B Technical appendix: Analysis of the enforcement of $N > 2$ agents

In the main text we restrict attention to the case in which the enforcement agency oversees two agents, and has the resources to investigate just one of them. In Section 5 we discuss how our results would change if instead the enforcement agency oversaw $N > 2$ agents, while still possessing the resources to investigate just one agent. In this Appendix we establish several results (omitted from the main text) that are required for this generalization.

Most of our analysis is conducted in terms of the investigation probability function q. As we discussed in the main text, this function can be defined for N agents also. With N agents, the function q is still increasing, and has the same comparative statics with respect to precision h. The only other properties of q used in the analysis are that (A) when $N = 2$, it satisfies relation (1), and (B) it is convex over negative values and concave over positive values (see Lemma 4). To summarize, with respect to (A) we show that for $N > 2$ the equality (1) becomes an analogous inequality, which is still sufficient to imply the existence of a pure-strategy symmetric equilibrium; and with respect to (B), our main results all hold even when q does not possess these curvature properties, provided that a_M is close to 1/2.

B.1 Generalizing equality (1)

On a formal level, our analysis makes repeated use of Lemma 1. This result relies in turn on the inequality

$$p\left(1, a_M\right) - p\left(1, 1\right) \geq p\left(a_M, a_M\right) - p\left(a_M, 1\right). \tag{10}$$

When $N = 2$, equation (1) implies that inequality (10) holds at equality for any investigation policy.

When $N > 2$ more work is needed to establish that (10) is satisfied. Now, an investigation policy defines a function $p : [0, 1]^N \to [0, 1]$, giving the probability that agent 1 is investigated as a function of an N-vector of action choices. Exactly as in the case of $N = 2$, it can be shown that the investigation policy "investigate the highest signal" maximizes $p(a, 0, \ldots, 0)$ and minimizes $p(a, 1, \ldots, 1)$. The proof of this result closely parallels that of Lemma 3 in the current paper, but requires considerable extra notation. The proof is included in an earlier working version of our paper, a copy of which is available upon request.

As such, "investigate the highest signal" is the investigation policy that maximizes the probability that no crime is an equilibrium, and minimizes the probability that severe crime is an equilibrium. Under this policy, for any pair of action choices a and a' the value $p(a', a, \ldots, a)$ depends only on the difference between a' and a. As before, we define a function q by $q(a' - a) = p(a', a, \ldots, a)$. In terms of q, inequality (10) is

$$q(1 - a_M) - q(0) \geq q(0) - q(a_M - 1). \tag{11}$$

Lemma 5 *For any $N \geq 2$, inequality (11) holds.*

Given Lemma 5, the equilibrium characterization results Propositions 3 and 6 are exactly as before.

B.2 Dispensing with the curvature properties of q

To conclude that marginal penalties have a negative effect absent crime waves, we need to establish that

$$\frac{S a_M q(1 - a_M)}{(1 - a_M) q(a_M) + a_M q(0)} < S \tag{12}$$

whenever (7) holds. When $N = 2$, this follows from the concavity of q over positive values. Although numerical simulations suggest that q is concave over positive values for general N, we have been unable to formally establish this. Nonetheless, (12) trivially holds for a_M close to 1/2.

Similarly, to conclude that marginal penalties have a positive effect under crime waves, we must show that

$$\frac{S a_M q(0)}{(1 - a_M) q(0) + q(a_M - 1) a_M} < S. \tag{13}$$

Again, this condition is trivially satisfied provided a_M is in the neighborhood of 1/2, or lower.

These observations together imply that the paper's main result, Proposition 4, holds for a_M in the neighborhood of 1/2.

B.3 The effect of changing N

Finally, in Section 5 we claim that the ratio $q_N(x)/q_N(0)$ is decreasing (respectively, increasing) in the number of agents N if $x<0$ (respectively, $x>0$). That is, as N increases the probability of investigation decreases faster for an agent who commits a lesser crime, holding the actions of other agents fixed. Here, we establish this result:

Lemma 6 *Suppose that the noise term ε is either normally distributed, or has a density function f such that $\frac{f(\varepsilon)}{f(x+\varepsilon)}$ is bounded. Then $q(x)/q(0)$ is a decreasing (increasing) function N when $x<0$ ($x>0$).*

B.4 Mathematical proofs

Proof of Lemma 5: Let ξ be the highest realization of the $N-1$ signals $\varepsilon^2,\ldots,\varepsilon^N$. Let H and h denote the distribution and density function of ξ. Thus for any a,

$$q(a)=\int \Pr\left(a+\frac{\varepsilon}{h}\geq\frac{\xi}{h}\right)f(\varepsilon)\,d\varepsilon=\int G(ha+\varepsilon)\,f(\varepsilon)\,d\varepsilon,$$

and so

$$q'(a)=h\int g(ha+\varepsilon)\,f(\varepsilon)\,d\varepsilon=h\int g(\varepsilon)\,f(\varepsilon-ha)\,d\varepsilon.$$

To establish (11) it suffices to show that $q'(a)\geq q'(-a)$. For this, it suffices to show that

$$g(\varepsilon)\,f(\varepsilon-ha)\geq g(\varepsilon-ha)\,f(\varepsilon),$$

for which in turn it suffices to show that

$$(\ln g)'\geq(\ln f)'.$$

In general, $G(x)=F(x)^{N-1}$ and so $g(x)=(N-1)\,F(x)^{N-2}f(x)$. Since certainly $(\ln F)'>0$, the result follows. ∎

Proof of Lemma 6: For expositional ease we prove the result for $h = 1$. The general case is identical. Clearly $q(0) = 1/N$, while for any $x \in \Re$

$$q(x) = \int_{-\infty}^{\infty} F(x+\varepsilon)^{N-1} f(\varepsilon)\, d\varepsilon.$$

Thus

$$\frac{q(x)}{q(0)} = \int_{-\infty}^{\infty} NF(x+\varepsilon)^{N-1} f(x+\varepsilon) \frac{f(\varepsilon)}{f(x+\varepsilon)} d\varepsilon.$$

Integration by parts gives

$$\frac{q(x)}{q(0)} = \left[F(x+\varepsilon)^N \frac{f(\varepsilon)}{f(x+\varepsilon)} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} F(x+\varepsilon)^N \frac{\partial}{\partial \varepsilon} \left(\frac{f(\varepsilon)}{f(x+\varepsilon)} \right) d\varepsilon.$$

For $x < 0$ the likelihood ratio $\frac{f(\varepsilon)}{f(x+\varepsilon)}$ is decreasing in ε. To evaluate the first term in the expression above, we need to evaluate

$$\lim_{\varepsilon \to -\infty} F(x+\varepsilon)^N \frac{f(\varepsilon)}{f(x+\varepsilon)}.$$

If $\frac{f(\varepsilon)}{f(x+\varepsilon)}$ is bounded above, this term is clearly zero. Otherwise, further conditions are required. L'Hôpital's rule gives

$$\lim_{\varepsilon \to -\infty} F(x+\varepsilon)^N \frac{f(\varepsilon)}{f(x+\varepsilon)} = \lim_{\varepsilon \to -\infty} \frac{f(x+\varepsilon)\, NF(x+\varepsilon)^{N-1}}{\frac{\partial}{\partial \varepsilon}\left(\frac{f(x+\varepsilon)}{f(\varepsilon)} \right)},$$

provided the righthand side exists. When the noise term is distributed normally,

$$\frac{f(x+\varepsilon)}{\frac{\partial}{\partial \varepsilon}\left(\frac{f(x+\varepsilon)}{f(\varepsilon)} \right)} = \frac{\exp\left(-\frac{1}{2\sigma^2} (x+\varepsilon)^2 \right)}{\frac{\partial}{\partial \varepsilon} \exp\left(-\frac{1}{2\sigma^2} (x^2 + 2x\varepsilon) \right)} = \frac{\exp\left(-\frac{1}{2\sigma^2} \left((x+\varepsilon)^2 - \left(x^2 + 2x\varepsilon\right) \right) \right)}{-\frac{x}{\sigma^2}} \to 0 \text{ as } \varepsilon \to -\infty.$$

So provided either $\frac{f(\varepsilon)}{f(x+\varepsilon)}$ is bounded above, or $\lim_{\varepsilon \to -\infty} \frac{f(x+\varepsilon)}{\frac{\partial}{\partial \varepsilon}\left(\frac{f(x+\varepsilon)}{f(\varepsilon)} \right)} = 0$,

$$\frac{q(x)}{q(0)} = \lim_{\varepsilon \to \infty} \frac{f(\varepsilon)}{f(x+\varepsilon)} - \int_{-\infty}^{\infty} F(x+\varepsilon)^N \frac{\partial}{\partial \varepsilon} \left(\frac{f(\varepsilon)}{f(x+\varepsilon)} \right) d\varepsilon.$$

For $x < 0$ the term $\frac{\partial}{\partial \varepsilon}\left(\frac{f(\varepsilon)}{f(x+\varepsilon)} \right)$ is everywhere negative, since f is log concave and so $\ln f(\varepsilon) - \ln f(x+\varepsilon)$ is a decreasing function of ε. It follows that the ratio $q(x)/q(0)$ is decreasing in N for $x < 0$. ∎