strategy. However, the result is relevant even when a search strategy is partly determined by an entity that has some commitment power and cares about reducing crime. We clarify this point by using the following variation of the model.

The baseline model did not explicitly model the information designer. We now include the designer as a player. The designer's objective is to minimize the crime rate. We may view the designer as a law enforcement organization, and the police as individual enforcers.

To capture the designer's partial commitment power over search strategies, we assume that there exists an exogenous signal structure $(\overline{S}, \overline{\pi})$ such that the designer can commit to how to act based on this information. Specifically, the designer can commit to any search strategy $\overline{p}:\overline{S}\to[0,1]$ subject to the capacity constraint (1) (we continue to impose Assumption 1). The realized signal from $(\overline{S},\overline{\pi})$ for each agent and search strategy $\overline{p}$ are publicly observable. For clarity, we call each signal $t\in\overline{S}$ a *pre-signal*.

We assume that the posterior distribution F_t over types induced by each pre-signal $t\in\overline{S}$ has a positive density on $[0,1]$. Thus, $(\overline{S},\overline{\pi})$ cannot be fully informative, which means that the designer has commitment power over limited information.

In addition to a search strategy, the designer chooses a signal structure (S_t,π_t) for each pre-signal $t\in\overline{S}$. The interpretation of $\{(S_t,\pi_t)\}_{t\in\overline{S}}$ is that it captures the information available to the designer in addition to $(\overline{S},\overline{\pi})$, but the designer cannot commit to how to act based on this additional information. Specifically, this additional information is used by the police: As in the baseline model, the police learn from signal structure (S_t,π_t) about the types of agents who have pre-signal t. The police then allocate search capacity $\overline{p}(t)$ across agents with each pre-signal t, and simultaneously, agents decide whether to commit crimes.

The designer could potentially set all (S_t,π_t) to be uninformative and force the police to adopt the designer's search strategy, $\overline{p}$. Alternatively, the designer might enable the police to base their search strategy on more information than $(\overline{S},\overline{\pi})$, but the resulting equilibrium outcome may go against the designer's objective. How should the designer trade-off the control over search strategies and the greater use of information?

To see how Theorem 1 helps us solve the designer's problem, suppose that the designer has committed to search strategy $\overline{p}$. Then, we can view the game between the police and the agents with pre-signal $t\in\overline{S}$ as our baseline model in which the prior type distribution is $F=F_t$ and the search capacity is $\overline{P}=\overline{p}(t)$. Therefore, the designer should provide the police with additional information (S_t,π_t) according to the truth-or-noise signal structure with cutoff $c_t\left(\overline{p}\left(t\right)\right)$, which solves

$$\mathbb{E}_{F_t}\left[\tilde{x}|\tilde{x}\leq c_t\left(\overline{p}\left(t\right)\right)\right]=\overline{p}\left(t\right).$$

The resulting crime rate conditional on pre-signal t is $1-F_t(c_t\left(\overline{p}\left(t\right)\right))$. The designer then optimizes over search strategies in order to minimize the crime rate,

$$\sum_{t\in\overline{S}} \mathbb{P}(t)\left[1 - F_t(c_t\left(\overline{p}\left(t\right)\right))\right],$$

subject to the search capacity constraint, where $\mathbb{P}(t)$ is the ex ante probability of pre-signal t.

One caveat is that we can no longer impose Assumption 1 on each $(F_t, \overline{p}\left(t\right))$, because it is endogenous. However, in Appendix A, we show that inequality (2) holds with weak inequality under the designer's optimal search strategy, and thus the cutoff $c_t\left(\overline{p}\left(t\right)\right) \in [0, 1]$ uniquely exists. Therefore, the truth-or-noise signal structures constitute the designer's optimal policy. The appendix also solves an example in which the designer provides the police with truth-or-noise signal structures with cutoffs that differ across pre-signals.

4 Endogenous Search Capacity

Under a certain condition, a version of Theorem 1 holds when the police can increase the total number of searches at cost. Formally, in this section, we assume that the police can choose any search strategy p at cost $C(P)$, where P is the total mass of searches induced by p, i.e.,

$$P \triangleq \int_0^1 \int_S p(s)\pi(ds|x)F(dx).$$

The police's payoff is the mass of successful searches minus cost $C(P)$. The rest of the model, such as the agents' payoffs and the timing, remains the same.

4.1 Relaxed Problem with Endogenous Search Capacity

First, we study the relaxed problem. Namely, we assume that the agents do not directly observe their types and solve for a joint signal structure (S_P, S_A, π) that minimizes the equilibrium crime rate.

When we analyze the original problem, we will impose a functional form assumption on cost function $C(\cdot)$. However, for the relaxed problem, we only assume that $C(\cdot)$ is strictly increasing, strictly convex, differentiable, and satisfies

$$C'(0) < 1 < C'\left(\int_0^1 xF(dx)\right).$$

The first inequality implies that the police search a positive mass of agents under any signal structure and any equilibrium. The second inequality ensures that the equilibrium crime rate is positive, which plays the same role as Assumption 1,

Figure 2 describes a crime-minimizing joint signal structure in the relaxed problem with endogenous search capacity. Similar to the case of the exogenous search capacity (Lemma 1), the agents receive signals *crime* or *not* depending on whether their types exceed some cutoff c^*, and in equilibrium, they follow action recommendations. The police's signal is a garbling of the agents' signals. In particular, the police privately identify—and therefore choose not to search—a fraction α^* of agents who receive signal *not*. For the remaining population, the police apply the same search rate $\rho^* > 0$. Parameter α^* can be 0, in which case the crime-minimizing joint signal structure takes qualitatively the same form as Lemma 1, i.e., the police receive no information, and the agents receive a cutoff signal.

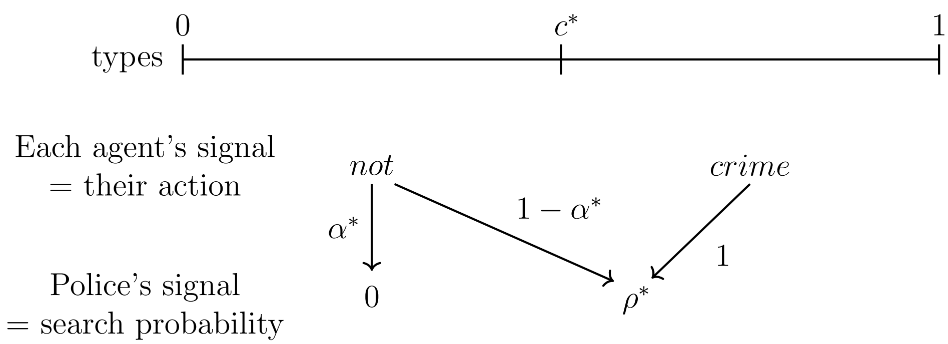

Figure 2: A solution to the relaxed problem with endogenous search capacity. The police identify a fraction α^* of the agents who do not commit crimes.

The following result formally describes the crime-minimizing signal structure of the relaxed problem. All omitted proofs are relegated to Appendix B.

Lemma 2. . *In the relaxed problem, the following joint signal structure (S_P^*, S_A^*, π^*), characterized by tuple $(\rho^*, c^*, \alpha^*) \in (0,1)^2 \times [0,1)$, minimizes a crime rate:*

1. *For every type $x > c^*$, the realized signal is $(s_i^P, s_i^A) = (\rho^*, crime)$ with probability*

1. *For every type $x < c^*$, the realized signal is $(0, not)$ or (ρ^*, not) with probability α^* or $1-\alpha^*$, respectively. In equilibrium, the agents and the police follow the action recommendations.*

2. *Tuple (ρ^*, c^*, α^*) satisfies $\mathbb{E}_F[\tilde{x}|\tilde{x} \leq c^*] = (1-\alpha^*)\rho^*$, i.e., the agents who observe signal "not" are indifferent between committing a crime and not.*

When $\alpha^* > 0$, the crime-minimizing signal structure of the relaxed problem enables the police to identify a fraction of agents who do not commit crimes in equilibrium. To see why such partial revelation can be optimal when the search capacity is endogenous, suppose that a mass $r \in (0,1)$ of agents are committing crimes. Compare the following two cases: In Case 0, the police have no information. In Case α, the police can privately identify a fraction α of the agents who do not commit crimes (or equivalently, a mass $\alpha(1-r)$ of such agents). Moving from Case 0 to Case α affects the police's strategy in two ways. First, the agents who do not commit crimes are on average less likely to be exposed to search, because the police do not search a fraction α of them. Second, the police choose a higher total search capacity in Case α than Case 0, because the police can detect a crime with probability $\frac{r}{1-\alpha(1-r)} > r$ by searching a mass $1-\alpha(1-r)$ of unidentified agents in Case α, whereas the probability of detecting a crime is r in Case 0. The crime-minimizing signal structure involves partial revelation (i.e., $\alpha^* > 0$) when the second effect dominates, so that the overall costs for the agents of committing crimes increase as we move from Case 0 to Case α.

At the same time, consistent with the case of exogenous search capacity, the crime-minimizing signal structure does not enable the police to identify "criminals," the agents who commit crimes. To see why, consider Case β in which the police can identify a fraction β of criminals. In contrast to Case α, moving from Case 0 to Case β could reduce the police's effort, because the probability of detecting a crime in the unidentified population is $\frac{(1-\beta)r}{1-\beta r} < r$. Moreover, as in the baseline model, the information about criminals distorts the allocation of searches and reduces the deterrent effect of search. Thus, allowing the police to predict crimes could increase a crime rate by both reducing search effort and distorting its allocation.

4.2 The Optimality of Truth-or-Noise Signals

When $\alpha^* = 0$ in Lemma 2, the solution to the relaxed problem with endogenous search capacity is qualitatively the same as the case of exogenous search capacity. Correspondingly, the truth-or-noise signal structures continue to solve the original problem.

Lemma 3. *Suppose that the solution to the relaxed problem described in Lemma 2 has $\alpha^* = 0$, i.e., the police receive no information. Then, the crime-minimizing signal structure of the original problem is the truth-or-noise signal structure with cutoff c^*, where c^* and the equilibrium search capacity P^* jointly solve*

$$\mathbb{E}_F[\tilde{x}|\tilde{x} \leq c^*] = P^*, \text{ and} \tag{4}$$

$$1 - F(c^*) = C'(P^*). \tag{5}$$

Proof. Equation (4) and the same argument as Theorem 1 imply that under the truth-or-noise signal structure with cutoff c^*, there exists a strategy profile such that: the agents optimally commit crimes if $x > c^*$ and not if $x < c^*$; the police adopt search strategy $p(s) = s$ for every $s \in [0, c^*]$; and the police cannot increase their payoffs by changing the search strategy while keeping the total search capacity P^* fixed. Thus, it remains to show that the police have no profitable deviation in terms of changing the total search capacity. The police's payoff from search capacity P is $(1 - F(c^*))P - C(P)$, because the posterior crime rate is equalized to be $1 - F(c^*)$ across all signals. Due to the convexity of $C(\cdot)$, equation (5) is sufficient for the optimality of P^*. $\square$

Lemma 3 provides a sufficient condition under which Theorem 1 extends to the case of endogenous search capacity. However, it does not tell when we obtain $\alpha^* = 0$ in the relaxed problem. The following result provides a sufficient condition for the primitives under which $\alpha^* = 0$ holds. The result relies on the following restriction on the cost function.

Assumption 2. The police's search cost function takes the form of

$$C(P) = \frac{L}{1+\beta} P^{1+\beta}$$

for some $\beta > 0$ and $L > 1$.

Proposition 1. *Assume that (F, β, L) satisfies*

$$\mathbb{E}_F\left[\tilde{x} \,\middle|\, \tilde{x} < F^{-1}\left(\frac{\beta}{1+\beta}\right)\right] > ((1+\beta)L)^{-\frac{1}{\beta}}. \tag{6}$$

We have $\alpha^* = 0$ in the relaxed problem. Therefore, the crime-minimizing signal structure of the original problem is a truth-or-noise signal structure, where the cutoff and the equilibrium search capacity jointly solve equations (4) and (5).

For a fixed β, inequality (6) is more likely to hold when the police face a greater search cost (i.e., L is high) or the agents face greater returns on crimes (i.e., $F^{-1}\left(\frac{\beta}{1+\beta}\right)$ is high). When these conditions hold, the police incur a high marginal cost of search in equilibrium. In such a case, revealing a fraction of agents with signal *not* does not much affect the police's search effort but decreases the probability $1 - \alpha$ with which the agents with signal *not* are exposed to search. To relax the obedience constraint for signal *not* as much as possible and attain a lower equilibrium crime rate, the crime-minimizing signal structure in the relaxed problem provides no information to the police. In this case, the crime-minimizing signal structure in the original problem becomes a truth-or-noise signal structure.

While the general analysis of the original problem is beyond the scope of the paper, the following example shows that even when $\alpha^* > 0$, a natural generalization of truth-or-noise signals may solve the original problem.

Claim 1. *Suppose that $F = U[0, 1]$ and $C(P) = \frac{L}{2}P^2$ with $L \geq \frac{3+\sqrt{5}}{4} \approx 1.31$. The equilibrium crime rate is minimized by a signal structure that reveals a fraction $\alpha^* = \max(0, 2 - \sqrt{2L})$ of the agents with types below c^* and discloses the truth-or-noise signal with cutoff c^* for the rest of the agents.*

In the above example, when $L \in [L^*, 2)$, the crime-minimizing outcome becomes as follows: The police privately identify a fraction α^* of agents who have types below c^* as signal ϕ. As to the remaining population, the police observe the truth-or-noise signal with cutoff c^*, i.e., a signal coincides with an agent's true type if the type is below c^*, and otherwise the signal is a noise drawn from $F(\cdot|\tilde{x} \leq c^*)$. These signals are private to the police, so the agents with type x below c^* do not know whether the police observe signal ϕ or x. In equilibrium, the police search agents with signal $s \in [0, c^*]$ with probability $\frac{s}{1-\alpha^*}$. Facing such a strategy, types below c^* are indeed willing to not commit crimes, because any agent with type $x < c^*$ believes that they will be searched with probability $(1-\alpha^*)\frac{x}{1-\alpha^*} = x$.[8]

5 Concluding Discussion

From the applied perspective, the crime-minimizing signal structure in Theorem 1 is unlikely to give a direct normative prescription on what information a law enforcement agency should have; the model is agnostic about how information is generated, and in practice, the set of feasible information structures will be limited by various constraints such as legal restrictions. Instead, we should view the results as a cautionary tale that tells why predicting crimes may backfire when the law enforcer cannot commit to how to act based on the information. The paper illustrates this idea by showing that the crime-minimizing signal is typically uninformative about the likelihood with which each agent commits a crime. From the theoretical perspective, the paper studies an information design problem in which the sets of players, actions, and states are infinite, and the feasible information structures are restricted (i.e., the agents must know their types). The paper uses an approach based on a relaxed problem, which may be useful for other information design problems with similar constraints.

References

Becker, Gary S (1968), "Crime and punishment: an economic approach." *Journal of Political Economy*, 76, 169–217.

Bergemann, Dirk and Stephen Morris (2019), "Information design: A unified perspective." *Journal of Economic Literature*, 57, 44–95.

[8]In general, the above signal structure and strategy profile may not solve the original problem. For example, the police's search strategy may be infeasible because we may have $\frac{s}{1-\alpha^*} > 1$. Also, some types above c^*, who anticipate search probability $\frac{\mathbb{E}[\tilde{x}|\tilde{x} \leq c^*]}{1-\alpha^*}$, may abstain from crimes, which leads to a different outcome from the solution to the relaxed problem.

Bjerk, David (2007), "Racial profiling, statistical discrimination, and the effect of a colorblind policy on the crime rate." *Journal of Public Economic Theory*, 9, 521–545.

Boleslavsky, Raphael and Kyungmin Kim (2018), "Bayesian persuasion and moral hazard." *Available at SSRN 2913669*.

Brayne, Sarah (2020), *Predict and Surveil: Data, discretion, and the future of policing*. Oxford University Press, USA.

Che, Yeon-Koo, Jinwoo Kim, and Konrad Mierendorff (2023), "Predictive enforcement." *mimeo*.

Cotton, Christopher and Cheng Li (2015), "Profiling, screening, and criminal recruitment." *Journal of Public Economic Theory*, 17, 964–985.

Curry, Philip A and Tilman Klumpp (2009), "Crime, punishment, and prejudice." *Journal of Public economics*, 93, 73–84.

Dworczak, Piotr and Giorgio Martini (2019), "The simple economics of optimal persuasion." *Journal of Political Economy*, 127, 1993–2048.

Eeckhout, Jan, Nicola Persico, and Petra E Todd (2010), "A theory of optimal random crackdowns." *American Economic Review*, 100, 1104–35.

Goldman, AJ and MH Pearl (1976), "The dependence of inspection-system performance on levels of penalties and inspection resources."

Guo, Yingni and Eran Shmaya (2019), "The interval structure of optimal disclosure." *Econometrica*, 87, 653–675.

Hernández, Penélope and Zvika Neeman (2022), "How bayesian persuasion can help reduce illegal parking and other socially undesirable behavior." *American Economic Journal: Microeconomics*, 14, 186–215.

Hörner, Johannes and Nicolas S Lambert (2021), "Motivational ratings." *The Review of Economic Studies*, 88, 1892–1935.

Jung, Christopher, Sampath Kannan, Changhwa Lee, Mallesh Pai, Aaron Roth, and Rakesh Vohra (2020), "Fair prediction with endogenous behavior." In *Proceedings of the 21st ACM Conference on Economics and Computation*, 677–678.

Kamenica, Emir (2019), "Bayesian persuasion and information design." *Annual Review of Economics*, 11, 249–272.

Kamenica, Emir and Matthew Gentzkow (2011), "Bayesian persuasion." *American Economic Review*, 101, 2590–2615.

Knowles, John, Nicola Persico, and Petra Todd (2001), "Racial bias in motor vehicle searches: Theory and evidence." *Journal of Political Economy*, 109, 203–229.

Kolotilin, Anton, Timofiy Mylovanov, and Andriy Zapechelnyuk (2022), "Censorship as optimal persuasion." *Theoretical Economics*, 17, 561–585.

Lazear, Edward P (2006), "Speeding, terrorism, and teaching to the test." *The Quarterly Journal of Economics*, 121, 1029–1061.

Liang, Annie, Jay Lu, and Xiaosheng Mu (2022), "Algorithmic design: Fairness versus accuracy." In *Proceedings of the 23rd ACM Conference on Economics and Computation*, 58–59.

Perry, Walt L (2013), *Predictive policing: The role of crime forecasting in law enforcement operations*. Rand Corporation.

Persico, Nicola (2002), "Racial profiling, fairness, and effectiveness of policing." *American Economic Review*, 92, 1472–1497.

Persico, Nicola (2009), "Racial profiling? detecting bias using statistical evidence." *Annu. Rev. Econ.*, 1, 229–254.

Persico, Nicola and Petra E Todd (2005), "Passenger profiling, imperfect screening, and airport security." *American Economic Review*, 95, 127–131.

Porto, Edoardo Di, Nicola Persico, and Nicolas Sahuguet (2013), "Decentralized deterrence, with an application to labor tax auditing." *American Economic Journal: Microeconomics*, 5, 35–62.

Rodina, David (2017), "Information design and career concerns." *Manuscript.[1079].*

Smolin, Alex and Takuro Yamashita (2022), "Information design in concave games." *arXiv preprint arXiv:2202.10883.*

Sun, Yeneng (2006), "The exact law of large numbers via fubini extension and characterization of insurable risks." *Journal of Economic Theory*, 126, 31–69.

Vasquez, Jorge (2022), "A theory of crime and vigilance." *American Economic Journal: Microeconomics*, 14, 255–303.

Zapechelnyuk, Andriy (2020), "Optimal quality certification." *American Economic Review: Insights*, 2, 161–176.

Appendix A: Omitted Proof for Section 3.3

We study the designer's problem in two steps. First, we fix a search strategy $\overline{p}$ arbitrarily and then solve for the optimal information policy. As discussed in the main text, if $\mathbb{E}_{F_t}[\tilde{x}] \geq \overline{p}(t)$, the equation $\mathbb{E}_{F_t}[\tilde{x}|\tilde{x} \leq c] = \overline{p}(t)$ has a unique solution $c_t(\overline{p}(t))$, and the designer can attain the minimal crime rate by providing the police with the truth-or-noise signal structure with cutoff $c_t(\overline{p}(t))$. If $\mathbb{E}_{\tilde{x}\sim F_t}[\tilde{x}] < \overline{p}(t)$, then we define $c_t(\overline{p}(t)) = 1$. In this case, the designer can reveal the type of each agent to the police, and the police can deter all crimes (by the agents with pre-signal t) with search strategy $p(x) = x$ for all $x \in [0, 1]$. Overall, the optimal (S_t, π_t) leads to the equilibrium crime rate of $1 - F_t(c_t(\overline{p}(t)))$. The designer's optimal search strategy then solves

$$\max_{(p(t))_{t\in\overline{S}}} \sum_{t\in\overline{S}} \mathbb{P}(t) F_t(c_t(\overline{p}(t))) \tag{7}$$

$$\text{subject to} \quad \sum_{t\in\overline{S}} \mathbb{P}(t)\overline{p}(t) = \overline{P}.$$

In the main text, we claimed that the case of $\mathbb{E}_{F_t}[\tilde{x}] < \overline{p}(t)$ never arises at the optimal search strategy $\overline{p}$. To see why, note that even in this extended model, the equilibrium crime rate is strictly positive for at least one pre-signal, say $t' \in \overline{S}$, by Assumption 1. Suppose to the contrary that $\mathbb{E}_{F_t}[\tilde{x}] < \overline{p}(t)$ holds. Then the designer can improve its payoff in the following way. First, the designer slightly decreases the search capacity allocated to pre-signal t and increases the search capacity allocated to a pre-signal (say t') that has a positive crime rate. Suppose that this change increases $\overline{p}(t')$ to $\overline{p}(t') + \epsilon$. Originally, the signal structure for pre-signal t' was the truth-or-noise signal with cutoff c where $\mathbb{E}_{F_{t'}}[\tilde{x}|\tilde{x} \le c] = \overline{p}(t')$. The cutoff now increases to $c_\epsilon > c$ that solves $\mathbb{E}_{F_{t'}}[\tilde{x}|\tilde{x} \le c_\epsilon] = \overline{p}(t') + \epsilon$. The designer can then replace $(S_{t'}, \pi_{t'})$ with the truth-or-noise signal with cutoff c_ϵ. Because $c_\epsilon > c$, this will reduce the equilibrium crime rate. But then we obtain a contradiction. Therefore, under the designer's optimal search strategy, $\mathbb{E}_{F_t}[\tilde{x}] \ge \overline{p}(t)$ holds, so $\mathbb{E}_{F_t}[\tilde{x}|\tilde{x} \le c] = \overline{p}(t)$ has a unique solution for every $t \in \overline{S}$.

It is beyond the scope of this paper to provide a general solution to problem (7). Instead, we provide an example.

Example 1. Suppose that each pre-signal is ex ante equally likely, i.e., $\mathbb{P}(t) = \frac{1}{|\overline{S}|}$ for every $t \in \overline{S}$, and each F_t is such that $\frac{F_t(x)}{x}$ is decreasing in $x \in [0, 1]$. We also assume that $\lim_{x \to 0} \frac{F_t(x)}{x} = \infty$. One example that satisfies these conditions is $\overline{S} \subset (0, 1)$ and $F_t(x) = x^t$ for each $t \in \overline{S}$.

Recall that if the designer allocates search capacity P to pre-signal t and provides the police with the crime-minimizing signal, the cutoff type c is given by

$$\mathbb{E}_{F_t}[\tilde{x}|\tilde{x} \le c] = \int_0^c \frac{x f_t(x)}{F_t(x)} dx = P.$$

Differentiating both sides in P yields $1 = \frac{cf(c)}{F(c)} \cdot \frac{\partial c}{\partial P}$. Thus, we obtain

$$\frac{\partial F_t(c)}{\partial P} = f_t(c) \frac{\partial c}{\partial P} = \frac{F_t(c)}{c}.$$

Therefore, at the solution to (7), the cutoffs induced by search rates allocated to different pre-signals must equalize $\frac{F(c)}{c}$:

$$\frac{F_t(c_t)}{c_t} = \frac{F_s(c_s)}{c_s}, \forall s, t \in \overline{S}.$$

Indeed, if we had $\frac{F_t(c_t)}{c_t} < \frac{F_s(c_s)}{c_s}$, the designer would commit to allocate less searches to pre-signal t and more to pre-signal s, because combined with the associated changes in the signals structures, the designer would be able to induce a lower crime rate. For example, if different posteriors $\{F_t\}_{t\in\overline{S}}$ are ranked in the first-order stochastic dominance, then the designer will allocate more searches to pre-signals that correspond to stochastically smaller posteriors (i.e., a higher $F_t(x)/x$).

Appendix B: Omitted Proofs for Section 4

Proof of Lemma 2

In what follows, we refer a joint signal structure simply as a signal structure or "signal," when it does not cause confusion. Take any signal structure (S_P, S_A, π) and any strategies of agents. For each $s \in S_P$, the *posterior crime rate*, $r(s) \in [0, 1]$, refers to the probability with which an agent is committing a crime conditional on the police's signal s.[9]

To prove Lemma 2, we first prove a lemma that enables us to restrict the class of signals we need to consider.

Lemma 4. *Consider the relaxed problem, and take any signal structure and any equilibrium. Let r denote the equilibrium crime rate. There is some $c \in [0, 1]$ such that the same crime rate arises under a signal structure and an equilibrium with the following properties:*

1. *Each agent receives signal "crime" and signal "not" if $x > c$ and $x < c$, respectively. In equilibrium, the agents follow action recommendations.*

[9]For every signal $t \in S_A$ of an agent, let $a(t)$ be the average probability with which an agent commits a crime after observing signal t. Then the posterior crime rate is equal to $r(s) = \mathbb{E}[a(\tilde{t})|s]$, where the expectation is with respect to an agent's signal $\tilde{t}$ conditional on the police's signal s.

2. *The police's signal is a garbling of an agent's signal.*[10] *In equilibrium, each signal of the police leads to a distinct posterior crime rate.*

Proof. Take any signal structure (S'_P, S'_A, π') and any equilibrium. Let p' denote the police's equilibrium search strategy. First, as in Lemma 1, we can replace the signal space S'_A of agents with their action space $S_A = \{crime, not\}$, and assume that each agent follows the action recommendation in equilibrium.

Second, we replace each signal $s' \in S'_P$ of the police with its posterior crime rate $r(s')$ induced by the agents' strategies. Let $S_P \subset [0,1]$ be the resulting signal space for the police. This modification reduces the police's information, because different signals may have the same posterior crime rate. We then assume that the police adopt search strategy $p(y) \triangleq \mathbb{E}[p'(s')|r(s') = y]$ for each $y \in S_P$, where the expectation is with respect to the police's original signal $s' \in S'_P$ conditional on that the posterior crime rate associated with the signal is y. The police find it optimal to adopt p because it ensures the same payoff as p' despite the police having less information. The agents' incentives remain the same, because strategies p' and p result in the same expected search rates conditional on each signal in S_A.

Finally, let $\hat{\pi} \in \Delta(S_P \times S_A)$ denote the joint distribution of the police's signal (i.e., posterior crime rate) and an agent's signal (i.e., action recommendation). Let $\hat{\pi}(\cdot|s) \in \Delta S_P$ denote the associated conditional distribution of the police's signal given an agent's signal $s \in \{crime, not\}$. We then modify the signal structure as follows. First, given equilibrium crime rate $r \in [0,1]$, we assume that types above and below cutoff $c \triangleq F^{-1}(1-r)$ receive signals *crime* and *not*, respectively. This modification relaxes the obedience constraints for the agents, so they are willing to follow action recommendations. Second, we assume that conditional on each agent i's signal $s_i \in \{crime, not\}$, the police observe signal $y_i \sim \hat{\pi}(\cdot|s_i) \in \Delta S_P$ (regardless of i's type). This modification preserves the joint distribution of posterior crime rates and action recommendations across the population. As a result, the agents continue to follow action recommendations, and the police optimally choose search straetgy p. The resulting equilibrium has the properties stated in the lemma. □

[10] By "garbling," we mean that there exist conditional distributions of the police's signal given an agent's signal, i.e., $\pi_P(\cdot|crime), \pi_P(\cdot|not) \in \Delta S_P$, such that for any $S \subset S_P$, $a \in \{crime, not\}$, and $x \in [0,1]$, we have $\pi(S \times \{a\}|x) = \pi_P(S|a)\pi_A(a|x)$. Here, $\pi_A(s|x)$ is the probability of an agent's signal being a conditional on type x.

We now prove Lemma 2.

Proof of Lemma 2. Take any signal structure (S_P, S_A, π) and any equilibrium with the properties described in Lemma 4. Part 2 of the lemma ensures that the police's signal is a garbling of an agent's signal, generated by conditional distributions $\hat{\pi}(\cdot|crime), \hat{\pi}(\cdot|not) \in \Delta S_P$. It is without loss to assume that the police's signals are recommended search probabilities the police follow in equilibrium.

First, we show $S_P \subset \{0, \rho, 1\}$ for some $\rho \in (0, 1)$. If S_P contains multiple interior search rates $\rho, \rho' \in (0, 1)$, they must have the same posterior crime rate, i.e., $r(\rho) = r(\rho')$. For example, if $r(\rho) < r(\rho')$, the police would profitably deviate by shifting search probabilities from signal ρ to ρ' without changing the total search capacity. However, $r(\rho) = r(\rho')$ contradicts Part 2 of Lemma 4 that each signal leads to a distinct posterior crime rate.[11] As a result, S_P contains at most one interior search rate, i.e., $S_P \subset \{0, \rho, 1\}$ for some ρ.

The unique interior search rate ρ (if exists) satisfies two properties. First, $\{\rho, 1\} \subset S_P$ implies $r(\rho) < r(1)$, i.e., the agents whom the police search with a higher probability (here, 1) have a higher posterior crime rate. Second, the police equate the marginal cost of search with the marginal probability of detecting a crime. Hence total search P solves $C'(P) = r(\rho)$.

In the second step, we show that if $\{\rho, 1\} \subset S_P$, we can replace signals ρ and 1 with the same signal σ to increase the search probability allocated to signal *not*. Indeed, after pooling signals ρ and 1, the posterior crime rate $r(\sigma)$ for signal σ satisfies $r(\sigma) > r(\rho)$. Thus if we let the police choose an optimal search strategy, the police will choose total search capacity $\tilde{P} > P$, because the marginal return on searches at P is now $r(\sigma) - C'(P) > r(\rho) - C'(P) = 0$. This pooling also reduces the police's information and thus the fraction of searches that go to the agents with signal *crime*. As a result, pooling signals ρ and 1 increases the expected search probability conditional on signal *not* and relaxes its obedience constraint. (In contrast, the obedience constraint for signal *crime* may be violated because of this modification.)

We now have $S_P = \{0, \sigma\}$ or $\{\sigma\}$ for some $\sigma > 0$, and the obedience constraint for signal

[11] To be precise, this argument only implies that there exists some $\rho \in (0, 1)$ such that the ex ante probability of a signal belonging to $S_P \cap (0, 1) \setminus \{\rho\}$ is 0 (instead of this set being empty). However, we can replace all signals in $S_P \cap (0, 1) \setminus \{\rho\}$ with signal ρ without affecting the equilibrium crime rate.

not holds.[12] We then gradually increase cutoff type c defined in Lemma 4 while maintaining conditional distributions $\hat{\pi}(\cdot|crime)$ and $\hat{\pi}(\cdot|not)$ that garble the agents' signals for the police. Increasing cutoff c decreases the mass of agents with signal *crime* and the posterior crime rate for every signal. Thus, as c increases, the police will choose lower search rates for every signal. At some cutoff $c^* \geq c$, the obedience constraint for signal *not* binds.[13] The obedience constraint for signal *crime* also holds, because the assumption $C'\left(\int_0^1 xF(dx)\right) > 1$ implies that the police's optimal searches never make all agents weakly prefer to abstain from crimes.

In the last step, we consider two cases. If $S_P = \{\sigma\}$, we obtain the desired result where $\alpha^* = 0$ (note that $(s_i^P, s_i^A) = (0, not)$ is irrelevant if $\alpha^* = 0$).

Otherwise, the signal space, in terms of the recommended search probabilities, is $\{0, \sigma\}$. In this case, we first change the police's signal back to the corresponding posterior crime rate, which results in a signal space $\{r_0, r_\sigma\}$. We then split signal r_0 into signals t_0 and t_σ that have posterior crime rates 0 and $r_\sigma > r_0$, respectively.

We now need to consider two cases. One is when $\sigma \in (0,1)$, i.e., before splitting signal r_0 into t_0 and t_σ, the search rate for signal r_σ is interior. In this case, we must have $C'(P) = r_\sigma$ before splitting. Thus, after splitting, it continues to be optimal for the police to allocate P to signal r_σ and not search agents with signal t_σ or t_0. In particular, searching agents with signal t_σ (in addition to those with r_σ) will entail a higher marginal cost than $C'(P)$. Signals t_σ and r_σ have the same posterior crime rate, so we can pool them into one signal, say u_σ, without changing the police's search capacity P or the agents' incentives.

The other case is when $\sigma = 1$, i.e., before splitting signal r_0 into t_0 and t_σ, the police search signal r_σ with probability 1. In this case, we may have $C'(P) < r_\sigma$ before splitting. Thus, after splitting, the police continue to search agents with signal r_σ with probability 1, and may also prefer to search agents with signal t_σ with a positive probability. As a result, the obedience constraint for signal *crime* may get violated. If so, we first pool signals t_σ and r_σ into signal u_σ. As in the previous paragraph, signals t_σ and r_σ have the same

[12] We cannot have $S_P = \{0\}$ because if the police do not search, all agents commit crimes, which incentivizes the police to choose a positive search rate because of $C'(0) < 1$.

[13] The reason is as follows. If $c^* = 1$, then all agents receive signal *not*, the police choose search probability 0, and the obedience constraint for signal *not* is violated. Thus at some $c^* \in (0,1)$, the agents become indifferent between committing crime and not after receiving signal *not*.

posterior crime rate, and this pooling does not change the police's search capacity or the agents' incentives. In particular, the obedience constraint for signal *crime* is still violated. We then keep increasing cutoff c as we did in the second step (see the underlined part above) until the cutoff hits the point at which the obedience constraint for signal *not* binds and the one for signal *crime* holds.

The police's signal is now binary, i.e., signal r_0 has posterior crime rate 0 and search rate 0, and signal u_σ has a positive crime rate and search probability. In terms of action recommendations, the police's signal space becomes $S_P^* = \{0, \rho^*\}$ for some $\rho^* \in (0, 1]$.

We now have signal structure (S_P^*, S_A^*, π^*) such that: $S_P^* = \{0, \rho^*\}$ and $S_A^* = \{crime, not\}$; the agents with types above and below cutoff c^* receive signals *crime* and *not*, respectively; and the police's signal divides the population into two groups, one with zero posterior crime rate and the other with a positive posterior crime rate. This signal structure and equilibrium satisfy Part 1. Part 2 follows from the second step of this proof. $\square$

Proof of Proposition 1

Proof. We first use Lemma 2 to solve the relaxed problem. We then turn to the original problem and show that inequality (6) ensures $\alpha^* = 0$.

To solve the relaxed problem, we focus on signal structures that take the form described in Figure 2. Instead of parameters (ρ^*, c^*, α^*), we use (ρ, c, α) to indicate that they may not be the crime-minimizing signal structure. Recall that α is the probability with which the police observe signal 0 conditional on that an agent observes signal *not*. When types above some cutoff commit crimes, a crime rate r pins down the cutoff type through $c = F^{-1}(1-r)$.

We fix $\alpha \in [0, 1]$ arbitrarily and then determine the cutoff type c and the unique positive search probability ρ from the mutual best responses of the agents and the police. By Part 2 of Lemma 2, the equilibrium crime rate $r(\alpha)$ is determined by the condition that the police's optimal search strategy given crime rate $r(\alpha)$ makes the agents who observe signal *not* indifferent between committing a crime and not.

We derive the expected search probability given signal *not*. As in Figure 2, if the crime rate is r, the posterior crime rate for signal ρ is $\frac{r}{r+(1-r)(1-\alpha)}$. The police's mass of searches

P then solves the first-order condition $C'(P) = LP^{\beta} = \frac{r}{r+(1-r)(1-\alpha)}$, or

$$P = \frac{1}{L^{\frac{1}{\beta}}} \cdot \left(\frac{r}{r + (1-r)(1-\alpha)} \right)^{\frac{1}{\beta}}.$$

For the moment, we ignore the constraint $\rho \leq 1$ and verity it later. The expected search probability conditional on signal *not* is $(1-\alpha)\frac{P}{r+(1-r)(1-\alpha)}$, or

$$I(r, \alpha) \triangleq \frac{1}{L^{\frac{1}{\beta}}} \frac{(1-\alpha) r^{\frac{1}{\beta}}}{[r + (1-r)(1-\alpha)]^{\frac{1+\beta}{\beta}}}.$$

The binding obedience constraint for signal *not* is written as

$$I(r, \alpha) = \mathbb{E}_{\tilde{x} \sim F}[\tilde{x} | \tilde{x} \leq F^{-1}(1-r)]. \tag{8}$$

At $r = 0$, we have $I(0, \alpha) = 0 < \mathbb{E}_{x \sim F}[x]$. At $r = 1$, we have $I(1, \alpha) = \frac{1-\alpha}{L^{\frac{1}{\beta}}} \geq 0$. Also, $I(r, \alpha)$ and $\mathbb{E}_{\tilde{x} \sim F}[\tilde{x} | \tilde{x} \leq F^{-1}(1-r)]$ are continuous in r. Thus equation (8) has a solution. Let $r(\alpha)$ denote the smallest solution. The minimal crime rate in the relaxed problem is given by $r^* \triangleq \min_{\alpha \in [0,1]} r(\alpha)$. However, instead of solving this minimization problem, we first derive $\alpha(r) \triangleq \arg\max_{\alpha \in [0,1]} I(r, \alpha)$ and then determine the minimal crime rate r^* through $I(r, \alpha(r)) = \mathbb{E}_{\tilde{x} \sim F}[\tilde{x} | \tilde{x} \leq F^{-1}(1-r)]$. Hereafter, we restrict attention to $\alpha \in [0, 1)$ and $r \in (0, 1)$, because $\alpha = 1$ or $r \in \{0, 1\}$ cannot be a part of a crime-minimizing equilibrium.

We have

$$\frac{\partial}{\partial \alpha} \log I(r, \alpha) = -\frac{1}{1-\alpha} + \frac{1+\beta}{\beta} \cdot \frac{1-r}{1-\alpha(1-r)}.$$

Note that $(1-\alpha)(1-\alpha(1-r)) > 0$, but the expression

$$(1-\alpha)(1-\alpha(1-r)) \frac{\partial}{\partial \alpha} \log I(r, \alpha) = -r + \frac{1}{\beta}(1-r)(1-\alpha)$$

changes its sign at most once from positive to negative as α increases, and $\frac{\partial}{\partial \alpha} \log I(r, \alpha) < 0$ for α close to 1. As a result, $\frac{\partial}{\partial \alpha} \log I(r, \alpha)$ is either always negative or changes its sign exactly once from positive to negative. Examining the first-order condition $\frac{\partial}{\partial \alpha} \log I(r, \alpha) = 0$, we obtain the following solution:

$$\alpha(r) = \begin{cases} 1 - \beta\frac{r}{1-r} & \text{if} \quad r \le \frac{1}{1+\beta}, \\ 0 & \text{if} \quad r \ge \frac{1}{1+\beta}. \end{cases}$$

and

$$I(r, \alpha(r)) = \begin{cases} \frac{1}{L^{\frac{1}{\beta}}} \cdot \frac{\beta}{(1+\beta)^{\frac{1+\beta}{\beta}}} \frac{1}{1-r} & \text{if} \quad r \le \frac{1}{1+\beta}, \\ \left(\frac{r}{L}\right)^{\frac{1}{\beta}} & \text{if} \quad r \ge \frac{1}{1+\beta}. \end{cases}$$

We now go back to the equation

$$I(r, \alpha(r)) = \mathbb{E}_{\tilde{x}\sim F}[\tilde{x}|\tilde{x} \le F^{-1}(1-r)].$$

The left-hand side is strictly increasing, and the right-hand side is strictly decreasing in r. Hence the equilibrium crime rate r^* is unique. Moreover, we have $r^* \ge \frac{1}{1+\beta}$ and thus $\alpha(r^*) = 0$ when

$$I\left(\frac{1}{1+\beta}, 0\right) < \mathbb{E}_{\tilde{x}\sim F}\left[\tilde{x} \mid \tilde{x} \le F^{-1}\left(\frac{\beta}{1+\beta}\right)\right],$$

which reduces to inequality (6). Finally, recall that when we derived the police's best response, we temporarily ignored the condition that the search probability ρ^* can be at most 1. This condition is satisfied at equilibrium because ρ^* satisfies $\rho^* \le \mathbb{E}_F[\tilde{x}|\tilde{x} \ge F^{-1}(1-r^*)] \le 1$. □

Proof of Claim 1

Proof. Substituting $F = U[0,1]$ and $\beta = 1$ into the solutions to the relaxed problem for Proposition 1, we obtain the following:

$$\alpha(r) = \begin{cases} \frac{1-2r}{1-r} & \text{if} \quad r \le \frac{1}{2}, \\ 0 & \text{if} \quad r \ge \frac{1}{2}. \end{cases}$$

and

$$I(r,\alpha(r)) = \begin{cases} \frac{1}{4L(1-r)} & \text{if} \quad r \leq \frac{1}{2}, \\ \frac{r}{L} & \text{if} \quad r \geq \frac{1}{2}. \end{cases}$$

Solving $I(r,\alpha(r)) = \frac{1-r}{2}$, we obtain the minimized crime rate under the relaxed problem:

$$r^* = \begin{cases} 1 - \frac{1}{\sqrt{2L}} & \text{if} \quad L \leq 2, \\ \frac{L}{2+L} & \text{if} \quad L \geq 2. \end{cases}$$

Because the type distribution is uniform, a crime rate of r^* means that an agent commits a crime if and only if their type exceeds $c^* = 1 - r^*$.

We now show that if $L \geq L^* = \frac{3+\sqrt{5}}{4}$, we can implement crime rate r^* in the original problem. To do so, we modify the signal structure in Theorem 1 as follows: If $x < 1 - r^*$, with probability $\alpha^* = \alpha(r^*)$, the police observe signal 0. Other parts of the signal structure follow the truth-or-noise signal structure with cutoff c^*: The police observe signal x with probability $1-\alpha^*$ if $x \leq c^*$ and observe signal $s \sim F(\cdot|\tilde{x} \geq c^*)$ whenever $x \geq c^*$.

If $L \geq L^*$, this signal structure has an equilibrium in which each agent commits a crime if and only if $x > c^*$, and the police search agents with signal s with probability $\frac{s}{1-\alpha^*}$. The agents' strategies are optimal: Any agent with a type below c^* is indifferent between committing a crime and not because they anticipate search probability $(1-\alpha^*)\frac{x}{1-\alpha^*} = x$ in expectation. Any type $x \geq c^*$ will face a search probability of $\frac{1-r^*}{2(1-\alpha(r^*))} < 1 - r^* \leq x$, where the first inequality uses $\alpha^* < 1/2$, which follows from $L \geq L^*$. The police's strategy is also optimal: The police never search signal 0 and are indifferent regarding how to allocate a given mass of searches across the positive signals. The choice of a total search capacity is also optimal: Indeed, the posterior crime rate for any positive signal is $\frac{r^*}{r^*+(1-r^*)(1-\alpha^*)}$, so the total search capacity induced by the above search strategy equates the marginal cost with the marginal crime rate because of the police's first-order condition in the relaxed problem. Also, $L \geq L^*$ ensures that the highest search probability $\frac{c^*}{1-\alpha^*}$ is below 1, so the police's strategy is feasible. Finally, if $L \in [L^*, 2)$, then we have $r^* = 1 - \frac{1}{\sqrt{2L}}$ and $\alpha(r^*) = \frac{1-2r^*}{1-r^*} = 2 - \sqrt{2L}$. If $L \geq 2$, we have $\alpha^* = 0$, so the police's signal reduces to the truth-or-noise signal structure. $\square$