Online Appendix
Strategic Abuse and Accuser Credibility

Harry Pei[*] Bruno Strulovici[†]

December 18, 2018

A Existence of Equilibrium

A.1 Proof of Proposition 3

We show that when L is large enough, there exists a symmetric equilibrium such that

1. $q(\boldsymbol{a}) > 0$ if and only if $\boldsymbol{a} = (1, 1)$.

2. The principal either abuses no agent or abuses only one agent.

The key step is to show the following proposition:

Proposition A.1. *For every $(c, \delta) \in \mathbb{R}_+ \times (0, 1)$, there exists $\overline{L} > 0$ such that for every $L > \overline{L}$, there exists a triple $(\omega_m^*, \omega_m^{**}, q_m) \in \mathbb{R}_- \times \mathbb{R}_- \times (0, 1)$ that solves the following three equations:*

$$\frac{q_m}{c}(\omega_m^* - c) = -\frac{1}{\delta\Phi(\omega_m^{**}) + (1-\delta)\alpha} \tag{A.1}$$

$$\frac{q_m}{c}(\omega_m^{**} - c + b) = -\frac{l^*}{l^*+2} \cdot \frac{1}{\delta\Phi(\omega_m^*) + (1-\delta)\alpha} - \frac{2}{l^*+2} \cdot \frac{1}{\delta\Phi(\omega_m^{**}) + (1-\delta)\alpha} \tag{A.2}$$

$$\frac{1}{\delta L} = q_m \Big(\delta\Phi(\omega_m^{**}) + (1-\delta)\alpha\Big)\Big(\Phi(\omega_m^*) - \Phi(\omega_m^{**})\Big). \tag{A.3}$$

Proof of Proposition A.1: The proof consists of two steps. In **Step 1**, we show that once fixing q_m to be 1, the value of the following expression:

$$A \equiv \inf_{(\omega_m^*, \omega_m^{**}) \text{ that solves (A.1) and (A.2) when } q_m = 1} \delta\Big(\Phi(\omega_m^*) - \Phi(\omega_m^{**})\Big)\Big(\delta\Phi(\omega_m^{**}) + (1-\delta)\alpha\Big) \tag{A.4}$$

[*]Department of Economics, Northwestern University. harrydp@northwestern.edu
[†]Department of Economics, Northwestern University. b-strulovici@northwestern.edu

is strictly bounded away from 0. We establish this by putting lower bounds on $\Phi(\omega_m^{**})$ and $\Phi(\omega_m^*) - \Phi(\omega_m^{**})$, respectively. To see this, first,

$$\omega_m^{**} \geq -b + c - \frac{c}{(1-\delta)\alpha}$$

and therefore,

$$\Phi(\omega_m^{**}) \geq \Phi\Big(-b + c - \frac{c}{(1-\delta)\alpha}\Big). \tag{A.5}$$

Next, let $\Delta \equiv \omega_m^* - \omega_m^{**}$, which has to be strictly between 0 and b. Deducting equation (A.2) from (A.1) and plugging in $q_m = 1$, we have:

$$\frac{b-\Delta}{c} = \frac{\delta l^*}{l^*+2}\Big(\delta\Phi(\omega_m^*) + (1-\delta)\alpha\Big)^{-1}\Big(\delta\Phi(\omega_m^{**}) + (1-\delta)\alpha\Big)^{-1}\Big(\Phi(\omega_m^*) - \Phi(\omega_m^{**})\Big). \tag{A.6}$$

Consider two cases:

1. When $\Delta \geq b/2$, then

$$\delta\Big(\Phi(\omega_m^*) - \Phi(\omega_m^{**})\Big) \geq \frac{b\delta}{2}\phi(\omega_m^{**}) \geq \frac{b\delta}{2}\phi\Big(-b + c - \frac{c}{(1-\delta)\alpha}\Big), \tag{A.7}$$

   which uses the assumption that the density of ω is increasing when $\omega < 0$.

2. When $\Delta < b/2$, then (A.6) implies that:

$$\delta\Big(\Phi(\omega_m^*) - \Phi(\omega_m^{**})\Big) \geq \frac{b(l^*+2)}{2l^*c}\Big(\delta\Phi(\omega_m^*) + (1-\delta)\alpha\Big)\Big(\delta\Phi(\omega_m^{**}) + (1-\delta)\alpha\Big).$$

$$\geq \frac{b(l^*+2)}{2l^*c}\Big(\delta\Phi\big(-b + c - \frac{c}{(1-\delta)\alpha}\big) + (1-\delta)\alpha\Big)^2. \tag{A.8}$$

Taking the minimum of the right-hand sides of (A.7) and (A.8), we obtain a lower bound for $\delta\Big(\Phi(\omega_m^*) - \Phi(\omega_m^{**})\Big)$. This together with (A.5) implies a lower bound for (A.4), which is strictly bounded above 0.

We use $\underline{A}$ to denote the lower bound we obtained in Step 1. In **Step 2**, we show that when $L > \underline{A}^{-1}$, there exists a solution to (A.1), (A.2) and (A.3) using a fixed point argument. For every $(\Phi^*, \Phi^{**}, q) \in [0,1]^2 \times [1/L, 1]$, let $f \equiv (f_1, f_2, f_3) : [0,1]^2 \times [1/L, 1] \to [0,1]^2 \times [1/L, 1]$ be the following mapping:

$$f_1(\Phi^*, \Phi^{**}, q) = \Phi\Big(c - \frac{c}{q(\delta\Phi^{**} + (1-\delta)\alpha)}\Big), \tag{A.9}$$

$$f_2(\Phi^*, \Phi^{**}, q) = \Phi\Big(-b + c - \frac{cl^*}{q(l^*+2)} \frac{1}{\delta\Phi^* + (1-\delta)\alpha} - \frac{2c}{q(l^*+2)} \frac{1}{\delta\Phi^{**} + (1-\delta)\alpha}\Big), \tag{A.10}$$

$$f_3(\Phi^*, \Phi^{**}, q) = \min\Big\{1, \frac{1}{\delta L} \frac{1}{\Big(\delta\Phi^{**} + (1-\delta)\alpha\Big)\Big(\Phi^* - \Phi^{**}\Big)}\Big\}. \tag{A.11}$$

Since f is continuous, the Brouwer's fixed point theorem implies the existence of a fixed point.

Next, we show that if (Φ^*, Φ^{**}, q) is a fixed point, then $q < 1$. This implies that every solution to the fixed point problem solves the system of equations (A.1), (A.2) and (A.3) as (A.11) and (A.3) are the same when $q < 1$. Suppose towards a contradiction that $q = 1$, then $\Phi^{-1}(\Phi^*)$ and $\Phi^{-1}(\Phi^{**})$ is a solution to (A.1) and (A.2) once we fix q to be 1. According to Part I of the proof, the assumption that $L > \underline{A}^{-1}$ implies that

$$\frac{1}{\delta L} \frac{1}{\Big(\delta\Phi^{**} + (1-\delta)\alpha\Big)\Big(\Phi^* - \Phi^{**}\Big)} < 1.$$

Therefore the RHS of (A.11) is strictly less than 1. This contradicts the claim that $(\Phi^*, \Phi^{**}, 1)$ is a fixed point of f, which implies that the value of q at the fixed point is strictly less than 1. □

Given the triple $(\omega_m^*, \omega_m^{**}, q_m) \in \mathbb{R}_- \times \mathbb{R}_- \times (0,1)$, one can then uniquely pin down $\widetilde{\pi}_m \in (0,1)$ via:

$$\frac{\delta\Phi(\omega_m^*) + (1-\delta)\alpha}{\delta\Phi(\omega_m^{**}) + (1-\delta)\alpha} = l^* \Big/ \frac{\widetilde{\pi}_m}{1 - \widetilde{\pi}_m}. \tag{A.12}$$

According to the analysis in subsection 3.4 of the main text, (A.1), (A.2), (A.3) and (A.12) are sufficient for the existence of a monotone-responsive equilibrium $\{\omega_m^*, \omega_m^{**}, q_m, \widetilde{\pi}_m\}$ satisfying $q(1,1) = q_m \in (0,1)$ and $q(0,0) = q(1,0) = q(0,1) = 0$.

A.2 Generalization

We generalize our existence proof by allowing for more than two agents and alternative specifications of the mechanical types' reporting strategies. Assume that when agent i is mechanical, he reports with probability p_0 when $\theta_i = 0$ and with probability p_1 when $\theta_i = 1$. We assume that $1 > p_0, p_1 > 0$.[1] We show that for every $\{c, \delta, p_0, p_1\}$, there exists $\overline{L} > 0$ such that for every $L > \overline{L}$, there exists a monotone-responsive equilibrium in symmetric strategies satisfying:

1. $q(\boldsymbol{a}) > 0$ if and only if $\boldsymbol{a} = (1, 1, ..., 1)$.

2. The principal either abuses no agent or abuses only one agent.

[1] In principle, we can also allow the mechanical types of different agents to adopt different reporting probabilities. For notation simplicity, we focus on environments where agents are symmetric.

Similar to the existence proof in the two-agent case where mechanical types' reporting probability does not depend on θ_i (Proposition 3 in the main text), the key step of the existence proof in this general environment is to show the following proposition:

Proposition A.2. *For every $(c, \delta, p_0, p_1) \in \mathbb{R}_+ \times (0,1) \times (0,1) \times (0,1)$, there exists $\overline{L} > 0$ such that for every $L > \overline{L}$, there exists a triple $(\omega^*, \omega^{**}, q) \in \mathbb{R}_- \times \mathbb{R}_- \times (0,1)$ that solves the following three equations:*

$$\frac{q}{c}(\omega^* - c) = -\frac{1}{(\Psi^{**})^{n-1}} \tag{A.13}$$

$$\frac{q}{c}(\omega^{**} - c + b) = -\frac{n}{n + (n-1)l^*} \frac{1}{(\Psi^{**})^{n-1}} - \frac{(n-1)l^*}{n + (n-1)l^*} \frac{1}{(\Psi^{**})^{n-2}\Psi^*} \tag{A.14}$$

$$\frac{1}{\delta L} = q(\Psi^{**})^{n-1}(\Psi^* - \Psi^{**}) \tag{A.15}$$

where

$$\Psi^* \equiv \delta\Phi(\omega^*) + (1-\delta)p_0 \text{ and } \Psi^{**} \equiv \delta\Phi(\omega^{**}) + (1-\delta)p_1.$$

The proof of Proposition A.2 follows from similar steps as that of Proposition A.1, which is available upon request. Notice that (A.13), (A.14), (A.15) together with (A.12) are sufficient for a monotone-responsive equilibrium, where $\widetilde{\pi}$ can be computed via (A.12) after fixing $\{\omega^*, \omega^{**}, q\}$.

B Proof of Theorem 1: Symmetry

We establish the symmetric properties of all monotone-responsive equilibria satisfying $q(0,0) = q(1,0) = q(0,1) = 0$. We show that such conviction rule apply to all monotone-responsive equilibria in the next section (Online Appendix C). The conclusion in this section is the following proposition:

Proposition B.1. *In every monotone-responsive equilibrium where $q(0,0) = q(1,0) = q(0,1) = 0$, the principal chooses $(\theta_1, \theta_2) = (0,1)$ and $(\theta_1, \theta_2) = (1,0)$ with the same probability and the two agents adopt the same reporting cutoffs.*

Proof of Proposition B.1: For $i \in \{1,2\}$, let β_i be the probability that $\theta_i = 1$ conditional on $\theta_j = 1$. We have the following expressions on the reporting cutoffs:

$$\omega_i^* = -c\frac{1 - qQ_{0,j}}{qQ_{0,j}} \tag{B.1}$$

and

$$\omega_i^{**} = -b - c\frac{1 - qQ_{1,j}}{qQ_{1,j}}, \tag{B.2}$$

with

$$Q_{0,j} \equiv \delta\Phi(\omega_j^{**}) + (1-\delta)\alpha$$

and

$$Q_{1,j} \equiv \delta\Big[\beta_j\Phi(\omega_j^{**}) + (1-\beta_j)\Phi(\omega_j^{*})\Big] + (1-\delta)\alpha.$$

Without loss of generality, suppose the probability with which $\theta_i = 0$ is weakly higher compared to the probability with which $\theta_j = 0$. then $\beta_i \leq \beta_j$ and moreover, given that the equilibrium probability of abuse is interior, the principal's incentive constraints imply that the cost of setting $\theta_i = 0$ conditional on $\theta_j = 1$ is no more compared to the cost of setting $\theta_j = 0$ conditional on $\theta_i = 1$:

$$\frac{\delta q\Phi(\omega_j^{**})\Big(\Phi(\omega_i^{*}) - \Phi(\omega_i^{**})\Big)}{\delta q\Phi(\omega_i^{**})\Big(\Phi(\omega_j^{*}) - \Phi(\omega_j^{**})\Big)} \leq 1,$$

which is equivalent to:

$$\frac{\Phi(\omega_i^{*})\Phi(\omega_j^{**})}{\Phi(\omega_j^{*})\Phi(\omega_i^{**})} \leq 1. \tag{B.3}$$

First, we show that $\omega_1^* = \omega_2^*$ and $\omega_1^{**} = \omega_2^{**}$ when the probability of $\theta_1 = 0$ and the probability of $\theta_2 = 0$ are equal, i.e. $\beta_1 = \beta_2$. In this case, both probabilities are interior, which implies that (B.3) holds with equality. Suppose towards a contradiction that $\omega_1^* < \omega_2^*$, then (B.1) implies that $\omega_1^{**} > \omega_2^{**}$. But then we have $\Phi(\omega_1^*)\Phi(\omega_2^{**}) < \Phi(\omega_2^*)\Phi(\omega_1^{**})$, contradicting the equality in (B.3).

Next, we show that $\beta_1 = \beta_2$ in every equilibrium. Suppose towards a contradiction that $\beta_1 < \beta_2$, i.e. $\theta_1 = 0$ occurs with strictly higher probability. Consider the following three cases:

1. If $\omega_1^* > \omega_2^*$, then (B.1) implies that $\omega_1^{**} < \omega_2^{**}$. This contradicts the requirement in (B.3) that $\Phi(\omega_1^*)\Phi(\omega_2^{**}) \leq \Phi(\omega_2^*)\Phi(\omega_1^{**})$.

2. If $\omega_1^* = \omega_2^*$, then (B.1) implies that $\omega_1^{**} = \omega_2^{**}$. On the other hand, (B.2) and $\beta_1 < \beta_2$ imply that $\omega_1^{**} \neq \omega_2^{**}$. This leads to a contradiction.

3. If $\omega_1^* < \omega_2^*$, then $\omega_1^{**} > \omega_2^{**}$. This implies that $\Phi(\omega_1^*)\Phi(\omega_2^{**}) < \Phi(\omega_2^*)\Phi(\omega_1^{**})$. Therefore, the principal faces strictly lower cost to set $\theta_1 = 0$. Therefore in equilibrium, he sets $\theta_1 = 0$ with

positive probability and sets $\theta_2 = 0$ with zero probability. This implies that $\beta_2 = 1$ and therefore

$$\omega_1^* = c - \frac{c}{q(\delta\Phi(\omega_2^{**}) + (1-\delta)\alpha)} \tag{B.4}$$

and

$$\omega_1^{**} = -b + c - \frac{c}{q(\delta\Phi(\omega_2^{**}) + (1-\delta)\alpha)} \tag{B.5}$$

On the other hand, we have:

$$\omega_2^{**} = -b + c - \frac{c}{q\Big(\beta_1(\delta\Phi(\omega_1^{**}) + (1-\delta)\alpha) + (1-\beta_1)(\delta\Phi(\omega_1^{*}) + (1-\delta)\alpha)\Big)} \tag{B.6}$$

According to our previous conclusion that $\omega_1^{**} > \omega_2^{**}$, we from (B.5) and (B.6) that:

$$\Phi(\omega_2^{**}) > \beta_1\Phi(\omega_1^{**}) + (1-\beta_1)\Phi(\omega_1^{*}). \tag{B.7}$$

This leads to a contradiction as $\omega_1^* = \omega_1^{**} + b > \omega_2^{**}$.

$\square$

C Proof of Theorem 1: Conviction Probabilities

We show that $q(0,0) = q(1,0) = q(0,1) = 0$ in all monotone-responsive equilibria when L is large enough. The conclusion is summarized by the following proposition:

Proposition C.1. *For every $\delta \in (0,1)$ and $c > 0$, there exists $\overline{L}(\delta, c) > 0$ such that $q(0,0) = q(1,0) = q(0,1) = 0$ and $q(1,1) \in (0,1)$ in every equilibrium when $L > \overline{L}(\delta, c)$.*

To prove Proposition C.1, we rule out the possibilities of other types of equilibria when L is large enough. For notation simplicity, let $\Phi_i^* \equiv \Phi(\omega_i^*)$ and let

$$\Psi_i^* \equiv \delta\Phi_i^* + (1-\delta)\alpha \text{ and } \Psi_i^{**} \equiv \delta\Phi_i^{**} + (1-\delta)\alpha \text{ for } i \in \{1,2\}.$$

For future reference, it is useful to recall a result that establishes the complementarity and substitutability between the principal's choices of θ_1 and θ_2, which we restate as Lemma C.1:

Lemma C.1. *In every monotone-responsive equilibrium, the principal's choices of θ_1 and θ_2 are* *strategic substitutes if*

$$q(1,1) + q(0,0) - q(1,0) - q(0,1) \tag{C.1}$$

is strictly positive, and are strategic complements if the value of (C.1) is strictly negative.

The proof is in subsection B.2 of the main appendix. The rest of this section is organized as follows. In subsection C.1, we examine equilibria in which θ_1 and θ_2 are strategic substitutes from the principal's perspective. In subsection C.2, we examine equilibria in which θ_1 and θ_2 are strategic complements. In subsection C.3, we examine equilibria in the knife-edge case where the value of (C.1) equals to 0.

C.1 Value of (C.1) is strictly positive

In this subsection, we focus on equilibria in which the principal's decisions are strategic substitutes, namely $q(1,0) + q(0,1) < q(0,0) + q(1,1)$.

First, we claim that if either $q(0,1)$ or $q(1,0)$ is strictly positive, then $q(1,1) = 1$. To understand why, suppose towards a contradiction that both $q(1,0)$ and $q(1,1)$ are interior. Then whether agent 2 report or not leads to the same posterior belief about $\theta_1\theta_2 = 0$. This can only be the case where a_2 is uninformative about θ_2, which implies that $\omega_2^* = \omega_2^{**}$ and hence $\Phi(\omega_2^*) = \Phi(\omega_2^{**})$. This implies that the principal's cost of abusing agent 2 is 0 as it is proportional to $\Phi(\omega_2^*) - \Phi(\omega_2^{**})$, contradicting the fact that his probability of abusing agent 2 is strictly less than 1.

Given that $q(1,1) = 1$ and $q(0,0) = 0$, we have the following expressions for player 1's reporting cutoffs when he has and has not been abused:

$$\omega_1^* \equiv -c\frac{(1-\Psi_2^{**})(1-q(1,0))}{q(1,0) + \Psi_2^{**}(1-q(1,0)-q(0,1))}, \tag{C.2}$$

$$\omega_1^{**} \equiv -b - c\frac{(1-X_2)(1-q(1,0))}{q(1,0) + X_2(1-q(1,0)-q(0,1))}, \tag{C.3}$$

where

$$X_2 \equiv \frac{1-p_1-p_2}{1-p_1}\Psi_2^{**} + \frac{p_2}{1-p_1}\Psi_2^* \tag{C.4}$$

and p_i is the probability with which $\theta_i = 0$. One observation is that ω_1^* is increasing in Ψ_2^{**} and $q(1,0)$, and is decreasing in $q(0,1)$; ω_1^{**} is increasing in X_2 and $q(1,0)$, and is decreasing in $q(0,1)$.

The distance between the two cutoffs is given by:

$$\omega_1^* - \omega_1^{**} = b - (\Psi_2^* - \Psi_2^{**})C_1 \tag{C.5}$$

where

$$C_1 \equiv c(1-q(0,1))(1-q(1,0)) \cdot \frac{p_2}{1-p_1}$$
$$\cdot \frac{1}{q(1,0) + X_2(1-q(1,0)-q(0,1))} \cdot \frac{1}{q(1,0) + \Psi_2^{**}(1-q(1,0)-q(0,1))}. \tag{C.6}$$

Symmetrically, one can obtain the expressions for ω_2^* and ω_2^{**} as well as the distance between them. Conditional on setting $\theta_2 = 1$, the probability with which the principal is convicted increases by:

$$(\Psi_1^* - \Psi_1^{**})\Big(q(1,0) + \Psi_2^{**}(1-q(1,0)-q(0,1))\Big) \tag{C.7}$$

once he sets $\theta_1 = 0$. Similarly, if the principal sets $\theta_2 = 0$ given that $\theta_1 = 1$, this probability is increased by:

$$(\Psi_2^* - \Psi_2^{**})\Big(q(0,1) + \Psi_1^{**}(1-q(1,0)-q(0,1))\Big). \tag{C.8}$$

In every equilibrium, both (C.7) and (C.8) are bounded below $1/L$. In what follows, we establish a lower bound for the maximum of these two expressions, which does not depend on L. This is sufficient to rule out equilibria of this form when L is large enough. Throughout the proof, we assume that $\omega_1^* \geq \omega_2^*$, which is without loss of generality. This leads to the following lemma on the comparison between $q(1,0)$ and $q(0,1)$, the proof of which can be found in subsection C.5:

Lemma C.2. *In every equilibrium where $\omega_1^* \geq \omega_2^*$, we have $q(1,0) \geq q(0,1)$.*

Lower Bound on ω_1^*: For every $\epsilon > 0$,

1. If $q(1,0) \geq \epsilon$, then

$$\omega_1^{**} \geq -b - c\frac{1-\epsilon}{\epsilon}. \tag{C.9}$$

2. If $q(1,0) < \epsilon$, then $q(0,1) \in (0,\epsilon)$ according to Lemma C.2. Therefore, we have:

$$\begin{aligned}
\omega_2^* &= -c\frac{(1-\Psi_1^{**})(1-q(0,1))}{q(0,1)+\Psi_1^{**}(1-q(1,0)-q(0,1))} \\
&\geq -c\frac{\Big(1-\delta\Phi(\omega_1^*-b)-(1-\delta)\alpha\Big)(1-q(0,1))}{q(0,1)+\Big(\delta\Phi(\omega_1^*-b)+(1-\delta)\alpha\Big)\Big(1-q(1,0)-q(0,1)\Big)} \\
&\geq -c\frac{\Big(1-\delta\Phi(\omega_2^*-b)-(1-\delta)\alpha\Big)(1-q(0,1))}{q(0,1)+\Big(\delta\Phi(\omega_2^*-b)+(1-\delta)\alpha\Big)\Big(1-q(1,0)-q(0,1)\Big)} \\
&\geq -c\frac{1-\delta\Phi(\omega_2^*-b)-(1-\delta)\alpha}{(1-\epsilon)\Big(\delta\Phi(\omega_2^*-b)+(1-\delta)\alpha\Big)}. \qquad \text{(C.10)}
\end{aligned}$$

As have shown in Online Appendix A that there exists a solution to the following equation:

$$\omega_2^* = -c\frac{1-\delta\Phi(\omega_2^*-b)-(1-\delta)\alpha}{(1-\epsilon)\Big(\delta\Phi(\omega_2^*-b)+(1-\delta)\alpha\Big)}.$$

Denote it by $\underline{\omega}^*(\epsilon)$ such that (C.10) is satisfied only when $\omega_2^* \geq \underline{\omega}^*(\epsilon)$. Since $\underline{\omega}^*(\epsilon)$ is decreasing in ϵ, a lower bound for ω_1^* is given by:

$$\underline{\omega}_1^* \equiv \sup_{\epsilon\in[0,1]} \Big\{ \min\big\{ -b-c\frac{1-\epsilon}{\epsilon}, \underline{\omega}^*(\epsilon)\big\}\Big\}, \qquad \text{(C.11)}$$

which is finite and moreover, does not depend on L.

Upper Bound on C_1: The key to bound C_1 is to bound the term

$$\frac{1}{q(1,0)+\Psi_2^{**}(1-q(1,0)-q(0,1))} \qquad \text{(C.12)}$$

from above. For every $\epsilon > 0$, consider the following two cases:

1. If $q(1,0) \geq \epsilon$, then (C.12) is no more than $1/\epsilon$.

2. If $q(1,0) < \epsilon$, then $q(0,1) < \epsilon$ according to Lemma C.2. Let $\underline{\omega}_2^{**}(\epsilon)$ be the smallest root of the following equation:

$$\omega \equiv -b-c\frac{1-\delta\Phi(\omega)-(1-\delta)\alpha}{\Big(\delta\Phi(\omega)+(1-\delta)\alpha\Big)(1-\epsilon)}, \qquad \text{(C.13)}$$

which is a lower bound for ω_2^{**} given that $q(1,0), q(0,1) \in [0,\epsilon]$. An upper bound on (C.12) is given by:

$$\frac{1}{q(1,0)+\Psi_2^{**}(1-q(1,0)-q(0,1))} \leq \frac{1}{\Phi(\underline{\omega}_2^{**}(\epsilon))(1-2\epsilon)}. \tag{C.14}$$

In summary we have:

$$C_1 \leq cY^2 \tag{C.15}$$

where

$$Y \equiv \inf_{\epsilon\in[0,1]} \Big\{ \max \big\{1/\epsilon, \frac{1}{\Phi(\underline{\omega}_2^{**})(1-2\epsilon)} \big\} \Big\}.$$

Lower Bound on the Maximum of (C.7) and (C.8): In this last step, we establish a lower bound on the maximum of (C.7) and (C.8). A useful inequality is that for every ω', ω'' with $\omega' > \omega''$,

$$\Phi(\omega') - \Phi(\omega'') \geq (\omega' - \omega'') \min_{\omega\in[\omega',\omega'']} \phi(\omega). \tag{C.16}$$

We consider two cases. First, consider the case in which $\Phi(\omega_1^*) - \Phi(\omega_1^{**}) \geq \Phi(\omega_2^*) - \Phi(\omega_2^{**})$. Using the fact that $\Psi_i^* - \Psi_i^{**} = \delta(\Phi(\omega_i^*) - \Phi(\omega_i^{**}))$, we have:

$$\frac{\delta}{\min_{\omega\in[\omega_1^{**},\omega_1^*]} \phi(\omega)} \Big(\Phi(\omega_1^*) - \Phi(\omega_1^{**})\Big) \geq \omega_1^* - \omega_1^{**} = b - C_1(\Psi_2^* - \Psi_2^{**}) \geq b - C_1(\Psi_1^* - \Psi_1^{**}). \tag{C.17}$$

This together with (C.15) gives an lower bound on $\Psi_1^* - \Psi_1^{**}$. Moreover,

$$\begin{aligned} q(1,0) + \Psi_2^{**}(1-q(1,0)) \quad &\geq \quad q(1,0) + \Psi_2^{**}\big(1-q(1,0)-q(0,1)\big) \\ &\geq \quad \frac{c\big(1-q(1,0)\big)(1-\Psi_2^{**})}{|\underline{\omega}_1^*|}, \end{aligned} \tag{C.18}$$

where the last inequality uses (C.2) as well as the previous conclusion that $\underline{\omega}_1^*$ is a lower bound of ω_1^*. This gives a lower bound on $q(1,0)$. The two parts together imply a lower bound on (C.7).

Second, consider the case in which $\Phi(\omega_1^*) - \Phi(\omega_1^{**}) < \Phi(\omega_2^*) - \Phi(\omega_2^{**})$. Let

$$\beta \equiv \frac{\omega_1^* - \omega_1^{**}}{b}. \tag{C.19}$$

Since $X_2 > \Psi_2^{**}$, we have $\beta \in (0,1)$. First, recall that $\underline{\omega}_1^*$ is the lower bound on ω_1, we have:

$$\frac{1}{\delta}(\Psi_1^* - \Psi_1^{**}) = \Phi(\omega_1^*) - \Phi(\omega_1^{**}) \geq \beta b \phi(\underline{\omega}_1^* - b). \tag{C.20}$$

On the other hand, (C.5) and (C.15) imply that:

$$\Psi_2^* - \Psi_2^{**} = (1-\beta)b/C_1 \geq \frac{(1-\beta)bY^2}{c} \tag{C.21}$$

Since the pdf of normal distribution increases in ω when $\omega < 0$, (C.21) leads to a lower bound on ω_2^{**}. We denote this lower bound by $\widetilde{\omega}(\beta)$. By definition, $\widetilde{\omega}(\beta)$ decreases with β.

1. When $\beta \geq 1/2$, (C.20) implies a lower bound for $\Phi(\omega_1^*) - \Phi(\omega_1^{**})$. Applying (C.18),[2] one can obtain a lower bound for $q(1,0)$. The two together lead to a lower bound on (C.7).

2. When $\beta < 1/2$, we have $\omega_2^{**} \geq \widetilde{\omega}(1/2)$ and furthermore,

$$\Psi_2^* - \Psi_2^{**} \geq \frac{b}{2C_1}.$$

The lower bound on ω_2^{**} also leads to a lower bound on $q(0,1) + \Psi_1^{**}(1 - q(1,0) - q(0,1))$, as (C.3) implies that:

$$\widetilde{\omega}(1/2) \leq \omega_2^{**} \leq \omega_2^* = -c\frac{(1-\Psi_1^{**})(1-q(0,1))}{q(0,1) + \Psi_1^{**}(1-q(1,0)-q(0,1))},$$

which leads to:

$$q(0,1) + \Psi_1^{**}(1-q(1,0)-q(0,1)) \geq \frac{(1-\Psi_1^{**})(1-q(0,1))}{-\widetilde{\omega}(1/2)/c}. \tag{C.22}$$

Since $1-\Psi_1^{**} \geq \delta - \delta\Phi(0)$ and $1-q(0,1) \geq 1/2$, the lower bound on $q(0,1)+\Psi_1^{**}(1-q(1,0)-q(0,1))$ is strictly bounded away from 0. This leads to a uniform lower bound on (C.8).

C.2 Value of (C.1) is strictly negative

Next, we study the case where $q(1,0) + q(0,1) > q(0,0) + q(1,1)$, or in another word, the choices of θ_1 and θ_2 are strategic complements from the principal's perspective. Lemma C.1 implies that conditional on abusing one agent, the principal has a strict incentive to abuse the other agent. Therefore in such equilibria, either both agents are abused or no agent is abused.

We start from two observations. First, $q(1,1) = 1$ in all such equilibria. This is because if $q(1,1) \in (0,1)$ and $q(1,0) + q(0,1) > q(0,0) + q(1,1)$, then one of the agent's report is uninformative about the state, leading to a contradiction. Second, due to the strategic complementarity between θ_1

[2] The validity of inequality (C.18) does not depend on the sign of $\Psi_1^* - \Psi_1^{**} - \Psi_2^* + \Psi_2^{**}$.

and θ_2, agent i's belief about agent j's probability of submitting a report is strictly higher when $\theta_i = 0$ compared to $\theta_i = 1$. This implies that:

$$\min\{\omega_1^* - \omega_1^{**}, \omega_2^* - \omega_2^{**}\} \geq b. \tag{C.23}$$

By setting $\theta_1 = \theta_2 = 1$, the principal's probability of being convicted is increased by at least

$$(\Psi_1^* - \Psi_1^{**})\Big(\Psi_2^*(1-q(0,1)) + (1-\Psi_2^*)q(1,0)\Big) + (\Psi_2^* - \Psi_2^{**})\Big(\Psi_1^{**}(1-q(1,0)) + (1-\Psi_1^{**})q(0,1)\Big), \tag{C.24}$$

compared to the case in which he sets $\theta_1 = \theta_2 = 0$. Therefore, the value of (C.24) cannot exceed $2/L$. The rest of this proof establishes a lower bound on (C.24) that applies uniformly across all L. This in turn implies that when L is large enough, equilibria that exhibit strategic complementarities between θ_1 and θ_2 do not exist.

First, $\max\{q(0,1), q(1,0)\} \geq 1/2$ since $q(0,1) + q(1,0) \geq 1$. Without loss of generality, we assume that $q(1,0) \geq 1/2$. Second, agent i has a dominant strategy of not reporting when $\omega_i > 0$, so $1 - \Psi_i^* \geq \delta(1 - \Phi(0))$. Third, player 1's reporting threshold when $\theta_1 = 0$ is:

$$\omega_1^* = -c\frac{(1-Q_2^H)(1-q(1,0))}{Q_2^H(1-q(0,1)) + (1-Q_2^H)q(1,0)} \tag{C.25}$$

where Q_2^H is the probability with which player 2 submits a report conditional on $\theta_1 = 0$. One can verify that the RHS of (C.25) is strictly increasing in Q_2^H. Therefore,

$$\omega_1^* \geq -c\frac{1-q(1,0)}{q(1,0)} \geq -c.$$

According to (C.23), we have:

$$\frac{1}{\delta}(\Psi_1^* - \Psi_1^{**}) = \Phi(\omega_1^*) - \Phi(\omega_1^{**}) \geq b \min_{\omega \in [-b-c,0]} \phi(\omega). \tag{C.26}$$

The uniform lower bound on (C.24) is then given by:

$$\underbrace{(\Psi_1^* - \Psi_1^{**})}_{\text{according to (C.26)}}\Big(\underbrace{\Psi_2^*(1-q(0,1))}_{\geq 0} + \underbrace{(1-\Psi_2^*)}_{\geq \delta(1-\Phi(0))}\underbrace{q(1,0)}_{\geq 1/2}\Big) + \underbrace{(\Psi_2^* - \Psi_2^{**})\Big(\Psi_1^{**}(1-q(1,0)) + (1-\Psi_1^{**})q(0,1)\Big)}_{\geq 0}$$

$$\geq \frac{\delta^2 b}{2}(1-\Phi(0)) \min_{\omega\in[-b-c,0]} \phi(\omega), \tag{C.27}$$

which concludes the proof.

C.3 Value of (C.1) is 0

Part I: We show that in every equilibrium where the value of (C.1) is 0, each agent is abused with strictly positive probability and moreover, $q(1,1)=1$. The implications of these conclusions are:

1. $q(1,0)+q(0,1)=1$.

2. The marginal cost of abusing each agent is the same.

$$(\Psi_1^* - \Psi_1^{**})q(1,0) = (\Psi_2^* - \Psi_2^{**})q(0,1). \tag{C.28}$$

Suppose towards a contradiction that agent 1 is abused with probability 0, then whether agent 1 reports or not does not affect the evaluator's posterior belief about $\theta_1\theta_2 = 0$. Given that $c > 0$, agent 1 never reports, which implies that the principal has a strict incentive to abuse him. This contradicts Lemma 3.1, which says that in every monotone-responsive equilibrium, the probability with which the principal is guilty is interior.

Suppose towards a contradiction that $q(1,1)\in(0,1)$, then either $q(1,0)\in(0,1)$ or $q(0,1)\in(0,1)$ or both. Without loss of generality, suppose $q(1,0)\in(0,1)$, then whether agent 2 reports or not does not affect the evaluator's posterior belief on whether $\theta_1\theta_2=0$ or $\theta_1\theta_2=1$. Given that $c>0$, agent 2 never reports, which implies that the principal has a strict incentive to abuse him. This contradicts the responsiveness requirement.

Part II: We place a lower bound on the value of (C.28) that uniformly applies across all L. Without loss of generality, we assume that $q(1,0)\geq q(0,1)$, and therefore, $q(1,0)\geq 1/2$. The expressions for agent 1's reporting cutoffs are given by:

$$\omega_1^* = -c\frac{q(0,1)}{q(1,0)}\Big(1-p_x\Psi_2^* - (1-p_x)\Psi_2^{**}\Big)$$

and

$$\omega_1^{**} = -b - c\frac{q(0,1)}{q(1,0)}\Big(1-p_y\Psi_2^* - (1-p_y)\Psi_2^{**}\Big)$$

for some $p_x, p_y \in [0,1]$, which are agent 1's beliefs about θ_2 conditional on each realization of θ_1. The difference between them is given by:

$$\omega_1^* - \omega_1^{**} = b - c\frac{q(0,1)}{q(1,0)}(p_x - p_y)(\Psi_2^* - \Psi_2^{**}). \tag{C.29}$$

where the absolute value of

$$c\frac{q(0,1)}{q(1,0)}(p_x - p_y)$$

is at most c. To bound the LHS of (C.28) from below, we proceed according to the following two steps.

Step 1: Lower bound on ω_1^* According to the expression for ω_1^* and using the assumption that $q(1,0) \geq q(0,1)$, we have:

$$\omega_1^* \geq -c\Big(1 - p_x\Psi_2^* - (1-p_x)\Psi_2^{**}\Big) \geq -c\delta(1 - \Phi(0)). \tag{C.30}$$

Let this lower bound be $\underline{\omega}_1^*$.

Step 2: Lower bound on (C.28) This can be accomplished by establishing strictly positive lower bounds on either of the following expressions: $\Psi_1^* - \Psi_1^{**}$ or $q(0,1)(\Psi_2^* - \Psi_2^{**})$. The former is sufficient since $q(1,0) \geq q(0,1)$ and $q(1,0) + q(0,1) \geq q(1,1) = 1$, implying that $q(1,0) \geq 1/2$.

The case in which $p_x - p_y \leq 0$ is trivial, as $\omega_1^* - \omega_1^{**} \geq b$. The lower bound on ω_1^* then implies a strictly positive lower bound on $\Psi_1^* - \Psi_1^{**}$. The case in which $p_x - p_y > 0$ follows similarly from the last step of subsection C.2. To illustrate the details, we consider two cases separately.

First, suppose $\Psi_1^* - \Psi_1^{**} \geq \Psi_2^* - \Psi_2^{**}$, then we have:

$$\frac{\Psi_1^* - \Psi_1^{**}}{\phi(\underline{\omega}_1^* - b)} \geq \omega_1^* - \omega_1^{**} = b - c(\Psi_2^* - \Psi_2^{**}) \geq b - c(\Psi_1^* - \Psi_1^{**}). \tag{C.31}$$

This yields a strictly positive lower bound on $\Psi_1^* - \Psi_1^{**}$.

Second, suppose $\Psi_1^* - \Psi_1^{**} < \Psi_2^* - \Psi_2^{**}$, then let $\beta \equiv (\omega_1^* - \omega_1^{**})/b$ which is between 0 and 1 due to the assumption that $p_x - p_y > 0$. Equality (C.29) implies that:

$$\omega_1^* - \omega_1^{**} = b - c\frac{q(0,1)}{q(1,0)}(p_x - p_y)(\Psi_2^* - \Psi_2^{**}) \geq b - c(\Psi_2^* - \Psi_2^{**})$$

which yields

$$\Psi_2^* \geq \Psi_2^* - \Psi_2^{**} \geq (1-\beta)b/c. \tag{C.32}$$

This leads to a lower bound on the cutoff ω_2^* as a function of β. We denote this lower bound by $\widetilde{\omega}(\beta)$, which we can show is a decreasing function of β. On the other hand, we also have:

$$\frac{1}{\delta}(\Psi_1^* - \Psi_1^{**}) = \Phi(\omega_1^*) - \Phi(\omega_1^{**}) \geq \beta b \phi(\underline{\omega}_1^* - b). \tag{C.33}$$

Now consider two subcases, depending on the comparison between β and $1/2$.

1. If $\beta \geq 1/2$, then (C.33) implies that

$$\Psi_1^* - \Psi_1^{**} \geq b\delta\phi(\underline{\omega}_1^* - b)/2. \tag{C.34}$$

2. If $\beta < 1/2$, then (C.32) implies that:

$$\Psi_2^* - \Psi_2^{**} \geq b/2c \tag{C.35}$$

Since

$$\omega_2^* = -c(1-Q)\frac{q(1,0)}{q(0,1)} \geq \underline{\omega}_2(\beta) \tag{C.36}$$

where Q is some number between 0 and $(1-\delta)\alpha + \delta\Phi(0)$. This yields the following lower bound on $q(0,1)$, namely

$$q(0,1) \geq \frac{-c(1-Q)q(1,0)}{\underline{\omega}_2(\beta)} \geq -\frac{c}{2\underline{\omega}_2(\beta)} \tag{C.37}$$

which is strictly bounded above 0 for all $\beta < 1/2$. This together with (C.35) lead to the following lower bound on the RHS of (C.28):

$$q(0,1)(\Psi_2^* - \Psi_2^{**}) \geq -\frac{b}{4\underline{\omega}_2(\beta)}. \tag{C.38}$$

C.4 Proof of Lemma C.2

Suppose towards a contradiction that in an equilibrium with the value of (C.1) being strictly positive, we have $\omega_1^* > \omega_2^*$ but $q(1,0) < q(0,1)$, then (C.2) implies that $\Phi(\omega_2^{**}) > \Phi(\omega_1^{**})$ or equivalently,

$\omega_2^{**} > \omega_1^{**}$. This together with $\omega_1^* > \omega_1^{**}$ and $\omega_2^* > \omega_2^{**}$ imply that:

$$\omega_1^{**} < \omega_2^{**} < \omega_2^* < \omega_1^*. \tag{C.39}$$

We start from showing that $p_1, p_2 > 0$. Suppose towards a contradiction that $p_1 = 0$ and $p_2 > 0$, then (C.4) implies that $X_1 = \Psi_1^{**}$. Therefore, $\omega_2^* - \omega_2^{**} = b > \omega_1^* - \omega_1^{**}$, contradicting (C.39). Suppose towards a contradiction that $p_1 > 0$ and $p_2 = 0$, then

$$p_1 \frac{\Psi_1^*}{\Psi_1^{**}} + p_2 \frac{1-\Psi_2^*}{1-\Psi_2^{**}} > p_2 \frac{\Psi_2^*}{\Psi_2^{**}} + p_1 \frac{1-\Psi_1^*}{1-\Psi_1^{**}}. \tag{C.40}$$

This is to say that the evaluator attaches higher probability to $\theta_1\theta_2 = 0$ when only agent 1 reports compared to the case where only agent 2 reports. This implies that $q(1,0) \geq q(0,1)$, which leads to a contradiction.

Given that we have already shown that $p_1, p_2 > 0$, while the responsiveness requirement implies that $\theta_1\theta_2 = 0$ with probability less than 1, i.e. $p_1, p_2 < 1$, we know that both of them are interior so that (C.7) and (C.8) must be equal. Applying the expression (C.2) to both agents, we have:

$$\begin{aligned} \left|\frac{\omega_1^*}{\omega_2^*}\right| &= \frac{1-\Psi_2^{**}}{1-\Psi_1^{**}} \cdot \frac{1-q(1,0)}{1-q(0,1)} \cdot \frac{q(0,1)+\Psi_1^{**}(1-q(1,0)-q(0,1))}{q(1,0)+\Psi_2^{**}(1-q(1,0)-q(0,1))}. \\ &= \frac{1-\Psi_2^{**}}{1-\Psi_1^{**}} \cdot \frac{1-q(1,0)}{1-q(0,1)} \cdot \frac{\Psi_1^*-\Psi_1^{**}}{\Psi_2^*-\Psi_2^{**}}. \end{aligned} \tag{C.41}$$

Since

$$\frac{1-\Psi_1^{**}}{1-\Psi_2^{**}} < \frac{\Psi_1^*-\Psi_1^{**}}{\Psi_1^*-\Psi_2^{**}} \leq \frac{\Psi_1^*-\Psi_1^{**}}{\Psi_2^*-\Psi_2^{**}}$$

Plugging this back in, we have:

$$1 \geq \left|\frac{\omega_1^*}{\omega_2^*}\right| > \frac{1-q(1,0)}{1-q(0,1)}. \tag{C.42}$$

The RHS of (C.42) is greater than 1 as we have assumed that $q(1,0) < q(0,1)$. This leads to a contradiction.

D Mitigating Punishments

In Appendix D.1, we prove Proposition 4 in the main text by constructing equilibria in which the principal's decisions are strategic complements when L belongs to an open interval. In Appendix D.2, we show there exist values of L such that focusing on equilibria where the principal's decisions are strategic complements is without loss of generality. Namely, the value of (C.1) is non-positive in all monotone-responsive equilibria.

D.1 Proof of Proposition 4

Consider equilibria where $q(1,1) = 1$, $q(1,0) = q(0,1) = q$ and $q(0,0) = 0$ with $q \geq 1/2$. The value of (C.1) is strictly negative. According to Lemma C.1, the principal's decisions to abuse agents are strategic complements. Therefore in equilibrium, the principal either chooses $\theta_1 = \theta_2 = 1$ or chooses $\theta_1 = \theta_2 = 0$ but he never abuses only one agent. When $\theta_i = 0$, agent i prefers to report if

$$\omega_i \leq \omega^* \equiv -\frac{c(1-q)(1-\Psi^*)}{q + \Psi^*(1-2q)}. \tag{D.1}$$

When $\theta_i = 1$, agent i prefers to report if

$$\omega_i \leq \omega^{**} \equiv -b - \frac{c(1-q)(1-\Psi^{**})}{q + \Psi^{**}(1-2q)}. \tag{D.2}$$

The principal's indifference condition is given by:

$$2/L = (\Psi^* - \Psi^{**})\Big((1-2q)(\Psi^* + \Psi^{**}) + 2q\Big), \tag{D.3}$$

where

$$\Psi^* \equiv \delta\Phi(\omega^*) + (1-\delta) \quad \text{and} \quad \Psi^{**} \equiv \delta\Phi(\omega^{**}) + (1-\delta).$$

Moreover, the equilibrium probability of crime, denoted by $\widetilde{\pi}_m$, is pinned down by:

$$\frac{\Psi^*(1-\Psi^*)}{\Psi^{**}(1-\Psi^{**})} = \frac{\pi^*}{1-\pi^*} \Big/ \frac{\widetilde{\pi}_m}{1-\widetilde{\pi}_m}, \tag{D.4}$$

where

$$\mathcal{I} \equiv \frac{\Psi^*(1-\Psi^*)}{\Psi^{**}(1-\Psi^{**})}$$

measures the aggregate informativeness of reports. This is because in such equilibria, one report is sufficient to convict the principal, and therefore, the evaluator is indifferent between $s = 1$ and $s = 0$ when there is exactly one report.

Comparing (D.1) to (D.2), we know that $\omega^* - \omega^{**} > b$. Rewrite (D.1) and (D.2) as:

$$\frac{\omega^*}{c} = -\frac{1-\Psi^*}{\Psi^* + (1-\Psi^*)\frac{q}{1-q}} \tag{D.5}$$

and

$$\frac{\omega^{**}+b}{c} = -\frac{1-\Psi^{**}}{\Psi^{**} + (1-\Psi^{**})\frac{q}{1-q}}, \tag{D.6}$$

notice that the RHS is bounded within $[-\frac{1-q}{q}, 0]$ and is continuous with respect to q. For every $c > 0$, both (D.4) and (D.5) admit a unique solution. Moreover, for every $\overline{q} < 1$ and $A \in \mathbb{R}_+$, there exists $\overline{c} > 0$ such that for every $c > \overline{c}$ and $q \in [1/2, \overline{q}]$, the solution satisfies $|\omega^*| > A$. Since $\omega^* - \omega^{**} > b$, we know that the aggregate reporting informativeness in the $\delta \to 1$ limit also goes to infinity, that is,

$$\lim_{c\to\infty} \lim_{\delta\to 1} \frac{\delta\Phi(\omega^*) + (1-\delta)\alpha}{\delta\Phi(\omega^{**}) + (1-\delta)\alpha} = \infty.$$

Therefore, when c is large enough and by setting L to be in an open set consisting of the values of (D.3) when $q \in [1/2, \overline{q}]$, the informativeness ratio $\mathcal{I}$ goes to infinity and the equilibrium probability of crime converges to 0.

D.2 Complementarity in Principal's Actions

For every (c, δ), let $(\omega_0^*, \omega_0^{**})$ be the unique solution to

$$\frac{\omega_0^*}{c} = \Psi_0^* - 1 \quad \text{and} \quad \frac{\omega_0^{**}+b}{c} = \Psi_0^{**} - 1,$$

where $\Psi_0^* \equiv \delta\Phi(\omega_0^*) + (1-\delta)$ and $\Psi_0^{**} \equiv \delta\Phi(\omega_0^{**}) + (1-\delta)$. Let

$$L_0 \equiv \frac{2}{\Psi_0^* - \Psi_0^{**}}. \tag{D.7}$$

According to the analysis in subsection D.1, we know that $(\omega_0^*, \omega_0^{**})$ are the agents' reporting cutoffs when L_0 is the punishment to the convicted and the conviction probabilities are given by $q(1,1) = 1$, $q(0,1) = q(1,0) = 1/2$ and $q(0,0) = 0$. We show the following proposition:

Proposition D.1. *There exists an open neighborhood of L_0 such that for every L belonging to this neighborhood, the value of (C.1) is non-positive in every equilibrium.*

The proof of Proposition D.1 are in subsections D.3 and D.4. It shows that whenever there exists a monotone-responsive equilibrium in which the principal's choices of θ_1 and θ_2 are strategic substitutes, then L needs to be strictly larger than L_0. This together with the existence of equilibrium implies that the principal's decisions are strategic complements in all equilibria. The proof considers two cases separately, depending on the conviction probabilities.

D.3 Case 1: $\mathbf{q}(\mathbf{1},\mathbf{0})$ or $\mathbf{q}(\mathbf{0},\mathbf{1})$ is Strictly Positive

Suppose towards a contradiction that there exists a monotone-responsive equilibrium in which either $q(0,1)$ or $q(1,0)$ is strictly positive or both. For notation simplicity, let $q_1 \equiv q(1,0)$ and $q_2 \equiv q(0,1)$. For $i \in \{1,2\}$, let p_i be the probability with which $\theta_i = 0$. Let ω_i^* and ω_i^{**} be the agent i's reporting cutoffs, with expressions given by:

$$\omega_i^* = -c\frac{(1-\Psi_j^{**})(1-q_i)}{q_i + \Psi_j^{**}(1-q_1-q_2)} \tag{D.8}$$

and

$$\omega_i^{**} = -b - c\frac{(1-X_j)(1-q_i)}{q_i + X_j(1-q_1-q_2)} \tag{D.9}$$

where $j \equiv 3-i$ and

$$X_i \equiv \frac{1-p_1-p_2}{1-p_i}\Psi_i^{**} + \frac{p_j}{1-p_i}\Psi_i^*.$$

It is without loss of generality to focus on equilibria in which the principal's cost of abusing the two agents are equal. This leads to the indifference condition:

$$L = \frac{1}{(\Psi_1^* - \Psi_1^{**})\Big(\Psi_2^{**}(1-q_1-q_2)+q_1\Big)} = \frac{1}{(\Psi_2^* - \Psi_2^{**})\Big(\Psi_1^{**}(1-q_1-q_2)+q_2\Big)}. \tag{D.10}$$

Without loss of generality, we assume $q_1 \leq q_2$. Since $q_1 + q_2 \leq 1$, we know that

$$L = \frac{1}{(\Psi_1^* - \Psi_1^{**})\Big(\Psi_2^{**}(1-q_1-q_2)+q_1\Big)} \geq \frac{2}{\Psi_1^* - \Psi_1^{**}}.$$

In what follows, we show that

$$\frac{2}{\Psi_1^* - \Psi_1^{**}} > L_0$$

or equivalently,

$$\Psi_0^* - \Psi_0^{**} > \Psi_1^* - \Psi_1^{**}. \tag{D.11}$$

According to the expression of ω_1^* in (D.8), we know that:

$$\omega_1^* = -c\frac{(1-\Psi_2^{**})(1-q)}{q+\Psi_2^{**}(1-2q)} \leq -c(1-\Psi_1^{**}) \leq c(\Psi_1^* - 1).$$

Therefore, ω_1^* is strictly below the unique solution of the equation:

$$\omega = c\Big(\underbrace{\delta\Phi(\omega) + (1-\delta)\alpha}_{\equiv\Psi(\omega)} - 1\Big)$$

which equals to ω_0^*. Furthermore, since $\omega_0^* - \omega_0^{**} > b > \omega_1^* - \omega_1^{**}$, one can obtain (D.11).

D.4 Case 2: $\mathbf{q(1,0) = q(0,1) = 0}$

Given that we have already shown in the proof of Theorem 1 that symmetric equilibria is without loss of generality when $q(0,0) = q(1,0) = q(0,1) = 0$, we focus on symmetric equilibria. Let $q \equiv q(1,1) \in (0,1]$. Let ω_1^* and ω_1^{**} be the agents' reporting cutoffs, which are the same across agents. The expressions for the cutoffs are the same as those for ω_m^* and ω_m^{**}, which are given by (3.8) and (3.9) in the main text respectively. The principal's indifference condition is given by:

$$\frac{1}{L} = \Psi_1^{**}(\Psi_1^* - \Psi_1^{**}).$$

To show $L > L_0$, one only needs to show that:

$$\Psi_0^* - \Psi_0^{**} > \Psi_1^{**}(\Psi_1^* - \Psi_1^{**}). \tag{D.12}$$

According to (3.8) in the main text, we have:

$$\omega_1^* = c - \frac{c}{q\Psi_1^*} \leq c - \frac{c}{\Psi_1^*} \leq c(\Psi_1^* - 1).$$

Similar to the previous case, we know that ω_1^* is strictly below ω_0^*. Since $\omega_0^* - \omega_0^{**} > b > \omega_1^* - \omega_1^{**}$, we know that

$$\Psi_0^* - \Psi_0^{**} > \Psi_1^* - \Psi_1^{**}. \tag{D.13}$$

This in turn implies (D.12).

E Proof of Proposition 6

Recall that $\Delta_1 \equiv t_1(1,0) - t_1(0,0)$ and $\Delta_2 \equiv t_2(0,1) - t_2(0,0)$. Without loss of generality, let $t_1(0,0) = t_2(0,0) = 0$. Then $t_1(1,0) = -t_2(1,0) = \Delta_1$, $-t_1(0,1) = t_2(0,1) = \Delta_2$. Let $t_1(1,1) = T$, then $t_2(1,1) = -T$. The two players' reporting cutoffs are then given by:

$$\omega_1^* = c + \frac{1}{q}(T + \Delta_2 - \Delta_1) + \frac{1}{q\Psi_2^{**}}(\Delta_1 - c), \tag{E.1}$$

$$\omega_1^{**} = -b + c + \frac{1}{q}(T + \Delta_2 - \Delta_1) + \frac{1}{qQ_2}(\Delta_1 - c), \tag{E.2}$$

$$\omega_2^* = c + \frac{1}{q}(\Delta_1 - \Delta_2 - T) + \frac{1}{q\Psi_1^{**}}(\Delta_2 - c), \tag{E.3}$$

$$\omega_2^{**} = -b + c + \frac{1}{q}(\Delta_1 - \Delta_2 - T) + \frac{1}{qQ_1}(\Delta_2 - c). \tag{E.4}$$

We consider three cases separately, depending on the signs of $\Delta_1 - c$ and $\Delta_2 - c$.

E.1 Case 1: $\Delta_1, \Delta_2 \geq c$

Suppose $\Delta_1, \Delta_2 \geq c$, then

$$\omega_1^{**} \geq -b + c + \frac{1}{q}(T + \Delta_2 - \Delta_1) \text{ and } \omega_2^{**} \geq -b + c + \frac{1}{q}(\Delta_1 - \Delta_2 - T)$$

Therefore,

$$\omega_1^{**} + \omega_2^{**} \geq -2b + 2c, \tag{E.5}$$

which implies that $\max\{\omega_1^{**}, \omega_2^{**}\} \geq -b + c$. Since $\mathcal{I}_m = \min\{\mathcal{I}_1, \mathcal{I}_2\}$, we know that

$$\mathcal{I}_m \leq \frac{1}{\delta\Phi(-b+c) + (1-\delta)\alpha} \leq \frac{1}{\delta\Phi(-b+c)}. \tag{E.6}$$

The above inequality establishes an upper bound on report informativeness.

E.2 Case 2: $\Delta_1, \Delta_2 < c$

Let

$$X \equiv \frac{1}{q}(\Delta_1 - \Delta_2 - T).$$

Without loss of generality, assume $X \geq 0$. Let $\beta \in (0,1)$ be the probability with which agent 1 is abused conditional on the principal being guilty. The expressions for the two cutoffs imply that:

$$\frac{\omega_2^* - c - X}{\omega_2^{**} + b - c - X} = \frac{(1-\beta)l^*\mathcal{I}_1 + \mathcal{I}_m}{(1-\beta)l^* + \mathcal{I}_m} \quad \text{and} \quad \frac{\omega_1^* - c + X}{\omega_1^{**} + b - c + X} = \frac{\beta l^*\mathcal{I}_2 + \mathcal{I}_m}{\beta l^* + \mathcal{I}_m}. \tag{E.7}$$

We start with the following Lemma:

Lemma E.1. *There exists a function $\epsilon : \mathbb{R}_+ \times [0,1] \to \mathbb{R}_+$ such that for every $\eta \in (0,1)$, if $\beta \leq 1-\eta$ and $\omega_2^* < -M$, then $\mathcal{I}_m < 1 + \epsilon(M,\eta)$.*

Proof of Lemma E.1: Since $\Delta_2 < c$, we know that $\omega_2^* - \omega_2^{**} < b$. Since $X \geq 0$, $\omega_2^* - c - X < 0$ and $\omega_2^{**} + b - c - X < 0$, we have:

$$\frac{\omega_2^* - c - X}{\omega_2^{**} + b - c - X} < \frac{\omega_2^* - c}{\omega_2^{**} + b - c} \leq \frac{M + c}{M + c - b} = 1 + \frac{b}{M + c - b}$$

On the other hand, since $\mathcal{I}_1 \geq \mathcal{I}_m \geq 1$, we know that:

$$1 + \frac{b}{M + c - b} \geq \frac{(1-\beta)l^*\mathcal{I}_1 + \mathcal{I}_m}{(1-\beta)l^* + \mathcal{I}_m} \geq \frac{(1-\beta)l^*\mathcal{I}_m + \mathcal{I}_m}{(1-\beta)l^* + \mathcal{I}_m} \geq \frac{\eta l^*\mathcal{I}_m + \mathcal{I}_m}{\eta l^* + \mathcal{I}_m}$$

This places an upper bound on $\mathcal{I}_m$, which converges to 1 as $M \to -\infty$. □

Lemma E.1 implies that for every $\eta \in (0,1)$, if $\beta \leq 1-\eta$, the informativeness of report is bounded from above by:

$$\max_{M \in \mathbb{R}_+} \left\{ \min\{1 + \epsilon(M,\eta), \frac{1}{\Phi(-M-b)}\} \right\}. \tag{E.8}$$

Expression (E.8) is bounded from above for every given η. Therefore, in order to establish a uniform upper bound on $\mathcal{I}_m$, we only need to show that unbounded informativeness cannot arise when α is close to or equal to 1. That is to say, it is without loss to consider cases in which $\beta \geq 1/2$. Therefore, agent 1 is abused with strictly positive probability, which implies that $\mathcal{I}_m = \mathcal{I}_1 \leq \mathcal{I}_2$.

Suppose towards a contradiction that for every $\overline{\mathcal{I}} > 0$, there exists $\{\Delta_1, \Delta_2, X\}$ under which there exists an equilibrium in which $\mathcal{I}_m > \overline{\mathcal{I}}$. In what follows, we consider two subcases separately.

Subcase 1: $\boldsymbol{\omega_1^* - \omega_1^{**} \geq \omega_2^* - \omega_2^{**}}$ Since $\mathcal{I}_1 \leq \mathcal{I}_2$, we know that $\omega_1^{**} \geq \omega_2^{**}$. According to the assumption that $\omega_1^* - \omega_1^{**} \geq \omega_2^* - \omega_2^{**}$, we know that $\omega_1^* \geq \omega_2^*$. Therefore:

$$X + \frac{\Delta_2 - c}{q\Psi_1^{**}} \leq -X + \frac{\Delta_1 - c}{q\Psi_2^{**}},$$

or equivalently,

$$X \leq \frac{1}{2}\Big(\frac{|c-\Delta_1|}{q\Psi_2^{**}} - \frac{|c-\Delta_2|}{q\Psi_1^{**}}\Big) \tag{E.9}$$

On the other hand,

$$\omega_2^* - \omega_2^{**} = b - \frac{|c-\Delta_2|}{q\Psi_1^{**}} \cdot \frac{\beta(\mathcal{I}_1 - 1)l^*}{\mathcal{I}_m + \beta l^* \mathcal{I}_1} > 0,$$

which implies that for every $\varepsilon > 0$, there exists $\mathcal{I}^*$ such that whenever $\mathcal{I}_m > \mathcal{I}^*$,

$$\frac{|c-\Delta_2|}{q\Psi_1^{**}} \leq b\frac{1+\beta l^*}{\beta l^*} + \varepsilon.$$

Since $\beta \geq 1/2$ and the RHS is decreasing in β, we know that when $\mathcal{I}_m$ is sufficiently large,

$$X \leq \frac{b}{2} \cdot \frac{1 + l^*/3}{l^*/3}. \tag{E.10}$$

Given this uniform upper bound on X, we know that as $\omega_1^* \to -\infty$,

$$\frac{\omega_1^* - c + X}{\omega_1^{**} + b - c + X} \to 1.$$

The second part of (E.7) together with $\beta \geq 1/2$ imply that $\mathcal{I}_2$ is uniformly bounded from above as $\omega_1^* \to -\infty$. This contradicts the assumption that $\mathcal{I}_m$ is unbounded.

Subcase 2: $\boldsymbol{\omega_1^* - \omega_1^{**} < \omega_2^* - \omega_2^{**}}$ Since $\beta \geq 1/2$, the distance between ω_1^* and ω_1^{**} is at most b and

$$\frac{\omega_1^* - c + X}{\omega_1^{**} + b - c + X} = \frac{\beta l^* \mathcal{I}_2 + \mathcal{I}_m}{\beta l^* + \mathcal{I}_m},$$

if $\mathcal{I}_m$ is unbounded, then $\omega_1^* - c + X$ is bounded from below. That is, there exists $A \in \mathbb{R}_+$ such that

$$|\omega_1^* - c + X| = \frac{|c-\Delta_1|}{q\Psi_2^{**}} \leq A. \tag{E.11}$$

Since $\omega_1^* - \omega_1^{**} < \omega_2^* - \omega_2^{**}$, we know that when $\mathcal{I}_m$ is sufficiently large,

$$\frac{|c-\Delta_1|}{q\Psi_2^{**}} \cdot \frac{1-\beta}{1+(1-\beta)l^*} \leq \frac{|c-\Delta_2|}{q\Psi_1^{**}} \cdot \frac{\beta}{1+\beta l^*}. \tag{E.12}$$

Therefore,

$$\frac{|c-\Delta_2|}{q\Psi_1^{**}} \leq \frac{|c-\Delta_1|}{q\Psi_2^{**}} \cdot \frac{1-\beta}{\beta} \cdot (1+l^*) \leq A(1+l^*)\frac{1-\beta}{\beta}. \tag{E.13}$$

According to Lemma E.1, $\beta \to 1$ and $\omega_1^* \to -\infty$ are required when $\mathcal{I}_m \to \infty$. Therefore, $X \to \infty$ and

$$\frac{|c - \Delta_2|}{q\Psi_1^{**}} \to 0.$$

But according to the expression that

$$\omega_2^* = c + X + \frac{|c - \Delta_2|}{q\Psi_1^{**}},$$

we know that ω_2^* is strictly positive when $\mathcal{I}_m$ is sufficiently large. Therefore $\omega_2^{**} \geq \omega_2^* - b \geq -b$ and therefore, $\mathcal{I}_m \leq \mathcal{I}_2 \leq 1/\Phi(-b)$, leading to a contradiction.

E.3 Case 3: $\Delta_1 \geq c$ and $\Delta_2 < c$

Define X in the same way as in the previous subsection. If $X \leq 0$, then

$$\omega_1^{**} \geq -b + c$$

which implies that $\mathcal{I} \leq 1/\Phi(-b + c)$.

If $X > 0$, then

$$\frac{\omega_2^* - c - X}{\omega_2^{**} + b - c - X} \to 1$$

as $\omega_2^* \to -\infty$. Since

$$\frac{\omega_2^* - c - X}{\omega_2^{**} + b - c - X} = \frac{(1-\beta)l^*\mathcal{I}_1 + \mathcal{I}_m}{(1-\beta)l^* + \mathcal{I}_m},$$

we know that in order for $\mathcal{I}_m \to \infty$, we need $\omega_2^* \to -\infty$ and $\beta \to 1$. Therefore, it is without loss to consider situations in which

$$\beta \geq \overline{\beta} \equiv \max\{1 - 1/l^*, 1/2\}.$$

When $\beta \geq \overline{\beta}$, we know that $\mathcal{I}_m = \mathcal{I}_1 \leq \mathcal{I}_2$. Since $\omega_1^* - \omega_1^{**} \geq b > \omega_2^* - \omega_2^{**}$, we know that $\omega_2^{**} < \omega_1^{**}$, which further implies that $\omega_2^* < \omega_1^*$. This implies that

$$X + \frac{\Delta_2 - c}{q\Psi_1^{**}} \leq -X + \frac{\Delta_1 - c}{q\Psi_2^{**}}$$

which is equivalent to:

$$X \leq \frac{1}{2}\Big(\frac{|\Delta_1 - c|}{q\Psi_2^{**}} + \frac{|c - \Delta_2|}{q\Psi_1^{**}}\Big).$$

Since $\omega_2^* - \omega_2^{**} > 0$, we know that for every $\mathcal{I}_m$ above some threshold,

$$\frac{|c-\Delta_2|}{q\Psi_1^{**}} \leq b\frac{1+\widetilde{\beta}l^*}{\widetilde{\beta}l^*},$$

where $\widetilde{\beta} \equiv \overline{\beta}/2$. Therefore

$$\begin{aligned} \omega_1^{**} &= -b+c-X+\frac{\Delta_1-c}{q\Psi_2^{**}}\cdot\frac{(1-\beta)l^*+\mathcal{I}_m}{\mathcal{I}_m+(1-\beta)l^*\mathcal{I}_2} \\ &\geq -b+c-\frac{|c-\Delta_2|}{2q\Psi_1^{**}}-\frac{|\Delta_1-c|}{2q\Psi_2^{**}}+\frac{|\Delta_1-c|}{q\Psi_2^{**}}\cdot\frac{(1-\beta)l^*+\mathcal{I}_m}{\mathcal{I}_m+(1-\beta)l^*\mathcal{I}_2} \\ &\geq -b+c-b\frac{1+\widetilde{\beta}l^*}{2\widetilde{\beta}l^*}+\frac{|\Delta_1-c|}{q\Psi_2^{**}}\Big(\frac{(1-\beta)l^*+\mathcal{I}_m}{\mathcal{I}_m+(1-\beta)l^*\mathcal{I}_2}-\frac{1}{2}\Big) \end{aligned} \tag{E.14}$$

The coefficient

$$\frac{(1-\beta)l^*+\mathcal{I}_m}{\mathcal{I}_m+(1-\beta)l^*\mathcal{I}_2}-\frac{1}{2}$$

is strictly positive when $\beta \geq \overline{\beta}$ and $\mathcal{I}_m$ is sufficiently large. Therefore (E.14) implies that

$$\omega_1^{**} \geq \overline{\omega}_1^{**} \equiv -b+c-b\frac{1+\widetilde{\beta}l^*}{2\widetilde{\beta}l^*} \tag{E.15}$$

which further implies the following upper bound on $\mathcal{I}_m$:

$$\mathcal{I}_m = \mathcal{I}_1 \leq \Phi(\overline{\omega}_1^{**})^{-1}.$$

F Proof of Proposition 7

First, we show that $\omega_n^* - \omega_n^{**} \in (0, b)$. Suppose towards a contradiction that $\omega_n^* - \omega_n^{**} \leq 0$, then the comparison between (5.3) and (5.4) suggests that $Q_{0,n} \geq Q_{1,n}$. Plugging this into (5.1) and (5.2), it implies that $\omega_n^* \geq \omega_n^{**} + b$. On the other hand, since $\omega_n^* - \omega_n^{**} > 0$, we know that $Q_{0,n} < Q_{1,n}$. The expressions for the cutoffs the imply that $\omega_n^* - \omega_n^{**} < b$, leading to a contradiction.

Next, we show that $\mathcal{I}_n \to 1$ as $\omega_n^* \to -\infty$. To see this, apply the expressions of ω_n^* and ω_n^{**} in (5.1) and (5.2), we have:

$$\frac{|\omega_n^*-c|}{|\omega_n^{**}+b-c|} = \frac{Q_{1,n}}{Q_{0,n}} = \frac{(n-1)l^*}{(n-1)l^*+n\mathcal{I}_n}\mathcal{I}_n + \frac{n\mathcal{I}_n}{(n-1)l^*+n\mathcal{I}_n}. \tag{F.1}$$

Since $\omega_n^* - \omega_n^{**} \in (0, b)$, the LHS converges to 1 as $\omega_n^* \to -\infty$, which implies that the RHS also converges to 1. This can only be the case when $\mathcal{I}_n \to 1$.

In the last step, we show that $\omega_n^* \to -\infty$ as $L \to \infty$. Suppose towards a contradiction that there exists an interior accumulation point $\omega^* \in \mathbb{R}_-$, then as the LHS of (5.6) converges to 0 when $L \to \infty$, we know that either $q_n \to 0$ or $\Phi(\omega_n^*) - \Phi(\omega_n^{**}) \to 0$ or both. The latter implies that $\omega_n^* - \omega_n^{**} \to 0$ as $\omega_n^* \to \omega^*$ and ω_n^* is interior.

Suppose $q_n \to 0$, then $\omega_n^* \to -\infty$ according to (5.1). Next, suppose towards a contradiction that q_n is bounded away from 0 along some subsequence, i.e strictly greater than some $\underline{q} > 0$, then we know that $\omega_n^* - \omega_n^{**} \to 0$. Subtracting the expression of ω_n^* from that of ω_n^{**}, we obtain:

$$\frac{q_n}{c}\Big(\omega_n^* - (\omega_n^{**} + b)\Big) = \frac{(n-1)l^*}{(n-1)l^* + n}\Big\{\frac{1}{\delta\Phi(\omega_n^*) + (1-\delta)\alpha} - \frac{1}{\delta\Phi(\omega_n^{**}) + (1-\delta)\alpha}\Big\}. \tag{F.2}$$

The absolute value of the LHS is no less than $\underline{q}b/c$ in the limit as $\omega_n^* - \omega_n^{**} \to 0$. The absolute value of the RHS converges to 0 as $\Phi(\omega_n^*) - \Phi(\omega_n^{**}) \to 0$, leading to a contradiction. This suggests that $\omega_n^* \to -\infty$ in every equilibrium as $L \to \infty$.

The three parts together imply that as $L \to \infty$, ω_n^* and ω_n^{**} go to $-\infty$, the aggregate informativeness of reports, $\mathcal{I}_n$, converges to 1 and the equilibrium probability of crime $\widetilde{\pi}_n$ converges to π^*.

G Proof of Proposition 9

We start from listing the sufficient conditions for equilibria in which $q(0,0) = q(1,0) = q(0,1) = 0$ and $q(1,1) \in (0,1)$. Since the posterior attaches to $\theta_1\theta_2 = 0$ reaches π^* after observing two reports, the relationship between the equilibrium probability of crime $\widetilde{\pi}$ and the informativeness of reports $\mathcal{I}$ is given by (5.11), which according to (5.10) can be rewritten as:

$$\widetilde{\pi} = \frac{l^* + \epsilon R - \epsilon R^2}{R + l^*}. \tag{G.1}$$

The expressions for the cutoffs are given by:

$$\omega^* = -c\frac{1 - qQ_0}{qQ_0} \quad \text{and} \quad \omega^{**} = -b - c\frac{1 - qQ_1}{qQ_1} \tag{G.2}$$

where $Q_i = \beta_i\Psi^* + (1-\beta_i)\Psi^{**} = (\beta_i + \frac{1-\beta_i}{R})\Psi^*$ for $i \in \{0,1\}$ with

$$\beta_0 \equiv \frac{2\epsilon(R + l^*)}{2R + l^* + \epsilon(l^* + R^2)} \tag{G.3}$$

and

$$\beta_1 \equiv \frac{l^* - \epsilon(R^2 + l^*)}{l^* + 2R - \epsilon(2R - R^2 + l^*)}. \tag{G.4}$$

Rewrite (G.2) by plugging into the definition of R, we have:

$$-\frac{\omega^* - c}{c} = \frac{1}{\xi_0(R,\epsilon)\Psi^* q} \text{ and } -\frac{\omega^{**} + b - c}{c} = \frac{1}{\xi_1(R,\epsilon)\Psi^* q} \tag{G.5}$$

where

$$\xi_0(R,\epsilon) = \beta_0 + (1-\beta_0)\frac{1}{R} \text{ and } \xi_1(R,\epsilon) = \beta_1 + (1-\beta_1)\frac{1}{R}. \tag{G.6}$$

Notice that both ξ_0 and ξ_1 are continuous functions with values no more than 1. The values of both functions equal to 1 when $R = 1$ and $\epsilon = 0$. More importantly, the values of ξ_0 and ξ_1 are fixed once we fix R and ϵ.

Next, consider the following mapping $f \equiv (f_1, f_2) : [0,1] \times [0,1] \to [0,1] \times [0,1]$:

$$f_1(\Psi^{**}, q) \equiv \Psi\Big(-b + c - \frac{c}{q\xi_1 R\Psi^{**}}\Big) \tag{G.7}$$

$$f_2(\Psi^{**}, q) \equiv \min\Big\{1, \frac{c}{\xi_0 \Psi^*(c - \omega^*)}\Big\}, \tag{G.8}$$

where for given Ψ^{**}, Ψ^* (and hence ω^*) is pinned down via $\Psi^*/\Psi^{**} = R$. Since f is continuous, the Brouwer's fixed point theorem implies the existence of a fixed point. In what follows, we show that $q = 1$ cannot be part of any fixed point of f when ϵ is close to 0 and R is close to 1. For this purpose, we need to show that:

$$\frac{c}{\xi_0 \Psi^*(c - \omega^*)} < 1 \tag{G.9}$$

for every Ψ^{**} solving the equation:

$$-\frac{\omega^{**} + b - c}{c} = \frac{1}{R\xi_1\Psi^{**}}. \tag{G.10}$$

To see this, first, (G.10) admits at least one solution as $\Psi^{**} \geq (1-\delta)\alpha$. Second, (G.9) is equivalent to:

$$\frac{\xi_1}{\xi_0}\frac{|\omega^{**}| + |c| - b}{|\omega^*| + |c|} < 1. \tag{G.11}$$

The above inequality holds as first, $\frac{\xi_1}{\xi_0} \to 1$ as $R \to 1$ and $\epsilon \to 0$; and second, $|\omega^*| > |\omega^{**}| - b$ whenever $\beta_0 < \beta_1$, the latter is true when ϵ is small enough.

Since every fixed point features $q \in (0,1)$, the fixed point of f is also the level of (Ψ^*, Ψ^{**}, q) in one of the monotone-responsive equilibrium. One can then pin down L_h via:

$$\frac{1}{\delta L_h} = \delta(\Phi(\omega^*) - \Phi(\omega^{**}))(\delta\Phi(\omega^{**}) + 1 - \delta). \tag{G.12}$$

Let $\widetilde{\pi}$ be given by (G.1). We know that $(\omega^*, \omega^{**}, q, \widetilde{\pi})$ is an equilibrium under (L_h, L_l, ϵ).

H Extensions

In this Appendix, we study two sets of extensions. In subsection H.1, we examine the robustness of our insights to alternative specifications of the agents' payoffs. In subsection H.2, we show that our results are not sensitive to the mechanical types' strategies.

H.1 Alternative Specifications on Agents' Payoffs

Social Preferences: We show that the our main insights are robust in a variant of the baseline model where agents have social preferences. Let agent i's payoff be:

$$\Big\{\underbrace{\omega_i + b\Big(\gamma\theta_i + (1-\gamma)\Pi_{j=1}^n \theta_j\Big) - ca_i}_{\text{payoff when the principal is acquitted}}\Big\} s. \tag{H.1}$$

Intuitively, an agent's payoff depends not only on whether the principal has committed a crime against him or not, but also on whether the principal is guilty or innocent. When $\gamma = 1$, it coincides with the baseline model.

In this setting, since agent 1 does not observe θ_2 and agent 2 does not observe θ_1, they decide whether to file reports or not based on their beliefs about $\Pi_{j=1}^n \theta_j$ after observing their own θ_i. As a result, an agent's strategy in every equilibrium is still characterized by two cutoffs: ω^* when $\theta_i = 0$ and ω^{**} when $\theta_i = 1$. Whether the principal's incentives to commit crimes are strategic complements or substitutes is still determined by the sign of (C.1). When L is large, we need two reports to convict the principal with positive probability and the principal's decisions to commit crimes are strategic substitutes. This results in negative correlation between the agents' private information.

Different from the baseline model, the expressions for an agent's reporting cutoffs are given by:

$$\omega^* = c - \frac{c}{q_m \Psi^{**}} \tag{H.2}$$

and

$$\omega^{**} = -b + c + \frac{b\Psi^*(1-\beta)(1-\gamma)}{\beta\Psi^{**} + (1-\beta)\Psi^*} - \frac{c}{q_m\Big(\beta\Psi^{**} + (1-\beta)\Psi^*\Big)}, \tag{H.3}$$

where $q_m \in (0,1)$ is the probability of conviction when there are two reports, $\Psi^* \equiv \delta\Phi(\omega^*) + (1-\delta)\alpha$, $\Psi^{**} \equiv \delta\Phi(\omega^{**}) + (1-\delta)\alpha$ and β is the probability of $\theta_j = 1$ conditional on $\theta_i = 1$. Compare (H.3) to the expression for ω_m^{**} in the baseline model, the novel term is:

$$\frac{b\Psi^*(1-\beta)(1-\gamma)}{\beta\Psi^{**} + (1-\beta)\Psi^*}, \tag{H.4}$$

which measures the impact of social preferences on an agent's equilibrium reporting strategy. This term equals to 0 when $1-\gamma$, the weight an agent attaches to $\Pi_{j=1}^n \theta_j$, equals to 0. Importantly,

$$0 \le \frac{b\Psi^*(1-\beta)(1-\gamma)}{\beta\Psi^{**} + (1-\beta)\Psi^*} \le (1-\gamma)b. \tag{H.5}$$

Using the same logic as Lemma 3.2 in the main text, we can show that $\omega^* - \omega^{**} \in [0, b]$. According to (H.2) and (H.3), we have:

$$\frac{|\omega^* - c|}{\left|\omega^{**} + b - c - \dfrac{b\Psi^*(1-\beta)(1-\gamma)}{\beta\Psi^{**} + (1-\beta)\Psi^*}\right|} = \frac{\beta\Psi^{**} + (1-\beta)\Psi^*}{\Psi^{**}} = \beta + (1-\beta)\frac{\Psi^*}{\Psi^{**}} = \frac{(l^*+2)\mathcal{I}_m}{l^* + 2\mathcal{I}_m}, \tag{H.6}$$

where $\mathcal{I}_m \equiv \Psi^*/\Psi^{**}$ measures the aggregate informativeness of reports.

As $L \to \infty$, one can show that both ω^* and ω^{**} go to $-\infty$ using the same argument as in subsection 3.4 of the main text. Since the difference between the denominator and the numerator of the LHS of (H.6) is at most b, the value of (H.6) converges to 1. This implies that $\mathcal{I}_m$ converges to 1, namely, the agents' reports are arbitrarily uninformative and the equilibrium probability of crime approaches π^*.

Ex Post Evidence & Punishing False Accusers: When an innocent principal is convicted, i.e. $\theta_1 = ... = \theta_n = 1$ and $s = 0$, then some ex post evidence arrives with probability p^* that reveals his innocence. After that, every agent who files a false accusation is penalized by k. Our analysis goes through in the same way and all our qualitative results remain robust as the presence of ex post evidence is equivalent to an increase in b. To see this, agent i's indifference condition when $\theta_i = 1$ is now given by:

$$q_m Q_1(\omega_i + b) = -c(1 - q_m Q_1) - q_m Q_1 p^* k. \tag{H.7}$$

The expression for the cutoff is then given by

$$\omega_m^{**} \equiv -b - p^*k - c\frac{1 - q_m Q_1}{q_m Q_1} = -b - p^*k + c - \frac{c}{q_m Q_1}. \tag{H.8}$$

The above expression is qualitatively the same as that in (2.9) of the main text except one needs to replace b with $\widetilde{b} \equiv b + p^*k$.

H.2 Alternative Mechanical Types

In this subsection, we examine the robustness of our findings against alternative specifications of the mechanical types' strategies. We allow the mechanical types' reports to be informative about the principal's innocence. We show that when commitment types are rare and the principal's loss from being convicted is sufficiently large, the informativeness of reports vanishes to 1 and the probability of crime converges to π^* as in the baseline model. This confirms the robustness of our findings. We focus on the comparison between the single-agent benchmark and the two-agent scenario.

H.2.1 Model & Result

Consider the following modification of the baseline model. With probability $\delta \in (0, 1)$, the agent is a strategic type maximizes payoff function given by (2.5) in the main text. With probability $1 - \delta$, the agent is a mechanical type whose reporting cutoff is $\overline{\omega}$ when $\theta_i = 0$ and $\underline{\omega}$ when $\theta_i = 1$. We assume that both $\overline{\omega}$ and $\underline{\omega}$ are finite with $\overline{\omega} \geq \underline{\omega}$, that is, the mechanical type's report could be informative about θ.[3]

When there is only one agent, his reporting cutoffs ω_s^* and ω_s^{**} are given by (3.1) and (3.2), respectively. The probability with which the principal is convicted after one report is q_s, with $(q_s, \omega_s^*, \omega_s^{**})$ satisfying:

$$q_s\Big(\delta(\Phi(\omega_s^*) - \Phi(\omega_s^{**})) + (1-\delta)(\Phi(\overline{\omega}) - \Phi(\underline{\omega}))\Big) = 1/L. \tag{H.9}$$

One can show that when $\delta \to 1$ and L is larger than some cutoff $L(\delta)$, the informativeness of report:

$$\mathcal{I}_s \equiv \frac{\delta\Phi(\omega_s^*) + (1-\delta)\Phi(\overline{\omega})}{\delta\Phi(\omega_s^{**}) + (1-\delta)\Phi(\underline{\omega})}$$

converges to ∞, namely, the agent's report becomes arbitrarily informative in the limit.

[3] Our analysis also applies when mechanical types are using arbitrary strategies contingent on (θ_i, ω_i), as long as conditional on each realization of θ_i, the probability with which the mechanical type reports is interior, and moreover, this conditional probability is weakly higher when $\theta_i = 0$ compared to $\theta_i = 1$.

In the two-agent case, for every $i \in \{1,2\}$, agent i's probability of filing a report is $\Psi^* \equiv \delta\Phi(\omega_m^*) + (1-\delta)\Phi(\overline{\omega})$ conditional on $\theta_i = 0$; his probability of filing a report is $\Psi^{**} \equiv \delta\Phi(\omega_m^{**}) + (1-\delta)\Phi(\underline{\omega})$ conditional on $\theta_i = 1$. The strategic agent's reporting cutoffs are given by:

$$\omega_m^* \equiv c - \frac{c}{q_m \Psi^{**}} \quad \text{and} \quad \omega_m^{**} \equiv -b + c - \frac{c}{q_m\Big(\beta\Psi^{**} + (1-\beta)\Psi^*\Big)}. \tag{H.10}$$

Let $\mathcal{I}_m \equiv \Psi^*/\Psi^{**}$. When L is large enough, the conviction probabilities in every monotone-responsive equilibrium satisfies $q(0,0) = q(0,1) = q(1,0) = 0$ and $q(1,1) \in (0,1)$. Therefore, the expressions for β and $1-\beta$ remain the same as in (3.15). The distance between the two cutoffs is given by:

$$\omega_m^* - \omega_m^{**} = b - \frac{c}{q_m}\frac{(1-\beta)(\mathcal{I}_m - 1)}{\Psi^{**}(\beta + (1-\beta)\mathcal{I}_m)} = b - \frac{c}{q_m\Psi^{**}}\frac{l^*}{2+l^*}\frac{\mathcal{I}_m - 1}{\mathcal{I}_m}. \tag{H.11}$$

One can then show that $\omega_m^* - \omega_m^{**} < b$. This is because for $\omega_m^* - \omega_m^{**}$ to be greater or equal to b, we need $\mathcal{I}_m \le 1$ which can only be true when $\omega_m^* \le \omega_m^{**}$, leading to a contradiction.

Different from the baseline model, when mechanical types' reports are informative about the principal's innocence, the strategic types' coordination motives can *reverse* the ordering between the two cutoffs. That is to say, ω_m^* can be strictly smaller than ω_m^{**} in equilibrium. As a result, the argument that shows $\mathcal{I}_m \to 1$ when $\omega_m^* \to -\infty$ in Lemma 3.3 in the main text no longer applies. This is because in principle, ω_m^* could be much smaller than ω_m^{**}, so the ratio between the absolute values in (3.17) can converge to something strictly above 1 as ω_m^* and ω_m^{**} converge to $-\infty$. To circumvent this problem, we take an alternative approach based on the comparison between ω_m^* and ω_s^*. The result in this subsection is the following proposition:

Proposition H.1. *There exists $\overline{L} : \mathbb{R}_+ \times (0,1) \to \mathbb{R}_+$ such that when $L > \overline{L}(c,\delta)$, an equilibrium exists. Compared to the single-agent benchmark, $q_m > q_s$, $\omega_m^* > \omega_s^*$ and $\omega_m^{**} > \omega_s^{**}$.*
Moreover, as $\delta \to 1$ and $L \to \infty$ with the relative speed of convergence satisfying $L \ge \overline{L}(c,\delta)$, we have $\omega_m^*, \omega_m^{**} \to -\infty$, $\mathcal{I}_m \to 1$ and $\widetilde{\pi}_m \to \pi^*$.

The proof is in the next subsection that treats two cases separately. Intuitively, in the *regular case* where $\omega_m^* \ge \omega_m^{**}$, one can still apply the ratio condition (3.17) to show that as $\omega_m^* \to -\infty$, the LHS converges to 1 which implies that $\mathcal{I}_m \to 1$. In the *irregular case* where $\omega_m^* < \omega_m^{**}$, the distance between $|\omega_m^* - c|$ and $|\omega_m^{**} + b - c|$ can be strictly larger than b and can explode as $\omega_m^* \to -\infty$. However, since $\omega_m^* > \omega_s^*$ and the informativeness in the single-agent benchmark grows without bound as $L \to \infty$, it places an upper bound on the informativeness of reports in the two-agent scenario.

Since informativeness is entirely contributed by the mechanical types in the irregular case, the value of the aforementioned upper bound converges to 1 as $\mathcal{I}_s \to \infty$. Summing up the two cases together, we know that the agents' reports are arbitrarily uninformative in the limit even when the mechanical types' reports are informative.

H.2.2 Proof of Proposition H.1

We start from establishing the comparisons between the single-agent benchmark and the two-agent scenario when mechanical types' reports can be informative about θ, captured by the two exogenous reporting cutoffs $\overline{\omega}$ and $\underline{\omega}$ with $\overline{\omega} \geq \underline{\omega}$.

Suppose towards a contradiction that $\omega_m^* \leq \omega_s^*$, the expressions for these cutoffs imply:

$$q_m\Big(\delta\Phi(\omega_m^{**}) + (1-\delta)\Phi(\underline{\omega})\Big) \leq q_s.$$

Therefore,

$$\begin{aligned}
&q_m\Psi^{**}\Big(\delta\Phi(\omega_s^*) + (1-\delta)\Phi(\overline{\omega}) - \delta\Phi(\omega_s^{**}) - (1-\delta)\Phi(\underline{\omega})\Big) \\
&\leq q_s\Big(\delta\Phi(\omega_s^*) + (1-\delta)\Phi(\overline{\omega}) - \delta\Phi(\omega_s^{**}) - (1-\delta)\Phi(\underline{\omega})\Big) = 1/L \\
&= q_m\Psi^{**}\Big(\delta\Phi(\omega_m^*) + (1-\delta)\Phi(\overline{\omega}) - \delta\Phi(\omega_m^{**}) - (1-\delta)\Phi(\underline{\omega})\Big)
\end{aligned}$$

or equivalently

$$\Phi(\omega_m^*) - \Phi(\omega_m^{**}) \geq \Phi(\omega_s^*) - \Phi(\omega_s^{**}). \tag{H.12}$$

On the other hand, since $\omega_m^* - \omega_m^{**} < b = \omega_s^* - \omega_s^{**}$ and $\omega_m^* < \omega_s^*$, we have:

$$\Phi(\omega_m^*) - \Phi(\omega_m^{**}) < \Phi(\omega_s^*) - \Phi(\omega_s^{**}). \tag{H.13}$$

which contradicts (H.12). This contradiction implies that $\omega_m^* > \omega_s^*$. Since $\omega_m^* - \omega_m^{**} < b = \omega_s^* - \omega_s^{**}$, we know that $\omega_m^{**} > \omega_s^{**}$. Moreover, $\omega_m^* > \omega_s^*$ implies that $q_m\Psi^{**} > q_s$. That is $1 \geq \Psi^{**} > q_s/q_m$, which implies that $q_m > q_s$.

Next, we establish the informativeness of the agents' reports when there are two agents and δ and L being sufficiently large. First, for every $X \in \mathbb{R}_+$, there exists $\overline{\delta} \in (0,1)$ and $L^* : (\overline{\delta}, 1) \to \mathbb{R}_+$ such that when $\delta > \overline{\delta}$ and $L > L^*(\delta)$, the resulting cutoffs in the single-agent case satisfies:

$$\frac{\delta\Phi(\omega_s^*) + (1-\delta)\Phi(\overline{\omega})}{\delta\Phi(\omega_s^* - b) + (1-\delta)\Phi(\underline{\omega})} > X, \tag{H.14}$$

which implies that

$$\delta\Phi(\omega_s^*) > (1-\delta)\Big(X\Phi(\underline{\omega}) - \Phi(\overline{\omega})\Big). \tag{H.15}$$

Next, we establish an upper bound on the informativeness of reports in the limit of the two-agent case. Consider a two-agent economy under parameter values (L, c, δ) such that $L \geq \overline{L}(c, \delta)$, i.e. monotone-responsive equilibria exist In equilibria where $\omega_m^* \geq \omega_m^{**}$, the expressions for ω_m^* and ω_m^{**} imply that:

$$\frac{|\omega_m^* - c|}{|\omega_m^{**} - c + b|} = \frac{(l^* + 2)\mathcal{I}_m}{l^* + 2\mathcal{I}_m}. \tag{H.16}$$

The LHS converges to 1 as $\omega_m^* \to -\infty$ so the RHS also converges to 1, which implies that $\mathcal{I}_m \to 1$.

In equilibria where $\omega_m^* < \omega_m^{**}$, since $\omega_s^* < \omega_m^*$, we have:

$$\mathcal{I}_m \leq \frac{\delta\Phi(\omega_m^*) + (1-\delta)\Phi(\overline{\omega})}{\delta\Phi(\omega_m^*) + (1-\delta)\Phi(\underline{\omega})} \underbrace{\leq}_{\text{since } \mathcal{I}_m > 1 \text{ and } \omega_m^* > \omega_s^*} \frac{\delta\Phi(\omega_s^*) + (1-\delta)\Phi(\overline{\omega})}{\delta\Phi(\omega_s^*) + (1-\delta)\Phi(\underline{\omega})}$$

$$\leq \frac{(1-\delta)\Big(X\Phi(\underline{\omega}) - \Phi(\overline{\omega})\Big) + (1-\delta)\Phi(\overline{\omega})}{(1-\delta)\Big(X\Phi(\underline{\omega}) - \Phi(\overline{\omega})\Big) + (1-\delta)\Phi(\underline{\omega})} = \frac{X\Phi(\underline{\omega})}{X\Phi(\underline{\omega}) - \Phi(\overline{\omega}) + \Phi(\underline{\omega})} \tag{H.17}$$

which also converges to 1 as $X \to \infty$.

To summarize, since $\omega_m^* \to -\infty$ and $X \to \infty$ as $\delta \to 1$ and $L \to \infty$, we know that the informativeness ratio $\mathcal{I}_m$ converges to 1 no matter whether $\omega_m^* \geq \omega_m^{**}$ or $\omega_m^* < \omega_m^{**}$.